

DOCTORAL THESIS

Training professional respeakers to subtitle live events in the UK A participative model for access and inclusion

Moores, Zoe

Award date:
2022

Awarding institution:
University of Roehampton

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Training professional respeakers
to subtitle live events in the UK:
A participative model
for access and inclusion

by

Zoe Moores BA Hons, MA, PGCE, MA

A thesis submitted in partial fulfilment
of the requirements for the degree of PhD

School of Arts

University of Roehampton

2022

The research for this project was submitted for ethics consideration under the reference MCL 15/025 in the Department of Media, Culture & Language and was approved under the procedures of the University of Roehampton's Ethics Committee on 27.01.16.

Abstract

Over the last fifteen years, respeaking, or the production of live subtitles by a person using speech recognition, has become a common technique on television. Whilst in many countries its use has expanded into other areas, this trend has not been seen anywhere near as widely in the UK, where respeaking is often considered a poor relation to other access modalities. This study explores how respeaking can be introduced at unscripted/partially scripted events to complement the access currently provided.

The action research methodology used in this research facilitated close collaboration with users and providers, who shared their expertise while outlining their key access needs. A bespoke training programme was designed to enable professional television respeakers to transfer their skills to this new setting and two rounds of research events allowed their respeaking to be evaluated and feedback to be collected from all involved. The data presented in this thesis is structured around these stages in the research design.

The respeaking for the most part met, and frequently exceeded, the benchmark of 98% accuracy set for live television subtitling; latency was similar to that seen on television. This technical analysis illustrates that respeaking can be considered as a viable way of providing access for d/Deaf, deafened and hard of hearing audience members in this new sector, access that a wider audience was also seen to benefit from. More importantly, support for this new service was voiced among the providers and audience members involved.

The theoretical underpinning of this study, embedded within intersectionality, social and epistemic justice and the social model of accessibility, demanded a deeper consideration of how the linguistic and sensory access respeaking brings must be integrated within the multiple layers of access that any event needs, and frames this thesis. Precisely what access entails will be shaped by the circumstances of each event and those who attend, but individual and collective awareness of and participation in access is vital, especially if broader equity and access are to be truly experienced in society.

Contents

List of figures	15
List of tables	17
List of appendices.....	19
Permissions	22
Acknowledgements.....	24
Stylistic conventions	28
Statement on accessibility	29
Abbreviations.....	30
Chapter 1: Introduction	32
1.1 Living Ζωή/Zoe.....	32
1.2 Identifying myself as researcher	34
1.3 Outline of the research	38
1.4 Chapter Overview.....	41
Chapter 2: Diversity, access and inclusion.....	48
2.1. Norms and labels as markers of identity.....	49
2.1.1 The use of norms and labels in society	49
2.1.2 Discriminatory normative bias	53
2.1.2.1 The impact of labels on individual identity.....	58
2.1.2.2 A non-binary approach	59
2.2 A triple theory approach	61
2.2.1 An intersectional approach.....	63
2.2.2 Social and epistemic justice	71
2.2.3 Social model of accessibility	75
2.2.3.1. From disability to accessibility	76
2.2.3.2 New discipline, new approach	78
2.3 Media accessibility as a bridge between audiovisual translation and accessibility studies	83

2.3.1 The scope of access and translation.....	83
2.3.1.1 Meeting points and divergences between media accessibility and audiovisual translation.....	84
2.3.1.2 The notions of vulnerability, respect and trust.....	87
2.4 Drawing together the theoretical threads.....	89
Chapter 3: Respeaking in a new setting.....	91
3.1 What is respeaking?.....	91
3.2 Working with speech recognition software.....	95
3.2.1 Converting spoken to written English.....	95
3.2.2 Dragon by Nuance.....	96
3.2.3 Tools available to the respeaker.....	99
3.2.3.1 Editing Dragon's vocabulary.....	100
3.2.3.2 House styles.....	100
3.2.3.3 Macros.....	101
3.2.3.4 Live correction.....	102
3.2.3 The effort involved in respeaking.....	103
3.3 Questions of quality and quantity.....	104
3.4 Defining live events in the context of this study.....	107
3.5 Who are subtitles for?.....	109
3.5.1 The traditional audience.....	109
3.5.2 The wider audience.....	113
3.6 Subtitle reception: Factors affecting the speed, content and presentation of subtitles.....	116
3.6.1 The question of speed.....	116
3.6.2 Three measures.....	117
3.6.3 Validity of the six-second rule.....	119
3.6.4 The question of live subtitles.....	122
3.6.4.1 Display.....	124

3.7 Respeaker training	125
3.7.1 Skills for intralingual respeaking.....	125
3.7.2 Skills for a live event subtitler	128
3.8 Joining the dots: Respeaking and accessibility	130
3.9 The research gap this study aims to fill.....	131
Chapter 4: Methodology and research design.....	133
4.1 Combining theory and action: development of a methodology and research design	134
4.1.1 Research questions	134
4.1.2 Action research as a methodology	135
4.1.3 The cyclic nature of action research.....	139
4.1.4 Research design.....	142
4.1.4.1 Cycle One	142
4.1.4.2 Cycles Two and Three.....	146
4.1.4.3 “Messiness” within the action research process.....	148
4.1.5 Rationale for mixed methods	149
4.1.6 Criticisms raised against action research	151
4.2 Stage-by-stage – the selection of research methods and tools.....	152
4.2.1 Focus groups in Cycle One and beyond	152
4.2.2 Observation and as-live respeaking	154
4.2.3 Respeaker training.....	155
4.2.4 Reception study at the research events	156
4.2.4.1 The use of reception studies.....	156
4.2.4.2 Event design: Case study approach	158
4.2.4.3 Data from the reception study.....	159
4.2.4.4 NER analysis.....	160
4.3 Analysis of the findings.....	161
Chapter 5: Engaging with stakeholders.....	162

5.1 Overview of Cycle One: Initial reconnaissance	162
5.2 What is the nature of the access currently provided at live events?	165
5.2.1 Consultation with Stagertext and related organisations	165
5.2.2 Observations at live events	167
5.3 What expectations do the key parties have for a respeaking service at live events?	170
5.3.1 Approach to data collection and analysis	171
5.3.2 Data from the audience.....	172
5.3.3 Views from d/Deaf, deafened and hard of hearing (DH) people	173
5.3.3.1 Development of the focus group sessions and online questionnaire...	173
5.3.3.2 Who were the participants?	175
5.3.3.3 Viewing habits	177
5.3.3.4 Conventions	180
5.3.3.5 Speed and style.....	189
5.3.3.6 Priorities in the different settings.....	194
5.3.3.7 Access at on television and at events: A thematic analysis of the qualitative responses	198
5.3.4 Views from Non-native speakers of English (NNE)	202
5.3.4.1 Development of the focus group sessions	203
5.3.4.2 Who were the participants?	203
5.3.4.3 Viewing habits	204
5.3.4.4 Priorities in the different settings.....	204
5.3.4.5 Comparison of the needs and expectations of the DH and NNE audience groups.....	205
5.3.5 Venues	206
5.3.5.1 Key themes	206
5.3.5.2 Priorities when using respeaking at live events.....	208
5.3.6 Respeakers	208
5.3.6.1 Data from respeakers	208

5.3.6.2 Who were the participants?	210
5.3.6.3 Respeaking training and knowledge	210
5.3.6.4 Working environment.....	211
5.3.6.5 Variations in content and subtitling style	212
5.3.6.6 Respeakers' opinions of respeaking vs STTR	215
5.3.6.7 The move to live events.....	216
5.3.6.8 Priorities for respeaking at events.....	219
5.3.6.9 The priorities of service providers	221
5.4 What is it like to subtitle a live event?	222
5.5 Review of Cycle One.....	223
Chapter 6: Respeaker training and event programme.....	225
6.1. The role of respeaking training and research events	225
6.2 Developing the training and events programme	227
6.2.1 Research Questions	227
6.2.2 Defining quality in the context of respeaking at live events	229
6.2.3 Principles behind the respeaker training and events	231
6.3 Who were the respeakers?.....	232
6.3.1 Selection Process	232
6.3.2 Respeaker Profiles	233
6.4 Equipment.....	234
6.4.1 Computer and speech recognition software.....	235
6.4.2 Microphones.....	235
6.4.3 Audio feed	236
6.4.4 Presentation software.....	236
6.4.5 Accessories	237
6.5 Respeaker training Round One: Design	238
6.5.1 Module overview	238
6.5.2 Training logistics	239

6.5.3 Theoretical knowledge and understanding.....	240
6.5.3.1 Module one – Project overview.....	240
6.5.3.2 Module two – Expectations at live events	241
6.5.3.3 Module five - Respeaking style.....	241
6.5.4 Practical knowledge and understanding on day one	242
6.5.4.1 Module three – Equipment.....	242
6.5.4.2 Module four – Practicalities of live events	242
6.5.4.3 Module six - Practice	243
6.5.5 Personal effectiveness – Day one and beyond	243
6.5.6 Follow-up training	244
6.6 Respeaker training Round One: Analysis	244
6.6.1 Professional approach	244
6.6.2 Respeakers’ needs	245
6.6.2.1 Core needs.....	245
6.6.2.2 Higher level needs.....	247
6.6.3 Respeaking style	249
6.6.3.1 Natural respeaking Style	249
6.6.3.2 Testing verbatim, standard and fully edited styles	250
6.7 Live event logistics	252
6.7.1 Sending out respoken text	252
6.7.2 Handing over	253
6.7.3 Cueing scripts	253
6.7.4 Saving transcripts	253
6.7.5 Dragon vocabulary and macros	254
6.7.6 Liaison with presenters	254
6.8 Research events: Design and purpose.....	255
6.8.1 Range of events.....	255
6.8.2 Simulating ‘real’ events	259
6.8.3 Early attempts at integrating access into the event	260

6.8.3.1	Publicity and recruiting audience members	260
6.8.3.2	Venue and presenter preparation	260
6.8.3.3	Availability and entry.....	261
6.8.4	Respeaker allocation	262
6.8.5	Event organisation	263
6.9	Technology used at the events.....	264
6.10	Style guidelines for live events	266
6.11	Updates to the training programme	271
6.11.1	Respeaker training Round Two.....	271
6.11.2	Future offerings of the training	272
6.11.2.1	Refining modules from the initial round of respeaker training.....	274
6.11.2.2	New modules.....	274
6.11.2.3	Innovativeness of the design	275
6.12	Resources for respeaking at live events	276
Chapter 7:	Reception study at the live events	278
7.1	A recap - event design and procedures.....	279
7.2	Data from those attending the events	279
7.2.1	Format	279
7.2.2	Working with the scales	281
7.3	Overview of the audience	285
7.3.1	Who attended the events?	285
7.3.2	Overview of the audience feedback	285
7.4	Round One.....	289
7.4.1	Event 1	289
7.4.1.1	Event outline	289
7.4.1.2	The space	290
7.4.1.3	Preparation and lead-up to the event.....	292
7.4.1.4	Event proper.....	293

7.4.1.5 Audience responses	294
7.4.1.6 Wider reflections on the event	295
7.5.2 Event 2	297
7.4.2.1 Event outline	297
7.4.2.2 The space	298
7.4.2.3 Preparation and lead up to the event.....	299
7.4.2.4 Event proper.....	299
7.4.2.5 Wider reflections on the event	302
7.4.3 Event 3	303
7.4.3.1 Event outline	303
7.4.3.2 The space	303
7.4.3.3 Preparation and lead up to the event.....	305
7.4.3.4 The event	305
7.4.3.5 Audience response.....	306
7.4.3.6. Wider reflections on the event	307
7.4.4 Event 4	308
7.4.4.1 Event outline	308
7.4.4.2 The space	308
7.4.4.3 Preparation and lead up to the event.....	309
7.4.4.4 Event proper.....	310
7.4.4.5 Audience response.....	310
7.4.4.6 Wider reflections on the event	311
7.5 Round Two.....	312
7.5.1 Events 5 and 6.....	312
7.5.1.1 Event outline	312
7.5.1.2 The spaces.....	313
7.5.1.3 Preparation and lead up to the events	314
7.5.1.4 Event proper.....	314
7.5.1.5 Audience responses.....	315

7.5.1.6 Wider reflections on the events	316
7.5.2 Event 7	317
7.5.2.1 Event outline and space	317
7.5.2.2 Preparation and lead up to the event.....	318
7.5.2.3 Event proper.....	318
7.5.2.4 Audience response.....	319
7.5.2.5 Wider reflections on the event	319
7.5.3 Event 8	319
7.5.3.1 Event outline	319
7.5.3.2 The space	320
7.5.3.3 Preparation and lead up to the event.....	320
7.5.3.4 Event proper.....	321
7.5.3.5 Audience response.....	321
7.5.3.6 Wider reflections on the event	321
7.6 Post-event focus groups – a thematic analysis.....	322
7.7 Presenter training.....	326
7.7.1 What would a training programme for presenters look like?	326
7.7.1.1 Know your audience.....	327
7.7.1.2 Being respoken	327
7.7.1.3 Communicating with the respeakers.....	328
7.7.1.4 How to handle different types of audio and visual content	329
7.7.1.5 Preparing for the event.....	330
7.7.1.6 Access for all = training for all.....	331
7.8 From process to product.....	332
Chapter 8: Analysis of respeaking at live events.....	334
8.1 Introduction	334
8.2 The RLE Corpus	335
8.3 Models for analysing respeaking quality	337

8.3.1 The NER model as a preferential model for assessing live subtitles	338
8.3.1.1 Requirements for a model	338
8.3.1.2 Models which preceded the NER.....	339
8.3.2 The NER workflow	341
8.3.3 NER Scoring.....	346
8.4 NER Analysis: Overview.....	348
8.4.1 Accuracy at the research events.....	349
8.4.2 Comparison between respeaking corpora.....	350
8.4.3 Latency.....	355
8.5 NER Analysis: Case Studies	357
8.5.1 Event 6: Post-screening discussion	358
8.5.2 Event 7: Museum tour.....	362
8.6 NERLE: The NER for Live Events	366
8.6.1 The need for the NERLE.....	366
8.6.2 Determining what content should be prioritised.....	367
8.6.3 Formal event content	370
8.6.3.1 Reversed asides.....	370
8.6.3.2 Vocabulary	372
8.6.4 Functional and problem-solving communication.....	372
8.6.5 Additional information	373
8.6.5.1 Speaker identification	373
8.6.5.2 How people speak.....	374
8.6.5.3 Event announcements.....	377
8.6.5.4 Visual information.....	379
8.6.6 Shared responsibility for errors	379
8.7 How do the NER analysis and audience feedback compare?.....	380
8.8 Reflections	383
Chapter 9: Conclusions, discussions and further work.....	384

9.1 Where are we now?	384
9.2 The impact of this study.....	385
9.3 Model for participatory engagement	387
9.4 A starting point for future research	391
9.4.1 Extending the range of the events.....	391
9.4.2 Wider audience research	392
9.4.3 Respeaker training.....	393
9.4.4 Process research.....	394
9.4.5 Links with other access modalities	394
9.4.6 The question of quality.....	395
9.4.7 Further exploration of the NERLE and subtitle introductions	396
9.5 Wider reflections	396
9.5.1 The spread of respeaking	396
9.5.2 Awareness of accessibility	400
9.5.3 The future of translation studies and accessibility studies	401
9.5.4 Be the change that is needed	401
Glossary	405
References	411
Filmography for the focus groups.....	456

List of figures

Fig. i: Basic Language Guide, adapted from Graeae Theatre Company, 2016.....	28
Fig. 2.1: Sign found on a toilet door (All Sign Images, 2021)	52
Fig. 2.2: Triple theory approach.....	62
Fig. 2.3: De Vries' Multidimensional Prism of Intersectionality (2015: 11).....	67
Fig. 2.4: Intersectionality web from Atewologun and Mahalingam (2018: 163), adapted from (Morgan, 1996)	70
Fig. 2.5: International Symbol of Access (<i>International Symbol of Access</i> , 2012)	75
Fig. 2.6: The three accounts of media accessibility (Greco, 2019a: 19)	85
Fig. 2.7: The formation of the field of accessibility studies (Greco, 2019a: 22)	86
Fig. 3.1: The ILSA (Interlingual Live Subtitling for Access) project's terminological map of translational activity	94
Fig. 3.2: Recognition error captured on BBC News, adapted from the screenshot cited in Jivanda, 2014.....	96
Fig. 3.3: Familiar sounds audiogram (KDH Research & Communication, 2012)	111
Fig. 4.1: Elliot's action research model, entitled 'What is action research?' (Elliot, 1991: 71 cited in Koshy <i>et al.</i> , 2010: 7).....	140
Fig. 4.2: O'Leary's cycles of research, (Koshy <i>et al.</i> , 2010: 8)	141
Fig. 4.3: Interweaving structure of focus groups and respeaker training (Moores, 2020a: 182)	145
Fig. 5.1: Respeakers and BSL interpreter at the second DH focus group session	173
Fig. 5.2: B1: Have you ever used subtitles/captions to access the following?	178
Fig. 5.3: Explaining verbatim, standard and edited subtitles, with extracts from O'Neill, 2016	191
Fig. 6.1: SmartMics in use.....	235
Fig. 6.2: Respeaking kit.....	237
Fig. 6.3: Doubled letters in the respoken output	247
Fig. 6.4: Technical set-up at stationary events	265

Fig. 6.5: Technical set-up at mobile events	266
Fig. 6.6: Subtitle position at event 2.....	269
Fig. 6.7: Examples of personalised tablet displays	270
Fig. 7.1: The cast of Quick Fix Theatre in The Imposter (Quick Fix Theatre, 2016) ..	289
Fig. 7.2: Quick Fix performing Alice in Wonderland (Quick Fix Theatre, 2017)	290
Fig. 7.3: Layout of event 1	291
Fig. 7.4: Key for the event layouts	291
Fig. 7.5: Layout of event 2.....	298
Fig. 7.6: Layout of event 3.....	304
Fig. 7.7: Layout of event 4.....	309
Fig. 7.8: Layout of event 5.....	313
Fig. 7.9: Layout of event 6.....	313
Fig. 7.10: Layout of event 7.....	317
Fig. 7.11: Layout of event 8.....	320
Fig. 8.1: WER calculation, adapted from Romero-Fresco, 2011: 151	339
Fig. 8.3: Calculating word correctness from deletion and substitution errors	340
Fig. 8.4: Calculating word correctness from deletion, substitution and insertion errors	340
Fig. 8.5: NER assessment flowchart (Moores and Romero-Fresco, 2019)	342
Fig. 8.6: A comparison of NER and WER scoring, adapted from Romero-Fresco, 2016: 58	347
Fig. 8.7: Interaction on TV and live events.....	366
Fig. 8.8: NERLE assessment flowchart (v3)	368
Fig. 8.9: Explanation of correct editions in the NERLE	380
Fig. 9.1: Model for participatory engagement	388
Fig. 9.2: Broader application of the model for participatory engagement	389
Fig. 9.3: Anti-Brexit placards	399

List of tables

Table 3.1: Speech rates on BBC programmes, based on Romero-Fresco, 2010:	177
.....	118
Table 3.2: Viewing speed and distribution of gaze between subtitles and images (Romero-Fresco, 2015: 338).....	121
Table 4.1: Research Design.....	143
Table 5.1: Cycle One: Initial reconnaissance.....	162
Table 5.2: Gantt chart of research activities	164
Table 5.3: Topics covered in the DH focus groups and online questionnaire.....	174
Table 5.4: Difficulties in using English and accessing content without subtitles/captions	179
Table 5.5: Audience preferences on subtitling/captioning conventions.....	181
Table 5.6: Audience opinions on the speed of subtitles/captions.....	190
Table 5.7: How verbatim, standard and edited subtitles are created	192
Table 5.8: Preferences for verbatim, standard and edited subtitles	192
Table 5.9: Preferences for verbatim, standard and edited subtitles by group.....	193
Table 5.10: Priorities for live subtitles on television	195
Table 5.11: Priorities for captions at the theatre	196
Table 5.12: Priorities for subtitles at live events.....	197
Table 5.13: Topics covered in the respeaker questionnaire.....	209
Table 5.14: Definition of verbatim, edited and respeaking styles for respeakers.....	214
Table 5.15: Priorities for respeaking on television and at live events.....	220
Table 5.16: Content for the respeaker training programme	221
Table 6.1: Cycle Two & Three Respeaker training and research events	228
Table 6.2: Module overview shared with the respeakers	238
Table 6.3: Event overview	256
Table 6.4: Conference in Vigo.....	258
Table 6.5: Respeaker allocation	263

Table 6.6: Suggested conventions for live events, compared to other forms of subtitles/captions	267
Table 6.7: Revised respeaker training programme	273
Table 7.1: Likert scales from the audience questionnaire	282
Table 7.2: Overview of the audience scores at each event by group	286
Table 7.3: Events ranked by audience preference.....	287
Table 7.4: Events with matching DH combined and total audience scores	287
Table 8.1: Key features of each event	337
Table 8.2: The 10-point conversion scale for NER accuracy (Romero-Fresco and Pöchhacker, 2018: 160) alongside the traditional accuracy ratings.....	348
Table 8.3: RLE Accuracy data (including software-specific errors)	349
Table 8.4: Revised accuracy data (RLE*) (excluding software-specific errors)	349
Table 8.5: RLE and RLE* accuracy data compared	350
Table 8.6: Summary of results for the RLE, the RLE*, Ofcom and LiRICS pilot	351
Table 8.7: Total edition and recognition errors	351
Table 8.8: RLE: Total number and percentage of edition and recognition errors	352
Table 8.9: RLE*: Total number and percentage of edition and recognition errors.....	353
Table 8.10: Total serious, standard and minor errors	353
Table 8.11: Total serious, standard and minor errors categorised as edition or recognition errors.....	354
Table 8.12: Latency at each event	356
Table 8.13: Speed, reduction rate and lexical density at each event	358
Table 8.14: Examples of errors at event 6.....	360
Table 8.15: Example of software-specific errors seen at mobile events.....	364
Table 8.16: Total audience and combined DH scores alongside data from the NER analysis	381
Table 8.17: The value of low, medium and high audience, NER and latency scores	381

List of appendices

Ethics

4.1 Sample consent form

Focus group resources

5.1.1 Focus group session plan

5.1.2 Posters of subtitle conventions

5.1.3 Sounds_Back In Time

5.1.4 Sounds_Beowulf

5.1.5 Background music_Beowulf

5.1.6 Songs_Back In Time_I like it

5.1.7 Songs_Back In Time_Space Oddity

5.1.8 Verbatim_Sunday Brunch

5.1.9 Standard_Sunday Brunch

5.1.10 Edited_Sunday Brunch

5.1.11 Comparison of the Sunday Brunch clips

5.1.12 Question Time transcript

Data collected from the focus groups

5.2 Guide to data analysis

5.3.1 DH Take home questionnaire

5.3.2.1 DH Take home responses

5.3.2.2 DH Take home responses cross-tabulated

5.4.1 DH Online questionnaire

5.4.1.1 Verbatim example

5.4.1.2 Standard example

5.4.1.3 Edited example

5.4.2.1 DH Online responses

5.4.2.2 Sample of DH online responses cross-tabulated

5.5.1 NNE questionnaire

5.5.2 NNE responses

5.6.1 Key themes in the respeaker interviews

5.6.2 Respeaker online questionnaire

5.6.3 Respeaker responses

5.7 Key themes in the venue interviews

5.8.1 Respeakers at the focus groups - questionnaire

5.8.2 Respeakers at the focus groups - responses

Respeaker training

6.1.1 Course booklet for RLE training

6.1.2 RLE training slides

6.1.3 Transcripts of clips to compare respeaking styles

6.1.4 Practice materials

6.1.5 Pre-training questionnaire

6.1.6 Mid-training questionnaire

6.1.7 Post-training questionnaire

Guidelines for using respeaking at live events

6.2.1 Original guidelines for presenters

6.2.2 Revised guidelines for presenters: Information for speakers

6.3 On-the-day: Equipment and technical checks

Data collection at the research events

7.1.1 Respeaker pre-event questionnaire

7.1.2 Respeaker post-event questionnaire

7.2.1 Venue pre-questionnaire

7.2.2 Venue post-questionnaire

7.3.1 Information about the event - questions for presenters

- 7.3.2 Presenter questionnaire
- 7.4.1 Audience demographic questionnaire
- 7.4.2 Audience feedback questionnaire
- 7.4.3 Audience response
- 7.5 Additional notes on the SPSS analysis
- 7.6 Audience data from the events
- 7.7 Qualitative feedback from the research events
- 7.8 Key themes in the post-event focus groups

NER-LE analysis

- 8.1 Guide to the NER-LE data analysis
- 8.2 NER-LE analysis
 - 8.3.1 Event 1 recording
 - 8.3.2 Event 2 recording
 - 8.3.3 Event 3 recording
 - 8.3.4 Event 4 recording
 - 8.3.5 Event 6 recording
 - 8.3.6 Event 7 recording
 - 8.3.7 Event 8 recording
- 8.4 NERLE scenarios

Making live events accessible using respeaking

- 9.1 Questions and considerations when making live events accessible using respeaking
- 9.2 Model for participatory engagement

Permissions

Figure 2.3 on p.67 'Figure 1. Multidimensional prism' is from Kylan Mattias de Vries, (2015) Transgender People Of Color At The Center: Conceptualizing A New Intersectional Model, *Ethnicities*, 15(1), pp. 3–27. Reprinted with the permission of SAGE Publications, Ltd. doi:10.1177/1468796814547058.

Figure 2.4 on p.70, Figure 7.1 Intersectional Identity web, Intersectionality as a methodological tool by Atewologun, D. & Mahalingam, R. © Lize A.E. Booyesen, Regine Bendl and Judith K. Pringle 2018. Reproduced with the permission of The Licensor through PLSclear.

Figure 2.6 on p.85 and Figure 2.7 on p.86 are reprinted by permission from Springer Nature Customer Service Centre GmbH : Springer International Publishing, Accessibility Studies: [Abuses, Misuses And The Method Of Poietic Design](#) by Gian Maria Greco, in Stephanidis, C. (eds) HCI International 2019 – Late Breaking Papers. HCII 2019. Lecture Notes in Computer Science, vol 11786. Springer, Cham © 2019

Figure 3.1 on p.94 is reprinted with the kind permission of Pablo Romero-Fresco and the ILSA Project.

Figure 3.3 on p.111 is reprinted courtesy of KDHC Research & Communication (KDHC). KDHC developed this image as a part of The Cochlear Implant School Toolkit, with U.S. federal funds from the National Institute on Deafness and Other Communication Disorders, part of the National Institutes of Health (grant # 4R44DC009357-02). The Cochlear Implant School Toolkit contents are solely the responsibility of KDHC and do not necessarily represent the official view of the National Institute on Deafness and Other Communication Disorders.

Figure 4.1 on p.140 'Figure 1.2 Elliot's action research model' is from Elizabeth Koshy, Valsa Koshy and Heather Waterman, (2010) *Action Research In Healthcare* p.7, SAGE Publications, Ltd. It was adapted from 'Fig 6.2' from John Elliot, (1991) *Action Research For Educational Change* © 1991 Buckingham: Open University Press. Reproduced with the permission of McGraw Hill.

Figure 4.2 on p.141 'Figure 1.3 O'Leary's cycles of research' is from Elizabeth Koshy, Valsa Koshy and Heather Waterman, (2010) *Action Research In Healthcare* p.8. Reprinted with the permission of SAGE Publications, Ltd.

Figure 7.1 on p.289 and figure 7.2 on p.290 are reprinted with the kind permission of Ben Wilson on behalf of Quick Fix Theatre.

Figure 8.5 on p.342 is printed with the kind permission of the co-author Pablo Romero-Fresco.

Some passages in this thesis are extracts from Moores, Z.(2020) 'Fostering access for all through respeaking at live events', *Journal of Specialised Translation* 33, pp.176-211, available at https://www.iostrans.org/issue33/art_moores.php

Acknowledgements

This PhD journey has been an eventful one. There have been periods of joy and excitement, fun and laughter, but also many bumps and moments of queasiness and fear along the way. Carrying out the research and writing this thesis have been my job and responsibility, but they were not tasks I could have completed alone. I have a great many people to thank.

Firstly, my immense gratitude to my supervisory team - Pablo Romero-Fresco, Lucile Desblache and Annabelle Mooney. You have been there for me throughout, giving me support and friendship, advice and guidance and continually challenging me to grow and explore as the study unfolded. This has meant more to me than you can know.

This research would not have been possible without the funding I received from AHRC TECHNE and I am indebted to you. A particular thank you to Jane Gawthrop and Carol Hughes for your hard work and dedication to us all. The training and support offered by TECHNE throughout have been greatly appreciated.

Stagetext, the input I received from your whole team, old and new, has been invaluable and helped me avoid so many (now obvious) mistakes. Particular thanks to Melanie Sharpe, Rosie Holmes, Deepa Shastri, Bethan Way, Emily Gallagher, Richard France, Liz Hilder, Claire McIntyre, Courtney Rudge and your speech-to-text team.

Projects involving action research often seek to make a real change for people and this one is no exception. In order to find out whether respeaking could be used at live events, I needed to talk to users and providers and to carry out a series of research events. A massive thank you to everyone involved in the running of these events – the venue teams, respeakers, presenters, interpreters, people helping on site and everyone who came to the events and volunteered their valuable time - without you, these events, and this research, could not have taken place. Thank you, too, to everyone who took part in the focus groups and replied to the questionnaires that were sent out at the start of the study. It is your voices that I have tried to share in this work – I hope I have done you

justice. A particular thank you to Dawn Jones, Maria Bayley and Mike Theobald. I have thoroughly enjoyed all our discussions.

Thank you to everyone who provided access at these events through live subtitles and BSL interpretation. Luke Barrett, Jason Cook, Laura Dunbabin, Sarah Huntington and Dee Johns, your commitment and dedication to this research has been amazing! And a big thank you to Alex Potterill, Lee Worth and Evan Dorrestein for getting us on our way. Thank you to all the teams at the venues who lent me their space to hold events and to test out respeaking; in particular to Nigel Greenaway and Emily Boulting at the Riverhouse Barn; Kiran Dhanoya and John McKnight at the BFI St Stephen Street, and Professor Richard Paterson; Orla O'Donnell at the Wellcome Collection; Louise Gardner at Watershed; Carmen Slijpen at Depot; Meg Parnell at Manchester Art Gallery and the various teams at the University of Roehampton.

Thank you to Quick Fix Theatre (and Ben Wilson in particular), Pete Allen, Pete Middleton and Jo-Jo Ellison at Archer's Mark, Julie McNamara at VitalXposure, Gillian Mciver, Elissavet Ntoulia, David Ellington, Sian Norris, Sarah Haynes, Pascale Maroney, Josh Branson, Julia Marwood and to Paul Connery for taking part in the events so that we had an opportunity to test the respeaking in action.

Thank you to the Lenna Cumberbatch, David Chapman and Liam Helm at the Royal Society, Jane Samuels at the Natural History Museum, Barry Ginley at the V&A and to the teams at Ai-Media, Deluxe and Ericsson for sharing your expertise; to the NADP (with a particular shout out to Simon Pearse and Lidia Best), Gillian Rollason and Johanna Taylor at Action on Hearing Loss, Gail Adams at the BSL Café and to all the other organisations who helped to share and spread the word about my research.

Thank you to my anonymous JoSTrans reviewers – your suggestions have undoubtedly improved my final thesis. Jen Ferguson, thank you for your work on the transcriptions. Mike O'Driscoll and Louise Fryer, thank you for your statistical input and analysis. Karen Parker, thank you for your patience during our BSL lessons. Sam Farrell and Chris Milton, thank you for your technical advice.

Finding the right working environment can also be a challenge. Thank you to London Writers' Salon for the space you create to write and, Shepperton Open Water Swim, thank you for offering such an amazing place to take a screen break!

Exploring this PhD has also been an opportunity for me to become part of the research community at Roehampton and beyond – my family has grown. In the Department of Media, Culture and Language, a particular thank you to Ahmed Abbasi, Andy Walker, Bruno Levasseur, Caroline Bainbridge, Cécile Renaud, Deborah Jermyn, Dionysios Kapsaskis, Elif Grant, Eva Eppler, Fjoralba Miraka, Gillian Mciver, Inma Pedregosa, Irene Artegiani, Jingue Zhang, Josh Branson, Judith Rifesser Nussbaumer, Lana Tubaltseva, Mandie Iveson, Michael Witt, Miguel Bernal-Merino, Sabela Melchor Couto, Serenella Massida, Stacey Abbott, Theodora Thomadaki, the late Tope Omoniyi, and William Brown. Clive Smith and Susie Hyde, thank you for all the technical support and loans of recording equipment. And a thank you to all my international research family, especially to Alessandra Rizzo, Aline Remael, Andrew Lambourne, Anja Jankowska, Anna Matamala, Agnieszka Szarkowska, Annalisa Sandrelli, Carlo Eugeni, Dan McIntyre, Daniela Eichmeyer, David Orrego Carmona, Elena di Giovanni, Federico Spoletti, Franz Pöchhacker, Gion Linder, Isabelle Robert, James Ward, Jan-Louis Kruger, Jan Pedersen, Jonathan Evans, Jorge Díaz-Cintas, Kim Starr, Lindsay Bywood, Lotta Lagerman, Łucja Biel, Łukasz Dukta, Monika Szczygielska, Nazaret Fresno, Pilar Orero, Rocío Bernabé, Sabine Braun, Sara Bonjoch, Sharon Black, Soledad Zarate, Sylvain Caschelin, Tomasz Korybski, Ulf Norberg and Veerle Haverhals. I hope we can meet again in person soon.

Elena Davitti, Hayley Dawson, Anna-Stiina Wallinheimo and Louise Fryer, it has been and is a pleasure to work with you on the SMART project. You have been a regular point of contact as we have worked from home and our discussions have been invaluable and stimulated my thinking as I write. Hayley, a particular thank you for second marking the NER samples and, Louise, for reading the final draft. Gian Maria Greco, I have enjoyed all the discussions we've shared and genuinely appreciated your help and feedback on my work as it progressed.

Thank you, too, to all my family and friends who have endured or enjoyed discussions about subtitling and access, and who have sat with me until the final credits rolled off the screen at the cinema over the last few years to see if the subtitlers and access professionals were mentioned. Janet Kohn, you won't get to see me graduate in a floppy hat, but your being at my MA graduation made it so special. Lisa Pitts, Marion & John Hadjikyriacos, Gail Adams, Claire Whiteley, Helen Earis, Jim Hurst, Barbara Schweinzer, H el ene Abiraad and Sumi, thank you all. Shanon Ferguson, you were there from the start and this would have been a very different journey without you; thank you for everything. Cassius Campbell, thank you for listening. Andrea Netscher, Jonathan Clifford, Brig MacLachlan, Martin Loxton, Diana Hanbury, Ros Langford and Renata Drtina, our time together online and in person has been a great support. Doug Duncan and Catherine Pawasarat, I take my inspiration from you.

And, Kate Dangerfield - this has been our journey! I am so grateful to have had a companion during the long evenings in the labs, at conferences, on road trips and, more recently, in the digital world. Thank you for the laughter, friendship, understanding and deep discussions!

May these efforts not have been in vain - may the future be a more equitable and accessible one.

Stylistic conventions

1. 'They', rather than 'he/she', is used as an inclusive general form.
2. Where possible, language relating to disability is phrased from the social model perspective, whereby "people are disabled by barriers within society, rather than being 'victims' of their impairments or conditions". Graeae's Guide for journalists has been used as a point of reference for this (Graeae Theatre Company, 2016) and is the main source of the basic language guide in figure one. On occasion, for example in quotations and to make a point, language from the opposite perspective is required.

✓	X
Disabled	handicapped, cripple, invalid
disabled people	the disabled, people with disabilities
has... (an impairment)	suffers from..., victim of...
non-disabled	able-bodied, normal, healthy
learning disabled	mentally disabled, retarded, backward
wheelchair user	wheelchair bound, confined to a wheelchair, in a wheelchair
deaf	the deaf
Deaf	the Deaf
Deaf sign language user, BSL user	deaf and dumb, deaf mute
blind or partially sighted people, visually impaired people (VIP)	the blind

Fig. i: Basic Language Guide, adapted from Graeae Theatre Company, 2016

Statement on accessibility

If you would like to read this thesis in a more accessible format, or to view the associated appendices, please contact me at zoe-moores@outlook.com so that we can discuss this further.

Abbreviations

AD - Audio description

AI - Audio introduction

AOHL - Action on Hearing Loss, now called RNID

AS - Accessibility studies

AVT - Audiovisual translation

BSL - British Sign Language

CE - correct edition, a feature of the NER model

CPS - Characters per second

DAC - Designated access co-ordinator

DH - d/Deaf, deafened and hard of hearing people

DNS - DragonNaturallySpeaking

DPI - Dragon Professional Individual

EASIT - Easy Access for Social Inclusion project

EEG - Electroencephalograms

GM - General Manager

HMM - Hidden Markov Modelling

ILSA - Interlingual Live Subtitling for Access

IRA - Idea rendition assessment

L2 - Second language

LTA - Live Text Access

LTM - Live Talks Manager

MA - Media accessibility

MAQ - Media accessibility quality

NADP - National Association of Deafened People

NER - NER model

NERLE - NER for Live Events

NNE - Non-native speakers of English

OSR - Original speech rate

RCP - Respeaking crisis point

RLE - Respeaking at live events, the public-facing title of this project; a shorthand form for the scores in Chapter Eight that are inclusive of software-specific errors

RLE* - a shorthand form in Chapter Eight for the scores that are exclusive of software-specific errors

RNID - Royal National Institute for Deaf People, formerly AOHL.

RSR - Respeaker's speech rate

SDH - Traditionally written, 'subtitling for the deaf and hard of hearing'; I prefer it to be read as 'subtitling for d/Deaf, deafened and hard of hearing people' (see Fig. i)

SI - Simultaneous interpreting

SLI - Sign language interpreting/interpreter

SMART - Shaping Multilingual Access through Respeaking Technology

SR - Speech recognition

STT - Speech-to-text

STTI - Speech-to-text interpreting

STTR - Speech-to-text reporter/reporting

TPM - Theatre programme manager

TS - Translation studies

VEM - Visitor Experience Manager

WER - Word error rate

WIRA - Weighted Idea Rendition Assessment

WPM - Words per minute

Chapter 1: Introduction

“Identity is both about similarity and difference,
but it is also about a state of becoming,
always moving towards a future self
that is made up of our past and present experiences.”

Hill, 2018: 18¹

1.1 Living Ζωή/Zoe²

I do not see myself as an academic who sits in an ivory tower, separate from the world that I am researching. Rather, I am as much part of the world as my research is. I have been shaped by this study and I am as much a creation of it, as it is something I have created³. The process of writing a PhD, and the learning and (self-)exploration that it has entailed has become part of my own story, journey and self-identity.

Languages, teaching and travel have featured strongly in my life from an early age. In some ways, a PhD in translation studies can be seen as a natural progression for a person with an academic approach to life. My undergraduate degree was in Classics and Modern Languages, French and Latin, and was followed with an MA in Linguistics after a year working as a Learning Support Assistant in a secondary school. I then taught English as a Foreign Language in Japan, before returning to the UK to train as a primary school teacher. Awareness of special educational needs in the classroom was always a focus of mine, and being creative when planning lessons was a joy. When I realised that primary school teaching was no longer for me, I retrained as an audiovisual translator through the Masters course at the University of Roehampton, where I returned for my PhD. Upon discovering audiovisual translation (AVT), I was immediately drawn to media accessibility as I saw it as a way to integrate advocacy and inclusion, which had become interests of mine, into work with languages in a way which allowed me to be creative. Whilst completing my MA, I began working at Ericsson as an accessibility subtitler and

¹ In these lines, Hill captures the essence of Hall’s 1996 discussion of Who needs ‘identity’? (in Hall and du Gay, 1996: 1-17).

² In Greek, Ζωή (Zoe) means life.

³ In this thesis, I use ‘study’ to refer to my doctoral research project as a whole. An outline of the research design, and steps involved within it, can be found in section 4.1.4 below.

respeaker and used my dissertation to explore accuracy in respeaking, as I developed this skill professionally myself. This doctoral study into how to introduce respeaking at live events is self-designed and is a culmination of my previous professional work, combining my experience of training and subtitling in the context of access and inclusion. Yet, this study goes deeper. It not only considers respeaking as a form of access at live events, it questions the broader provision of access within society. Whilst it might seem that in the 21st century, in the UK, there should already be good access, all too often this is not the case. A degree of access provision is now incorporated into products, services, architecture and urban planning, yet for many people, it is insufficient and more specific access is essential (United Nations, 2016: 6-7; Hunt and Guardian Readers, 2017). Gaining this access often requires identifying as belonging to a certain group recognised as having an access need. Take, for example, the idea of a wheelchair user. Once a provider has the image of a wheelchair user in mind, they may envisage that staircases and steps will be problematic, that a lift will be needed. Even then, more subtle, yet related requirements, may be missed. However, not all access needs are this visible. Many people have been discriminated against because onlookers judge that they do not deserve to use the access provided – for example, seemingly non-disabled people using a disabled parking slot or toilet (Al-Othman, 2015; McPolin, 2019). Without some kind of visual element to latch onto, it can be more difficult for people to comprehend the need for alternate forms of access, and consequently to provide them, especially if not explicitly requested. Despite being a visual form of access itself, subtitling appeared as a provision for aspects of our identity that are less visible: traditional subtitle users either speak a different language to the one spoken in a film or programme's audio, or are d/Deaf, deafened or hard of hearing (DH) people requiring sensorial access. In both cases, the need being met is not immediately visible. Rather, on meeting that person, it is likely that normative assumptions are applied, and the onlooker will assume that they either understand or hear the original audio, depending on the type of access in question. This deeper level of investigation into and consideration of access and the questions it raises about the expression of identity, normative assumptions and the ability of people

to access the rights and justice they are entitled to in our modern world is as much a natural progression of my life story, as the professional path I have followed. For this reason, I open my PhD with a Bildungsroman-esque reflection, illustrating the personal journey of discovery, growth and understanding I have followed, and am still following, alongside this academic study. Carrying out this research has allowed me to better understand the intersections within my own identity, how I see myself in the world and how my research and my approaches fit into the deeper philosophical and sociological theories of how the world is shaped and acts.

1.2 Identifying myself as researcher

Who is Zoe? How do I appear? What labels do I take on and which am I given?

By appearance, I am a white woman, straight, cisgender, in my 40s, with no obvious disability or physical needs. Begin to speak to me, and you will hear that I am English, and most likely judge that I am middle class, from London. You might guess my religion as Christian, knowing the norms for London in the late 1970s, and perhaps take a guess at where I stand politically. All in all, my appearance and voice would suggest I probably 'fitted in well' growing up and still do, and that I appear conventional.

However, this image is based on assumptions – your assumptions of me and mine of how you will identify and think about me. And such assumptions are problematic, as are the norms and labels around which they are based, especially when they begin to function in a discriminatory way (an idea which is explored further in Chapter Two). To what extent does any label capture a person's identity? And how closely does appearance reflect internal identity? In the course of this PhD, and in particular in Chapter Two, I will argue that identity is a very personal matter. A single label may mean very different things to different people: deafness could be a disability to one person and a defining identity to another. Every person has numerous intersecting identities, and the importance assigned to each will vary between individuals and within a single individual as they show up and appear in different situations.

I would not call my childhood/teenage years conventional. Whilst not disabled myself, I gained experience of disability through my family. My brother was autistic and had Tourette's Syndrome. Only three years younger than me, we grew up together. My life was very different from that of my friends, but it was my *normal*. The fact that he appeared as different was noticed, judged, and usually, but not always, accommodated by people we met. This early experience certainly shaped me and meant that I saw the world in a particular way, opening my eyes to what life can be like for disabled people. Still, I only had a narrow understanding of the barriers society imposes. Before I began my MA in AVT, the world of subtitling, dubbing and audio description was one I had barely given a thought to. Despite this, with the knowledge I had, I tried from an early age to act as an ally and advocate of disabled people. The position I took with this was striving for inclusion in the classroom, and then, in this thesis, an orientation towards accessibility, a decision which will be examined further below.

Alongside this early experience of disability, came death in my family, of my mother when I was in my early twenties, and of my brother eight years later. He was living in a care home and his death was a result of multiple Health and Safety failures (Daily Mail Reporter, 2009). With this came a visceral understanding of what happens when provision is insufficient and the appalling, albeit extreme, consequences that can result when peoples' needs are not met. Whilst I did not then recognise what social or epistemic injustice, the theories that underpin the current research, were, I know that the seeds to act on them had already been sown. It was following my brother's death that I recall a very clear example of epistemic injustice being voiced: upon hearing that my brother had died, a colleague remarked, "At least it wasn't your stepbrother," who at the time was almost ten years old, with no disability. I did not want either brother to die, but I was struck by the value system in place: not being disabled outranked being disabled and being young and disability-free outranked it even more.

As I grew up, I also realised that many of the conventional, assumed labels I suggested above, did not apply to me. I am bisexual rather than straight. I am not Christian, but Buddhist. Both are labels that have taken me time and struggle in different ways to fully

embody, in the sense of accepting them, living them fully and taking on the difference and otherness - and also belonging - that they bring. To me, embodying an identity is more of a process than a point that is reached, and the degree to which I choose to embody either one might differ according to the situation I am in. As the epigraph at the start of this chapter states, identity is “about a state of becoming, always moving towards a future self that is made up of our past and present experiences.” Reflecting on the notion of *normal* in my own life, I feel like I cross the barrier between being *normal* and not, staying in and stepping out. The italicisation reflects my discontent with the preferential or hierarchical idea that *normal* implies, and the *not normal* or *abnormal* which apparently results. Through my lived experience, I have felt the prejudice this idea of normal and lack of acceptance can bring – not so much through outright confrontation or spoken prejudice as by knowing that although my appearance would suggest that normative accounts should apply to me, they often do not. This experience is not the same as being DH, yet both often involve a disparity between what is seen or assumed by others and the lived experience.

Other aspects of my identity have grown or changed in depth as my own life experience has evolved. I was born in England with British, then, later, European citizenship. However, it was only when I lived in France and Japan that I really felt English or British for the first time. By experiencing new education systems, work and lifestyles abroad, what had previously been simply normal in the UK, stood out as being different and so became more visible and familiar to me. Similarly, it was in Japan that my status and privilege changed for the first time: although I spoke some Japanese, without a proper command of the language, despite my desire to, I was unable to fully participate in the society around me. I was visibly and linguistically ‘other’. In the wake of the 2016 referendum, my status changed again. When asked whether the UK should remain or leave the EU, I did not even need to think how I would reply (though I wished the question had never been posed). However, overnight I had become a remainer, or a remoaner depending on your political inclination, and an embodied European. As I write this chapter on 1st February 2020, there is a disconnect as my European citizenship has

been removed without my consent. My internal identity remains the same, though I am no longer politically or externally able to claim it.

Only a few months after Britain's exit from the European Union, world events once more led me to reflect on my own identity and the place I have and want to hold in the world. Having first appeared in December 2019, the coronavirus was declared a global pandemic on 11th March 2020 (*BBC News*, 2020a). In the year and a half since then, life has moved from the unrecognisable through a series of 'new normals'. Lockdown, in its various forms, has been an experience shared by many around the world. Terms such as 'vulnerable' and 'key worker' have been redefined and privilege and inequality have come to be understood in new ways. Whilst Covid was first defined as a leveller (Jones, 2020; Ryan, 2020a), with the potential to attack everyone, it quickly became clear that some groups were affected by both the virus and the ensuing societal changes far more than others. In particular, the divisions in society that privilege causes were starkly exposed. If ever I had questioned it, my position of health, wealth and privilege became clear.

All these strands of my identity, the intersections within them and my deepening understanding of how they have and continue to shape me, are embedded within the research study that follows. As I explain in Chapter Three, action research was a natural methodology for me to adopt as a result of my reflexive practice in education. However, the deep-rooted way in which action research is engrained within the study and the principles which underpin it have also grown from who I am as a person and my desire to create opportunities for social and epistemic justice and to give a voice to people in shaping their own access. We live in a global society, and the choices each person makes have a significant impact on everyone and everything around them. Raising awareness of inequalities that exist and seeking a more equitable society through this research is one way I can contribute.

I have chosen to situate this research within accessibility studies, because I believe that access concerns us all (Greco, 2018) and that we all have a need for it. The normative/non-normative account, especially when discriminatory normative

assumptions are at play, seems to isolate one form of identity from another, creating hierarchies and separation. The approach it presents is a binary one, whereas I see identity, and indeed access, as something more fluid. In contrast, the accessibility account seeks to remove barriers and strives to foster access for all. It recognises that not only can everyone benefit from this, but that, more importantly, at some point, in some situation, everyone will need access in some form. Whilst I approached this research by taking non-native speakers of English (NNE) as an example of a wider audience group, the final guidelines which result from the study are written from the broadest possible point of view. Rather than giving a prescriptive checklist to achieving access, they encourage a questioning consideration of how access can cater for everyone.

These approaches are not without their challenges. Trying to provide access for everyone on every level simultaneously is unlikely to be possible. This research will necessarily involve exploring where and when natural barriers arising from multiple forms of access being provided may appear. For myself as researcher, another challenge arises; acting as an ally or advocate for others means learning from and listening to them, not taking charge or taking over (Louise, 2017). This research is user-focused - it is my responsibility to ensure that the users' voices are present in the words I write. As Charlton wrote in 1998, "Nothing about us without us"⁴. In the context of a PhD which also involves a personal journey, it is all the more important that I strike the correct balance and continue to give a voice to the key parties in this study, to those who will make use of this access, to the venues and respeakers who will help provide it and to the presenters who will be respoken.

1.3 Outline of the research

The ultimate aim of this study is to find out whether respeaking, a technique where a

⁴ Charlton published a work under this title in 1998, which I have referenced, though at the start of Chapter One (p.3), he acknowledges his own encounter with the slogan. Michael Masutha and William Roland, two leaders of Disabled People South Africa used it; they, in turn, had heard it used at an international disability rights conference by a person from Eastern Europe.

person creates live subtitles using speech recognition, can be introduced into the live event setting as a means of providing (sensorial) access for the audience and to fill a gap in the access provision that exists there.

The number of events that are currently accessible is (too) low. In 2018/2019, Stagertext, the national captioning charity with whom I have worked closely on this research, provided access at 373 captioned performances and supported access at an additional 154; working with MyClearText⁵, they made 206 talks and tours accessible with live subtitles (Stagertext, 2019: 9-10). Stagertext's aim is to provide access for the diverse audience that experience hearing loss, access that does not depend on an understanding of sign language interpretation. However, sign language interpreters are sometimes present at their live talks and tours to accommodate the diverse audience that attend them.

Many other events are also made accessible through sign language interpretation. Signed Culture⁶ is a charity that supports and promotes BSL access to the arts in the UK at diverse cultural events including theatre, musicals, operas, talks, tours and dance. While tours in multiple languages may be available at some museums, either with an in-person tour guide or through audio guides, these are not naturally accessible and when it comes to the theatre multilingual performances are far less likely⁷.

There is no single record for the percentage of live events that are accessible for people who are DH across the UK. However, when we consider the number of venues around the country and the number of events taking place on a daily basis, the information suggests that the proportion of accessible events is low. For example, the State of Museum Access Report (Cock *et al.*, 2018), which audits museums considered to be the best in the UK, reveals that only 3% mentioned BSL interpreted talks on their websites and only 1% referred to subtitled talks or tours. Whilst the actual provision may vary, as

⁵ Information about MyClearText can be found at mycleartext.com.

⁶ Information about their work can be found at signedculture.co.uk.

⁷ One exception to this is the work of the theatre company LegalAliens (Morash, 2017), yet examples of this are few and far between.

the report states, “access and inclusion start online” (*ibid.*: 5). The Museum and Heritage Access 2020 Survey (Cock *et al.*, 2020: 4) confirmed this, with 70% of those who participated stating they checked online for access information before visiting. Potential visitors use the information that is available online to decide whether or not to visit in person. By failing to mention accessible services that are on offer, future visitors may be lost. Respeaking could provide a solution to this by providing another form of access that can be used to make live events accessible.

In order to investigate the potential of respeaking as a solution, three key questions were posed, that guide this study:

1. What training and technical set-up is needed to allow experienced television respeakers to transfer their skills to the live event setting?
2. How can high quality respeaking and access be ensured?
3. Can the findings from this UK study be applied to the provision of access across borders and in society more broadly?

The first two questions are practical in their nature and directly shaped the methodology and the research design. They necessitated both the process and product of respeaking to be examined and an expected outcome for both was materials which could be used by others to establish this service, once quality expectations had been reached. The third research question, which asks whether there is a broader application for the findings of this study picks up the discussion from the previous section (1.2). A person is not and cannot be defined by a single label and, similarly, the use of an accessible service should not, in my opinion, be restricted to people who hold or are given that label. This conviction has led me to approach the question of providing respeaking at live events from wider discussions of translation, accessibility and access and framed by social and epistemic justice, intersectionality and within a universalist account of access. By doing so, I hope that changes in society can be explored which might lead to increased opportunities for social and epistemic justice.

These three questions have accompanied me over the last three years as I have gone about my own life. On one level, this has meant that I have sought out and attended

accessible productions and events and followed the work of charities that advocate for and support d/Deaf, deafened and hard of hearing people that respeaking is traditionally created for. On another level, awareness of issues surrounding accessibility have travelled with me as I travel, read the news and social media and live life on a daily basis. Access affects everything, or, perhaps, everything is affected by access.

It is for this reason, that I believe awareness is so important for change. As a person and as a researcher, I want and need to talk to people and hear their actual experiences, so that I can do my best to ensure that my research is tackling something of actual use, that will have real impact. I believe that it is only through increased awareness about particular needs and wider access that understanding and openness will grow; and if change is going to happen, this awareness, and the voices of users, must also reach the people who make decisions that affect access.

1.4 Chapter Overview

This thesis unfolds in the following way. **Chapters Two, Three and Four** present the theoretical, research-related and methodological underpinnings of the study and outline the gap in knowledge that this study seeks to fill.

Chapter Two, Diversity, Access and Inclusion, presents the theoretical standpoints adopted in this study. Situated in both translation studies and accessibility studies, this research draws on different aspects of each discipline to present an interdisciplinary study. Translation and access are both shown to be “ubiquitous” concepts (Blumczynski, 2016) which necessarily involve and impact all areas of life. The theoretical framework is comprised of three theories – intersectionality, social and epistemic justice and the social model of accessibility. There are many overlaps between them, yet they all contribute a significant element to the unified theory, one which highlights the changes within society that are needed for it to be a fairer and more equitable place. Intersectionality brings awareness to how the different layers of identity that an individual holds affects their interactions with the world around them. Social and epistemic justice capture this need for change and focus, in particular, on how resources are distributed

in society and how different groups of people are (under)valued. The social model of accessibility brings these principles together into a framework that suggests how access can be provided, with a series of principles that must be in place if these theoretical stances are to be upheld. Through these theories, some key principles behind the research are introduced. An underlying principle and desire for respect pervades this study, which could be considered as a new development on the theme of trust which has long been discussed in TS. Similarly, where norms are current parlance in translation terms, in this societal context, they appear as labels which shape and constrain us and the problematic relationship around these, when talking about accessibility, is particularly relevant. The chapter closes by showing how this study acts as a bridge between the two disciplines.

Chapter Three places a spotlight on the research gap that this study is intended to fill. The definition of a live event, and the variations within this setting are considered more closely, as is the gap that respeaking as a form of access would need to fill. The gap is considered from the perspective of legislative requirements, current access provision and expectations that such a service would be expected to meet. Alongside this, consideration is given to the practice of respeaking and the inherent challenges that exist within this modality, both as a form of subtitling and as a form of live access provision. A thematic review of key areas of research into respeaking and access is presented here to further highlight where this research sits within the broader disciplines of AVT, TS and AS.

Chapter Four presents the methodology and research design of this study and introduces the research methods and tools that have been used. Action research was chosen as an overarching methodology as it presented a framework which allowed the core principles behind the study to be embedded within the research design, whilst offering cycles of action and reflection required to answer the three research questions and develop an effective respeaking service at live events. In this study it functions as both a philosophical approach and overarching methodology and, for this reason, it is the central feature of the project design. This approach is not without criticism in the

literature, and key criticisms are addressed in the course of the chapter. The research design involves three cycles of action and research. The first is a cycle of observation and reconnaissance in the field (Elliott, 1991). In addition to observing live events in action, I conducted focus groups with users (DH and NNE) and providers (respeakers and venues). Semi-structured interviews and questionnaires were used here. The second and third cycles involved two rounds of respeaker training and research events. Individually, the research events acted as case studies into how different types of events are run and together they formed a large reception study. During both cycles, data collection involved a NER analysis of the respeaking at the events and further questionnaires and focus group discussions. The data collected across the study was a mixture of qualitative and quantitative and the rationale behind this is also provided in this chapter.

Together, **Chapters Five, Six, Seven** and **Eight** present the findings of the different research cycles of the study.

In **Chapter Five**, I analyse the findings from the first cycle of action research. The data collected during the four-fold focus groups is presented. Focus groups, conducted through semi-structured interviews and questionnaires, were used to explore experiences of access at live events and to delve more deeply into what expectations each of these groups had in respect to the use of respeaking at live events. The data comes in the form of statistical analyses from the questionnaires and a reflexive thematic analysis of the qualitative data in its different forms. Following the analysis of the responses from each group, commonalities and differences in their perspectives are considered, as is the way that this will feed into the initial respeaker training programme that results from these focus groups. Alongside this, some observations are shared from the event observation and as-live respeaking that I conducted.

Chapter Six explains how the respeaker training programme developed over the course of the research. The initial round of respeaker training which resulted from Cycle One of the action research and training methodology behind it is presented. The respeakers who took part in the programme are introduced, and their role as informed participants

is explained, as is the development of this working Community of Practice. Within this discussion, the processes involved in providing access at live events are outlined, the technical set-up used is described in detail and stylistic guidelines for respeaking at live events are presented. In line with action research, changes in procedures were implemented as the events progressed and new requirements for this training programme were discovered, which allowed the respeakers to deepen their understanding of the role they would need to take on when providing access at live events as well as allowing them to acquire the additional technological expertise that they would need to troubleshoot any challenges they encountered. The research events which provided these opportunities for exploration are presented in this chapter, and the diversity of the features they contained, as they ranged from public speakers, film panels, tours and Q&A discussions, is highlighted. Following this, the revisions to the training modules are discussed and by the end of the chapter, the final training programme, which would be offered to future respeakers entering the live event environment, is presented. Fundamental to this discussion is how the quality of the training and events will be measured. The concepts of the quality of experience and service (Greco and Jankowska, 2019: 8) are central to this.

Chapter Seven examines both rounds of research events from the perspective of a reception study - an audience reception study, as is traditional, but here the examination is expanded to include a respeaker-presenter-venue reception study as well, although on a smaller scale than that conducted with the audience. This is an opportunity to explore the feedback received from all involved in the events. The data comes from the questionnaires completed at each event and the experiences shared in the focus group discussions held after each event and by post-event reflections sent by some participants. Much of the data provided is qualitative, though a descriptive statistical analysis of the data from the audience questionnaires completed at each event is also included.

Whilst Chapter Seven focuses on the process involved in providing the respeaking and how the different parties experienced it, **Chapter Eight** focuses on the respeaking

produced; in translation studies terms, it is the product of the translation process that is considered. The features of respoken subtitles at live events are then presented, and the choices behind and origins of some of the conventions adopted at live events are discussed. A quality analysis of the respoken output follows, which paves the way to a broader consideration of this first corpora of respoken live events in English. An NER analysis of the accuracy of the respeaking provided during the events in Cycles Two and Three of the research allow comparisons to be drawn across events as well as between this corpus and the larger Ofcom and LiRICS corpora of respeaking in English. The latency seen at each event is also discussed. Case studies of respeaking in action from specific events accompany the data, so that the types of error seen and implication of the range and spread of errors can be understood in context. Analysing the data in this way allows further comment on the formative way in which learning happened, with input coming from all involved at each of the events. In the course of carrying out the NER analysis, it became clear that while the scoring system of the NER could be applied to the live event setting, new analytical pathways were required when making judgements on the nature of the errors seen. Following the analysis of the live events, I present a revised version of the NER for live events, the NERLE, which is adapted to this new setting. Guidelines for using this model can be found in the appendix. The chapter closes with a comparison of the ranking for each event seen in Chapters Seven and Eight, to determine to what extent the audience responses align with the industry-style technical analysis.

Finally, **Chapter Nine** provides an opportunity for the findings of this study to be shared in the shape of conclusions, discussion and suggestions for further work. The chapter begins with a moment of reflection. The research questions are revisited and the success of the methodology and research design are reviewed as the findings from Chapters Five, Six, Seven and Eight are recalled. The impact of this study is shared, including both the impact that has already been effected as the research was conducted and the future impact that is hoped will result. At the same time, the limitations of this study are

discussed and ways in which this study could be extended, relating to genre-, audience- and technique-specific considerations, are also suggested.

This review of the respeaker training programme, reception study, corpus of respeaking at live events and the NERLE and the innovation they represent are the responses to the first two research questions, which investigated what training and technical set-up would be needed to allow experienced television respeakers to transfer their skills to the live event setting and how high quality respeaking and access could be ensured. More detailed consideration is then given to the third research question, which encourages us to recall the theoretical approach and reflect upon whether the core principles behind the research remained embedded in it as the research progressed, asking how the findings from this study can be applied to the provision of access in broader terms in society.

It becomes clear that this question has already been answered to an extent in the course of the previous chapters, since this close consideration of how respeaking can be used to provide access has highlighted many questions which are often raised when discussing access provision. These findings are drawn together in the form of a model for participatory engagement when providing access through respeaking at live events. This model illustrates how participation, action and communication between all parties are essential if the best possible access is to be provided. Much emphasis is placed on the chain of access which must be maintained if events are to be as accessible as they can be (Greco *et al.*, 2012). This model for participatory engagement when respeaking is used is then extended into a participatory model for access, which aligns with the role of access co-ordinator, which was developed by the ACT project (Remael *et al.*, 2019) in the timeframe of this research. Appended to this chapter are questions and considerations to reflect on when making live events accessible using respeaking, which highlight the importance of pre-, mid- and post-event awareness and action for everyone involved in an event.

The chapter and thesis close by returning to the start. I consider the way the different strands included within this chapter, and the thesis as a whole, contribute to the disciplines of both translation studies and accessibility studies as well as to a number of

connected, yet interdisciplinary fields. I reflect on the direction in which I see accessibility studies moving, and how I hope that attitudes are changing towards access and d/Deaf awareness in society, and how I hope they will continue to change. Finally, I reflect on my own personal journey, with which this chapter opened.

A **glossary** of key terms follows Chapter Nine and the **appendices** to the thesis include the research tools used in this study and the data collected through them, the training resources for the respeaker training programme presented in Chapter Six, and guidelines for making live events accessible through respeaking.

Chapter 2: Diversity, access and inclusion

“What is the space between the both/and, where are the borders,
and who is drawing them?

...

Think of the twilight of dusk or the emerging light of dawn.

These are the spaces in between, the spaces where something is not one thing or another;
it is not yet day, and yet it is no longer night.”

Barker and Iantaffi, 2019: 203

In the first chapter of this thesis, I explained how intricately my own life story and experiences are woven within this research and how they contributed to the theories and methodological approaches I have adopted. In the current chapter, I situate the study on a theoretical and disciplinary level and outline the gap in knowledge that I seek to address.

Drawing from diverse areas of knowledge, I intend the account that follows to be interdisciplinary in its nature. Multidisciplinarity allows ideas from different disciplines to be considered alongside each other, leading to an increase in knowledge that remains within the boundaries of each individual discipline (Choi and Pak, 2007). In interdisciplinary research, the intention is to integrate these different strands of knowledge and provide an opportunity for new solutions to be explored and developed as the boundaries of what has gone before shift and evolve (*ibid*; Greco, 2018: 215). It is the latter that I seek to do. I hope that this study will contribute to this process of integrating and creating knowledge and act as a bridge between translation studies and accessibility studies, building on work already done in the former regarding the practice of (live) subtitling for d/Deaf, deafened and hard of hearing people, and adding an original contribution to the latter, emerging discipline, which is itself an interdisciplinary field (Greco, 2018: 215).

Approaching this research through the lenses of both allows for a more nuanced understanding of the factors at play when introducing respeaking into the live event setting and leads the way to this study being framed under the lenses of intersectionality,

social and epistemic justice. It is through these theoretical stances that I explore the concept of access and the framework I am using to present it, the social model of accessibility (Greco 2019a: 25 ; 2019b: 28-29). In this model, barriers to access are framed as a social issue, which necessitates social change; yet, the needs of the individual must be considered alongside decisions that are made in and for society. Such an approach aligns features of this research with the sociological turn of (audiovisual) translation, with its investigation of agency and structures within society (Chaume, 2018: 52). The social model of accessibility draws on a universalist account of access, which is based on the idea that, rather than focusing on a single sensory or linguistic need, media accessibility instead encompasses a variety of “products, services and environments, for all persons who cannot access them in their original form” (Greco, 2018: 211).

I begin the account that follows with an exploration of norms and labels as markers of identity and use this to introduce intersectionality, social and epistemic justice, and the social model of accessibility before situating access through these lenses. I then illustrate how this research acts as a bridge between (audiovisual) translation and accessibility studies.

2.1. Norms and labels as markers of identity

2.1.1 The use of norms and labels in society

Making judgements and assumptions about what we see and the people we meet are a natural part of life. Many of the actions we take are done automatically (from detecting that one object is further away than another to driving a car on an empty road) and these quick judgements and assumptions are no exception (Kahneman, 2011: 20-21). There are two systems at play in decisions we take, one is fast and automatic (System 1), and one more reflexive and considered (System 2). Kahneman calls these fast and slow thinking. Whilst we might like to think that we live and function in System 2, very often our impressions, intuitions, intentions and feelings come from System 1 (p.24). Nevertheless, System 2 is able to step in and take over (p.23).

As humans, we have a tendency to link what we see and encounter to what we already know, creating patterns and ways of acting. “We are pattern seekers, believers in a coherent world” (p.115). Many of the actions we perform are innate, genetically determined or pre-wired, although they may be modulated by experience (Costa, 2011: 579). Gradually, these patterns form units of behaviour or practices (Baum, 2000: 186), which develop through rule-making and systems of action, context and reinforcement. Selection further determines which practices are likely to be replicated and passed on and in what situations (p.203-4). Whilst these patterns and practices may vary between cultures, our fast and slow thinking draw on them and inform how we navigate society. Kahneman’s (2011: 71) System 1 maintains and updates a model of the personal world of each one of us, assessing and representing what is normal in it for us. The associations we make, “that link ideas, circumstances, events, actions and outcomes that co-occur with some regularity” feed into it. Patterns within these associated ideas begin to represent the structure that we expect to encounter in our life (*ibid.*) and these expectations determine what we consider normal or not. Norms exist for categories in all areas of life, and these norms specify both the range of plausible values and the most typical cases (p.74). It is System 1, which understands language, that has access to these norms. This act of using norms and categories to navigate the world around us has become embedded in our society, as our use of labels reveals⁸.

Labels guide and direct us through life, from which toilet we choose to the age-determined class a child enters at school. They are used to capture characteristics including age, gender, race, sexual orientation, (dis)ability and many more. They highlight a particular aspect of a person and either provide specific information about them, or, if a label is unfamiliar, provide a reference point for learning more. Arranging classes by age, means that teaching can be targeted towards a specific level, even if variations exist within the abilities in that class. Knowing that a child has a specific

⁸ I have chosen to talk of *labels* rather than *signs*, although the latter might seem a more natural choice given the use of signs when talking of language and translation (for example, Saussure, 1916/1983). Since the argument I am making is broader in scope than linguistics alone, *label* seems the most appropriate term to use.

learning need may lead to a particular approach to learning being adopted. Parents who discover that their child identifies as gay (or as any gender or sexuality within the LGBTIQ+ umbrella) may find any one of these labels a useful starting point for further research if this is an unfamiliar world to them. And if you have a guest coming for dinner, knowing that they are vegetarian or vegan is something you would probably prefer to know before you plan the menu.

Society has long treated different groups of people in different ways as a result of certain characteristics they hold: the different rights and responsibilities that men and women have traditionally held and racial discrimination are both examples of these. Today, with human rights provision in place, it might seem logical to assume that similar rights and opportunities exist for all people. However this is not the case and the two examples just given are themselves not yet consigned to history. Many services exist to ensure that provision and (legal) protection are available for those who need it, and adopting a label is a necessary requirement to accessing these. Sometimes self-identification is sufficient, but in other cases official categorisation is needed.

Labels are therefore an essential tool within our society, serving as a useful point of reference, an “admission ticket” to services (Ho, 2004: 87) and a vital source of protection. Yet, the fast categorisation associated with labelling can also be problematic. Although they provide a useful starting point in imagining what a person’s needs will be, there are limits to how much labels can accurately convey. Just like the symbolic image on a toilet door (Fig. 2.1), a label gives an indication of who may be found inside, but the person behind the door looks very different. If this reality is lost, needs may be addressed in an oversimplified manner and a lack of provision may result.



Fig. 2.1: Sign found on a toilet door (All Sign Images, 2021)

When a label is used to categorise a group of people, it creates (for the outsider) a sense that similarity and homogeneity exists between the group members, which at the same time sets them apart from those who are not part of the group. This label comes, in some way, to define and represent who group members are. What results is often a fixed and potentially generalised notion of who and what a particular label refers to and, from an access perspective, the provision that will be required. If we take the case of deafness, hearing loops can be found in many buildings. Whilst many may see them and think access is available, the reality is that the loops do not provide access for all d/Deaf people. Similarly, the suitability of the access provided by BSL interpretation is often overestimated by the hearing community – whilst for BSL users, BSL interpretation is very likely to be preferred as it offers access in their first language, not all d/Deaf people are BSL users; once again, it is not an access option that meets the needs of the *whole* group. Extending the group to include hard of hearing people adds further complexity to the discussion.

By focusing entirely on the impairment, which often happens when labels are used, we are taken into medical-model thinking, whereby it is the person with the impairment who is the problem and needs fixing. In contrast, social-model thinking is focused on identifying the barriers that person faces and finding solutions (Shape Arts, no date)⁹.

⁹ The medical and social models of disability are discussed further in section 2.2.3.1 below.

Through its focus on the barriers society creates, the latter approach is better equipped to dealing with the richness of experience that exists behind any label. Quite what deafness means and how it manifests for any individual is likely to be affected by many factors. The onset and degree of deafness are certainly important (see section 3.5.1 for further discussion of this), but how deafness is experienced may also change over time; the surrounding environment has an effect, as does a person's lived experience. Despite the supposed categorisation that a label implies, the reality is that there is no single, knowable experience of reality (for example, Braun and Clarke, 2013: 6).

This pull between the generalised experience a label's categorisation suggests and the individuality of actual experiences must also be considered. We have already noted that in categorising people, a label includes some and excludes others. Another, more subtle layer is at play in this process: the normative bias that most labels hold and the consequent positioning of the labelled groups in relation to this norm.

2.1.2 Discriminatory normative bias

Whether we recognise it or not, these normative frameworks drive society. Fast thinking is based on norms and, as humans, we have a tendency to relate more easily to people who are in similar groups to us. As a consequence, many of our ideas are built on what we are familiar with and many normative assumptions result. Such assumptions are beliefs about how things are or should be, for example, the idea that we are all equal and that we should all have human rights (United Nations, 2012: 10). Some normative assumptions, however, are discriminatory in their nature, for example when they are used as a conscious or unconscious background to judge others. This often presents as implicit bias and may happen when the fast thinking of System 1 occurs without the reflexive input from System 2. And this, in my opinion, is where normative assumptions become problematic and potentially result in people being judged in a discriminatory way. Rather than being a static construct, society is continually evolving and thinking and value systems change over time. There may be occasions where, even though, for example, legislation has evolved, our own actions and reactions may be harder to rewire,

and fast thinking and implicit bias take us back to more outdated notions. It is at times like this when discrimination may result.

This discrimination may happen on an individual level or be embedded in society. For example, Criado Perez (2019: 125-127) cites alarming examples from the area of safety research where standardisations in Protective Personal Equipment (PPE) are based on average male data (norms) and therefore fail to account for the experience of those who are 'other'¹⁰. As a result, the PPE equipment issued is poorly fitting, causing pain and discomfort and, in one case, proving fatal: in 1997, a female police officer in the UK died from a stab wound inflicted upon her as she entered a flat with a hydraulic ram; she had removed her body armour as using the ram while wearing it was too difficult (p.127). Further research reveals that often, a standard *white* male norm is in use, making facial PPE gear unsuitable for men from black and minority ethnic groups or with facial hair (TUC, 2017: 4). Labels can often mask a hidden or underlying attitude that is discriminatory or exclusive, or which privileges a particular group over another. When used alone, *Fiction* is an inclusive term; however, when subdivisions of *Women Writers and Fiction* (Amazon.co.uk, 9th May 2020) but not *Men Writers and Fiction* appear, suddenly *Fiction* seems to represent a mostly male domain.

In the case of disability, this normative perspective becomes all the more apparent as it is also ableist, with discrimination happening in favour of non-disabled, *able-bodied* people. Unless the situation is one when there is a very strong drive for access and inclusion to be at the forefront of planning and design, the status quo is that the normative citizen, the one in mind when decisions are taken, is an able-bodied citizen.

When the coronavirus lockdown was adopted in the UK, one of the first directives from the government was to stay at home (Prime Minister's Office, 10 Downing Street and The Rt Hon Boris Johnson MP, 2020); travel was only permitted for work where this was absolutely necessary. Whilst achieving this was not without its problems, organisations

¹⁰ When explaining her use of 'other', Criado Perez refers to de Beauvoir (1949/1990), where women were regarded as 'other' to the male norm. In this thesis, I use the term more broadly to refer to any group that sits outside the norm, whether due to gender, race, disability or any intersectional identity.

swiftly adapted. Homeworking became a reality, online conference platforms made remote meetings a possibility, education moved online and much guidance was published to support people with this transition (Health and Safety Executive, no date). Yet, long before the pandemic, many disabled people had requested these exact adjustments through *The Equality Act 2010*. Many disabled people offered advice on how to deal with the challenges of isolation that might result (for example, Rosa, 2020; Ryan, 2020b) and at the same time, the resistance that many had encountered in response to these requests was also revealed: “As well as joy at being offered new opportunities, many feel frustrated that it took the non-disabled world to become house-bound before access was granted” (Ryan, 2020c). The pandemic at once highlighted the ableist norms and expectations which drive how society functions - and how quickly these norms *can* be altered when the conditions demand it.

Even when such changes were made in response to our restricted movement, other ableist norms persisted; the experience of d/Deaf people was one of many that was overlooked. The use of non-clinical facemasks to limit the spread of coronavirus was mooted early in the pandemic, but only much later did the isolating impact that this could have on d/Deaf people get picked up in the news (Betteley, 2020). Whilst masks with transparent panels were quickly shared and promoted by individuals and charities, take-up at a clinical level took longer to follow (Taylor-Coleman, 2020). The daily briefings from the government, which ran from March 16th to June 23rd (*BBC News*, 2020b), were an important source of information and analysis, yet at no point was a BSL interpreter present in the briefing room despite the devolved governments and many others around the world having one and multiple requests through the #WhereIsTheInterpreter campaign. As well as restricting access to vital health information, this lack of visibility was widely perceived as treating BSL users as second-class citizens (Pring, 2020)¹¹.

¹¹ In fact, sixteen months later, the failure of the UK Government to provide an in-person BSL interpreter for these data briefings was ruled illegal by the High Court, who said it amounted to discrimination under the *Equality Act 2010* (O'Dell, 2021).

Whilst coronavirus may have brought increased attention to such injustices, these ableist trends are not new. Banning plastic straws in the fight against climate change overlooked their vital role as an accessibility tool and their original use among people who were ill or disabled, before becoming a mainstream device (Schultz, 2019). However, in the process of implementing the ban, the voice of disabled people was “left out of the conversation” (*ibid.*).

Similar cultural marginalisation can be seen in many other areas. It took Mattel, the producers of Barbie, 60 years to have a permanent group of disabled dolls within their diverse Barbie Fashionistas range (Ibbetson, 2019). Prior to that, a wheelchair using friend of Barbie called Wheelchair Becky had been a temporary addition to their line in 1997 (Max, 2013). A little ironically, Barbie’s Dream House was at first inaccessible to her and it was only after her wheelchair was redesigned that Share a Smile Becky, now rebranded, could enter (*ibid.*). #ToyLikeMe, an arts and play not-for-profit organisation, took a different approach and, in 2015, began giving makeovers to existing toys to give them disabilities. They wanted to “normalise disability for what it really is, part of the natural spectrum of human life” (#ToyLikeMe, no date b) and “separate disability from tired hospital, baddie and geek associations and instead create a new more celebratory and fun aesthetic by giving fairies guide dogs and wizards wheelchairs” (#ToyLikeMe, no date a).

There were two purposes behind their actions: firstly, they wanted to prevent the sense of self-isolation and low self-esteem that can result when positive representation of people similar to oneself is missing or lacking (#ToyLikeMe, no date b); secondly, they had observed that toys representing disability reduced the anxiety and prejudice non-disabled children sometimes felt towards disabled peers (O’Neill *et al.*, 2018: 1).

The stereotypes noted by #ToyLikeMe are widespread in the representations of disabled people on screen and many others exist as well (for example, Carr and Darke, 2012:40), although representation is gradually beginning to change. Liz Carr, a disabled actress and activist, stated that many writers do not know how to write for disabled actors, and much “risk aversion and fear” remains (Carr, 2020), as she reflected on her role in *Silent*

Witness (2013-2020). She spoke of her pride about the representation of disabled people that the show had achieved, but also of the policing she had done and input she had given to ensure that problematic lines and moments were avoided (*ibid.*). In their play *Still No Idea*, Hammond *et al.* (2018: 56) conclude that things will only change when “disabled people [are put] into positions of power: writing the stories, commissioning, programming”. 2020 saw a series of storylines on the BBC where nuances of d/Deaf experiences were brought to the screen in collaboration with d/Deaf scriptwriters. *EastEnders* (2020) used creative subtitles¹² in conjunction with muted audio to illustrate what the character Ben was able to hear (Lomax, 2020) and in *Casualty* (2020), layers of discrimination surrounding deafness were brought out through Jade’s story, alongside the impact of different sounds in her daily life (Dainty, 2020). It seems that achieving diversity in all respects on screen necessarily involves diversity at every level; only in this way will there be a platform for representation and the authenticity needed to escape the stereotypes that might otherwise dominate. While diversity is growing on film and television screens for certain groups, the white ableist norm prevails (Pulver, 2020). More change is needed.

In fact, the more closely we investigate the implicit value and affinity that is attached to the prevailing norm and the discriminatory normative assumptions that result, the more Goodley’s (2014:22) description of the valued citizen of the twenty-first century seems to ring true:

Cognitively, socially and emotionally able and competent; biologically and psychologically stable, genetically and hormonally sound and ontologically responsible, hearing, mobile, seeing, walking, normal, sane, autonomous, self-sufficient, self-governing, reasonable, law-abiding and economically viable, white, heterosexual, male, adult, breeder, living in towns, global citizen of [West Europe and North America].

¹² I acknowledge the debate surrounding the use of ‘creative subtitles’ and ‘creative AVT’ (for example, Romero-Fresco, 2021b; Sokoli and Pedersen, 2021; Romero-Fresco and Chaume, 2022), but use it for simplicity in this case to denote the difference from more traditional forms of subtitling which I have discussed in this thesis.

2.1.2.1 *The impact of labels on individual identity*

Since access provision is so often determined by these labels, and the normative assumptions that underpin them, it naturally raises the question of what impact adopting or being assigned a label in order to access it has on the individual. In the 1960s, labelling theorists wrote of the stigma that resulted from a label. For Goffman (1963), this stigma centred on difference and deviation from the norm, which prevented a person from being fully accepted in society. This stigma was a “deeply discrediting attribute”, one which reduced a person “in our minds from a whole and usual person to a tainted, discounted one” (p.13). Revisiting this notion in 2001, Link and Phelan redefine stigma as existing when elements of labelling, stereotyping, separating into ‘us and them’, status loss and discrimination “co-occur in a power situation that allows these processes to unfold” (p.382). The role of power in this instance is a subtle yet significant one. Members of any group might have stereotypes about others, but when one group has power, the other will endure “important consequences” (p.372). As a result, this label and the associated stigma, has the potential to impact all areas of life, from access to employment and housing, to psychological well-being and sense of identity (*ibid.*)

In the case of deafness and hearing loss specifically, a great deal of stigma remains (Wallhagen, 2010). Whilst some of this stigma stems from altered self-perceptions, in particular with regard to ageing, it is also reinforced through society (p.68). Hearing loss is something people feel able to joke about (RNID, 2018) and an embarrassment around wearing hearing aids is suggested by the current trend for discreet and invisible hearing aids (*ibid.*). In fact, it can take a person seven years from experiencing hearing loss to actively looking for solutions such as hearing aids (Clason, 2021). If a person is in denial about their experience of hearing loss, it is unlikely they will be aware of provision such as a captioned theatre performance or live subtitles, especially if they are targeted at DH people. Would they associate that label with themselves? In contrast, if such access were provided in a more inclusive way, a regular feature in any production, there would be more awareness of deafness and it would be normalised as an experience, perhaps

resulting in less stigma in society. Such access might even help someone acknowledge their own experience of hearing loss in a supported way.

For all the problems I have raised surrounding their use, labels and all they represent *can* be an incredibly powerful and self-affirming tool. Used this way, labels often confer identity, but an identity that is determined and defined by the individuals who take them on. Rather than accepting “boring old grandparent/artificial limb coloured hearing aids” (Gallant, 2012), many d/Deaf people now pimp and personalise their hearing aids, using them as an expression of their personality and identity (SARAHSMITH76, 2013); the same trend is being seen with canes and crutches (Merrick, 2015). From *queer* to *crip*, many terms that were once used pejoratively have been reclaimed and embodied (Rand, 2014; Hamscha, 2016). Increasingly, the protection and identity offered in law by labels is in demand: in 2020, a court ruled that they recognised ethical veganism as a philosophical belief that should be a protected characteristic under the *Equality Act, 2010* (McCulloch, 2020). Even when labels are assigned externally, for example in medical contexts, their impact on the individual can also be a positive one. Mogensen and Mason (2015), when exploring autism, found that in cases when a diagnosis with a label facilitated agency, knowledge and control, it could be advantageous (p.266), often relieving a person of their need to conceal or mask their difference.

2.1.2.2 A non-binary approach

Moving away from the duality and fixed positions that specific, normative labels imply through the either/or thinking that is inherently behind them is one way of addressing the us/them divide that labels can cause. As the epigraph to this chapter suggests, seeing the fluidity with(in) identities is an important step, as is seeing the implicit attitude which shapes them: “What is the space between the both/and, where are the borders, and who is drawing them?” (Barker and Iantaffi, 2019: 203).

The case of disability provides an effective example of this non-binary state:

The fact is that most citizens will have some level of impairment, some degree of physical difference from others. Most humans, as they age, will find themselves less able to see, hear, walk or think as well as they did before. One disability activist recently spoke at a

convention to 'normal' people and said, 'We are 500 million strong and growing. Come back in twenty years and a lot of you will be with us!' (Davis, 1995: xv)

Approaching labels in a more fluid way is likely to better capture the range of experiences that individuals may experience; certainly, seeing a range of experiences within deafness, according to whether an individual is deaf, Deaf, deafened or hard of hearing, is more realistic than suggesting there is a deaf/hearing or even deaf/hard of hearing divide that can be pinned down in a straightforward way. This non-binary approach also provides space for individual interpretations of identity to be accommodated.

As Davis (*ibid.* p. xvi) points out, many people who are not impaired in the usual sense of the word may consider themselves part of the disabled community. Similarly, there are other people that society might consider to have a disability, but who may not self-identify as disabled. Certainly, deafness is considered by many as a cultural identity, rather than disability (for example, Jones, 2002; Ladd, 2003)¹³. It only becomes a disability when barriers to access arise in society as a result of failures to meet Deaf people's needs (Üstün *et al.*, 2001: 7). CODA, children of deaf adults, may consider themselves bicultural, and part of the hearing and deaf communities (Davis, 1995: xvii), although medically speaking, they would not be considered deaf. From an ableist perspective, the very etymology of 'disability' suggests an inability to do something, and may lead to assumptions that disabled people are precluded from certain activities. A more authentic experience of disability may mean doing things in a different way and while becoming disabled may necessitate change, it may also lead to new and different perspectives and opportunities as any change might. The neurodiversity movement provides another example of this, where, for example, the "atypical communication" of autism is seen as natural and positive, rather than something to be changed or cured

¹³ It is also worth noting that whilst in legislation such as the *Equality Act, 2010*, adjustments for d/Deaf people would be made through the characteristic of disability, many inclusion-oriented groups now talk of Deaf and disabled people to highlight this cultural identity. Inclusion London (<https://www.inclusionlondon.org.uk/>) is one example.

(autisticallity, 2014). In fact, the very naming of this type of communication as “atypical” reveals how difference is pathologized because autistic people are a minority (*ibid.*).

The normative assumptions surrounding the term *disabled* may also work in reverse. People with hidden or less visible disabilities may not appear disabled and may therefore not be considered disabled by many who see them; once their disability is revealed, others may judge that they are not disabled *enough* (Guffey, 2018: 193). Gender and sexuality provide another case where a person’s visual identity may not match the identity within and ‘non-binary’ has itself become a label in this context.

2.2 A triple theory approach

Therefore, while labels may seem like a very straightforward device to use when creating and determining access provision, they actually hide a far deeper complexity, one which potentially impacts on the range of access provided and how any person who uses that access both sees themselves and is seen. In today’s society, there is a drive to value and celebrate difference, and, at the same time, through the implicit normative assumptions that underlie many decisions, there remains a fear around that which is different and not understood. This may result in differences, and those who are ‘different’ not being (fully) acknowledged, and consequently erased. Such erasure means that dominant cultural assumptions about what is normal persist (Chon, 2000: 440) and that stigma and discomfort around difference are reinforced (Hughes, 2012; Aho and Alter, 2018). For this reason, society and social policy have a vital role to play in determining both the extent to which stigmatisation and marginalisation are experienced and the impact that labels bring, be it positive or negative (McDonald, 2009; Hughes, 2012; Aho and Alter, 2018).

These considerations have led me to situate my research through three interrelated lenses. Before presenting each in more detail, I will briefly outline how they work together and complement each other (Fig. 2.2).

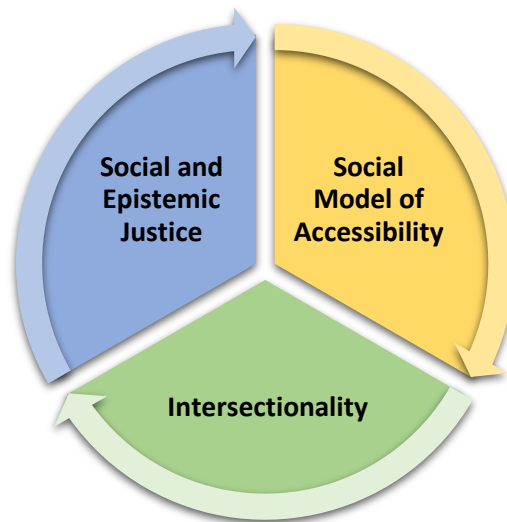


Fig. 2.2: Triple theory approach

The fact that we have complex internal and external identities is a fundamental aspect of this study and, for that reason, the first of the three lenses is intersectionality. Notions such as the non-binary view discussed above and theories of complexity (Burnes, 2008) inform how individual identity could be treated in relation to access provision, but intersectionality seems to offer the most appropriate perspective. It takes account of the interconnected nature of the different labels any individual has and acknowledges how these categorisations act in a simultaneous and mutually constitutive way (lumencandela, no date): when experienced together, the discrimination that results from the combined effect of two or more labels, is greater than the sum of experiencing the effect of any of these labels individually. Thus, a person at any marginalised intersection is potentially more vulnerable to societal discrimination. However, as well as offering an analytical framework for identifying this, it is also a tool for empowering people (Collins and Bilge, 2016: 37).

Social justice, with its demand for a fairer and more equitable world with respect to how wealth, opportunities and privileges are distributed within a society (distributive justice) (LegalDictionary Content Team, 2016), is a fundamental aspect of intersectionality. For this to be achieved, change at many levels within society is needed. Epistemic justice

adds an additional slant to this demand: in addition to achieving a distributive change, a shift in our values and understanding with respect to people who are different to us is also required. If equitable access is to be provided, we first need to understand the individual and diverse ways in which society is experienced. This combination of social and epistemic justice brings an ethical dimension within the study, which is also reflected in the social model of accessibility (Greco, 2018: 220).

In addition to providing an alternative to the normative biases that are all too often seen in society, the social model of accessibility (Greco 2019a: 25; 2019b: 28-29) also introduces a sense of both poietic agency and epistemic value towards those who benefit from access provision. Turning on its head the idea that access is provided for a single, homogenous group of people, at the core of this model is the idea that a single type of access may be used by a range of people in different ways (*ibid.*) and that knowledge and planning is required to embed such access within a given context. In this way, the empowering potential of intersectionality and social and epistemic justice can be actualised: as the disability rights slogan so clearly states: Nothing about us without us. Everyone must be involved in the discussion and shaping of this provision for its nuances to be understood and for its maximum impact to be felt.

Combined, these three lenses allow a participatory model for access and engagement to be posited in the final chapter, which at once accommodates the diverse experiences and needs that individuals have in an inclusive way, whilst encouraging agency and a mutual respect for those involved in this provision.

2.2.1 An intersectional approach

In essence, intersectionality is “a way of understanding and analysing the complexity in the world, in people and in human experiences” (Collins and Bilge, 2016: 2). With a strong focus on the inequalities that exist in society and how they are created, intersectionality recognises that multiple factors work together in shaping social and political dimensions of life and provides a framework for examining it:

When it comes to social inequality, people's lives and the organization [sic] of power in a given society are better understood as being shaped not by a single axis of social division be it race or gender or class, but by many axes that work together and influence each other. (*ibid.*)

Whilst the term intersectionality was coined by Crenshaw in the late 1980s/early 1990s, the understanding behind it long predated this (Collins and Bilge, 2016: 63-87). Many had spoken and written of their own lived experiences of seeking spaces where they could show up in the fullness of, for example, their race and gender (Harris, 2001: 292-294), rather than having to effectively set one of these aspects of their identity aside to join the debate. In naming intersectionality, Crenshaw (1991) brought these understandings into academia and enabled them to be reconfigured as a form of critical enquiry and praxis (Collins and Bilge, 2016: 81). Over time, attention has spread beyond the original intersections of gender, race and social class.

How is intersectionality operationalised within this research? On a methodological level, rather than focusing in on a defined cluster of intersections, I apply the 'both and' principle (Collins, 1993; Barker and Iantaffi, 2019) to the social axis of d/Deaf-hearing, since I want to investigate which intersections arise alongside it in relation to access within the live event setting. As Crenshaw (1991: 1245) herself acknowledged, many different intersections may be critical in shaping experience. I believe these may be revealed more naturally by inviting conversation about access, rather than trying to prescriptively determine them in advance. From an analytical point of view, it means my data may be on the qualitative and experiential side; nevertheless, I hope a variety of intersections along this axis which affect access will be revealed. In my call for volunteers, I invite people who are d/Deaf, deafened and hard of hearing during the initial round of data collection (Chapter Five) and extend the invitation to people at the hearing end of the spectrum to join the live events (Chapter Seven). In doing so, I hope to find participants along this axis and leave the interpretation of the labels used as open as possible. Following the discussion above and in line with the comments of de Vries (2015: 17), I

know that the labels used will influence who chooses to participate, since ultimately, some identification with the label used is needed if a person is to respond.

In addition to the d/Deaf-hearing axis, I also consider a second social axis within this study, that of language. Although the link between language and identity is clear, and the fact that the linguistic factor can be a “crippling obstacle” to social inclusion (Piller and Takahashi, 2010: 550), I did not choose this axis from an intersectional perspective – in fact, it is only in more recent studies that language has been studied as an intersection (de Vries, 2015; Atewologun and Mahalingam, 2018). I hoped that including a language axis would provide insights into how access, here respeaking, is used in different ways by a range of people. What was interesting was that the two axes naturally intersected through BSL signers. Far fewer participants from this second social axis joined the study, but there was nevertheless an opportunity to explore how the needs of categorically different groups complement each other and can be built on to create a more universalist approach to access (see section 2.2.3.2 below for a more detailed explanation of this point).

McCall (2005: 1773) outlined three methodological approaches to intersectionality, defined by their stance towards and use of analytical categories “to explore the use of intersectionality in social life”. These ranged from a rejection and deconstruction of categories (anticategorical complexity) to a provisional adoption of said categories to be used in a strategic way (intercategorical complexity). Between these two points on the continuum comes intracategorical complexity, which interrogated the implicit boundaries of categories, whilst acknowledging how they can be used. McCall also notes that some studies fall between the anticategorical and intracategorical complexities due to the ambivalent status they assign to categories (p. 1783) since categories do not represent the diversity of experiences; it is here that I would position my methodological stance. However, whilst McCall suggests such ambivalence results from the attempted homogenisation of standard groups, in contrast to the diversity of the group under consideration, my study is ambivalent for a different reason. I do place more attention on the non-standard group, but I see the range from ‘standard’ to ‘non-standard’, to use

McCall's terms, as a continuum, with no end of the spectrum more homogenous than the other, in line with the non-binary approach referred to above. Similarly, since within deafness, there are so many potential subcategorisations within the d/Deaf-hearing axis in particular, an element of intracategorical complexity is also present within this research. In the analysis, where participant numbers allow, some conclusions will be drawn about the views expressed by groups at different points along the d/Deaf-hearing continuum.

De Vries (2015) offers a new approach to the complexity of intersectionality, through his multidimensional prism. Here (Fig. 2.3) he expands the non-binary approach and examines the multiple ways that twelve planes (his term for social categories) intersect. Within this study, the focus is on individual experiences of access within the live event setting, and an intersectional perspective is used to explore this more deeply. However, structural and political perspectives are at work behind these individual experiences; as Anderson (2012: 171) states, "Structural injustices call for structural remedies". The dynamics of power within and across these multiple levels must be considered if any conclusions drawn and suggestions made are to be truly embedded within them. Whilst a desire for change may be felt at a personal level, all three theories that I have adopted make a very clear statement that change has to happen at the structural, societal level. Whatever model of access and provision is ultimately adopted, it will certainly affect the individual experience of access; political awareness and influence will also be required. In fact, many criticisms that have been levelled against intersectionality stem from the misguided notion that individual identity is considered to the exclusion of structural and political factors (Collins and Bilge, 2016: 124-131).

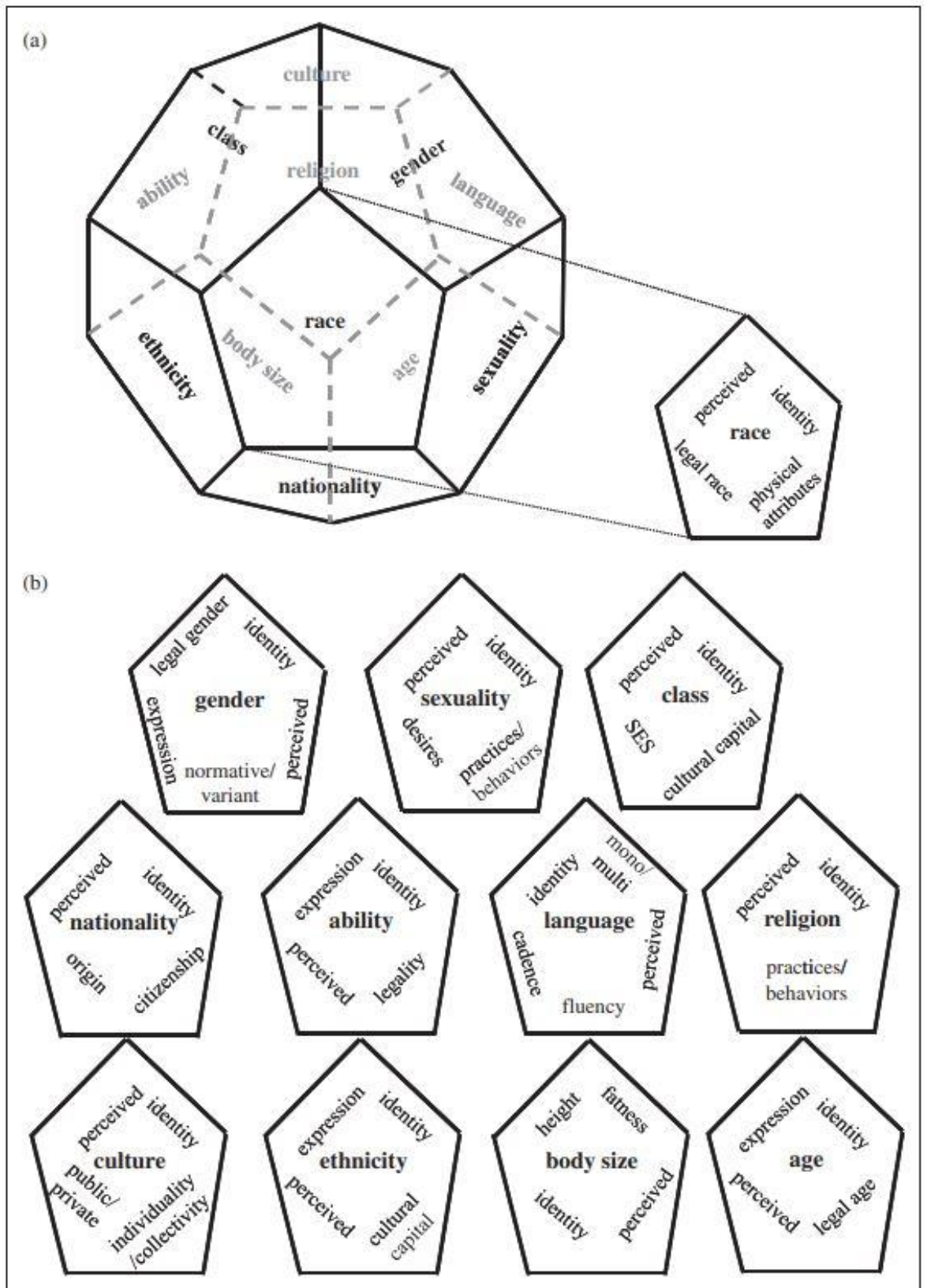


Fig. 2.3: De Vries' Multidimensional Prism of Intersectionality (2015: 11)

Some have suggested that intersectionality is too essentialist, conceptualising individuals as having fixed or unchanged identities (p.124). I hope from the discussion so far that it is clear that this is neither my stance nor one that I believe intersectionality promotes. Others have criticised intersectionality for having “too much identity politics within it” (*ibid.*), whereby individual groups might move away from broader-based party politics to fight for their own cause (Merriam-Webster.com, no date). However, as Crenshaw (1991) explains, there is no advocacy for any kind of political promotion within intersectionality, and any shift in power would simply lead to a different discriminatory dynamic, working against the principles of social justice. Rather, intersectionality frames the injustice that these groups experience and in that way could be considered as advancing the struggle that identity politics was designed to address.

Alongside this, some have criticised intersectionality for being “exclusionary” (Collins & Bilge, 2016: 126) with such an intense focus on marginalised communities, the implication being that it should instead be used for a more universal approach (Hutchinson, 2001). This discussion of where to focus the lens of intersectionality, including within this study of access, is one I have found complex to negotiate, even as I have pursued an account of access. The association of intersectionality and diversity have raised similar dilemmas: when the two are closely entwined, there is a risk that fundamental principles of both may suffer (Collins & Bilge: 188). Ahmed (2012: 14), reflecting on this in the context of race within academic institutions, suggests that when diversity becomes a view, concern with structural issues and power relations may disappear:

If diversity is a way of viewing or even picturing an institution, then it might only allow some things to come into view. Diversity is often used as shorthand for inclusion, as the “happy point” of intersectionality, a point where the lines meet. When intersectionality becomes a “happy point”, the feminist of color [sic] critique is obscured. All differences matter under this view¹⁴.

¹⁴ With reference to this, in the title of this chapter, I use ‘diverse’ in its sense of difference, rather than as a shorthand for inclusion.

This is indeed a difficult path to tread. I do not see the two approaches as incompatible, yet, both need to be negotiated with awareness, so that fundamental dimensions of intersectionality do not “recede from view” (*ibid.*) and injustice is not recreated in a different way (Penney, 2001). I hope, as Collins and Bilge (2016: 190) suggest, that a critical stance within both may enable the social justice dimension of both to be enhanced and this is a point that I return to frequently, in the first instance in 2.2.2 and 2.2.3 below. I believe that a critical stance and awareness is central to any study which utilises intersectionality. A reflexive stance is inherent within my own approach, something which will be discussed further in Chapter Four, and intersectionality has added a new slant to this. This research relies heavily on interaction with multiple focus groups and organisations. I step in and out of different functional roles within this research – a respeaker, a researcher, the event organiser, an audience member – and my own intersections and experiences show up in different ways. Intersectionality highlights the complexity of these interactions and the different dynamics within each of these relationships. Atewologun and Mahalingam (2018) encourage researchers to be able to *shift* and *root*. The former is being able to understand the “situated nature of other participants’ positionalities” (p.158); the latter is having a reflexive understanding of one’s own. They provide an intersectional identity web as a tool for mapping and visualising how different strands of identity are more or less dominant at any point in time (Fig. 2.4). Whilst I do not necessarily agree with the exact expressions they have chosen for each of the strands, I found the concept behind the web very useful as a personal reference point and as a way of acknowledging more consciously where my own power and privilege lie, and how to navigate that as a researcher. Since the public-facing element of this study was not named as an intersectional one, I did not use this tool with participants at any point in the study; reference to some strands were included within the demographic section of the questionnaires used.

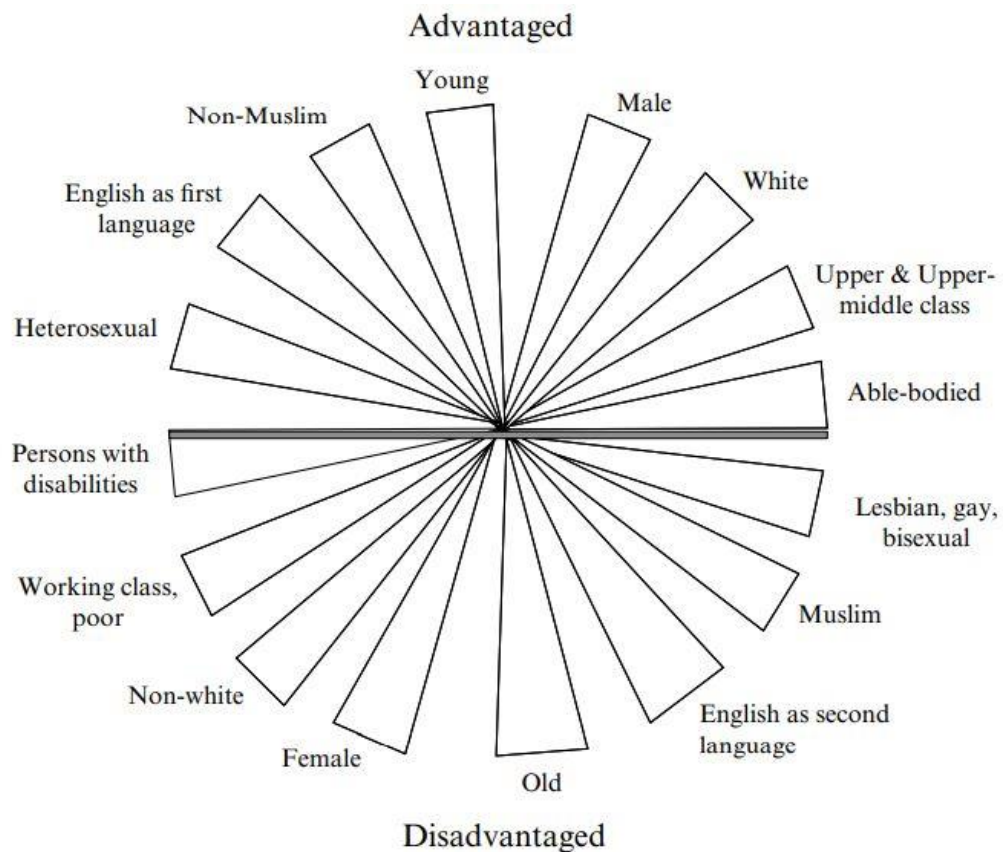


Fig. 2.4: Intersectionality web from Atewologun and Mahalingam (2018: 163), adapted from (Morgan, 1996)

SOURCE: Figure 7.1 Intersectional Identity web, Intersectionality as a methodological tool by Atewologun, D. & Mahalingam, R. © Lize A.E. Booysen, Regine Bendl and Judith K. Pringle 2018. Reproduced with the permission of The Licensor through PLSclear.

In summary, taking an intersectional approach within this study means that the complexity within the research is kept to the fore, alongside the different dynamics that play out within access provision. It is at once a way of examining the different accounts shared within the research and navigating the complexity seen, given that this is inherent as a result of the multiple dimensions and categories under consideration (McCall, 2005: 1772). In addition, intersectionality acts as a beacon, as it highlights the way in which injustice and inequality has been embedded within our society: “Injustice is never about just one dimension of being, and therefore is not remediable through a focus on just that one dimension. In such a process, injustice is revealed as far more complex than public discourses would often have us believe” (Block and Corona, 2016: 519).

Faced with the different stances over how diversity and inclusion are negotiated in the light of the pervasive injustice that remains, and how to balance adopting a more universal view without losing sight of the importance of individual needs, it is ultimately this need to account for intersectionality that leads me towards the broader perspective.

2.2.2 Social and epistemic justice

By framing this study through the lens of social and epistemic justice, the principle that everyone should have full access to the world around them is placed at the centre of everything this research entails and ensures that this fundamental dimension of intersectionality is not lost. As Opatow (2016: 41) writes, social justice “is a construct bound up in hopes and ideals” and this dual lens captures the contribution I hope this thesis will make to our knowledge, understanding and practice of access and accessibility.

I consider social justice from the perspective of distributive justice, as originally proposed by Rawls (1971). In considering how resources are distributed and allocated, the question of accessing these resources is immediately raised. Epistemic justice adds a new slant to this distributive focus, which further grounds this study in its ethical intent. Embedded alongside an action research methodology, presented in full in Chapter Four, there is a drive for change and an opportunity to see it being actualised, if only on a small scale, within the lifetime of the project, as well as in the contribution it offers.

In its broadest sense, social justice is concerned with how benefits and burdens are shared within different individuals or groups within society (Clayton and Williams, 2004: 1) and ensuring that people have access to wealth, knowledge, power and privileges. Central to this idea when Rawls (1971) first presented his Theory of Justice was a principle of fairness. He envisaged a society where “each person is to have an equal right to the most extensive scheme of equal basic liberties compatible with a similar scheme of liberties for others” (Rawls, 1971/1999: 53). Such rights included political liberty, freedom of speech, freedom of thought and the ability to hold personal property amongst others. Yet, he also acknowledged that inequalities would still exist in such a

society. He stated that these inequalities could only be accepted if what resulted was ultimately fair for all people: “The basic structure should allow these inequalities so long as these improve everyone’s situation, including that of the least advantaged, provided that they are consistent with equal liberty and fair opportunity” (Rawls, 2004: 75-76).

This measure of fairness remains crucial in discussions of social justice and (in)equality today (see for example, Joyce and Xu, 2019: 16). Increasingly, as well as examining how resources can be better allocated, the question has been raised of how level the playing field actually is. Investigations such as the Grenfell fire (Meding *et al.*, 2017), the Windrush scandal (JCWI, 2017) and the impact of Covid 19 on marginalised communities in the UK and increased international attention on Black Lives Matters (Campbell, 2021) have highlighted the embedded impact of social injustice (Thompson, 2020; Barnard, 2021). Wherever there is a sense of criticism rather than acceptance of the status quo (Collins & Bilge: 30), social justice has the potential to cut across all areas of life and new items on the social justice agenda range from issues of globalism and migration to environmental justice and the impact of climate change (Pearce and Paxton, 2005: xvi-xvii). From an arts and cultural perspective, “the drive for equality has motivated the LGBT community, ethnic minorities, disabled people and those with mental health issues to actively pursue their rights and to seek expression through the arts” (Gillieron and Robson, 2015: 11) and the transformative effect that the arts can have is widely understood. Social justice as a concept has also found its way into principles of design as a solution for resolving the humanitarian crises we face¹⁵ and as an instrumental factor in creating more just and fair societies (Tai, 2009), something which will be relevant as we explore the social model of accessibility in the next section.

Within this thesis, the coupling of social justice with epistemic justice places an additional emphasis on the importance of the distribution and access to ideas and knowledge in society, on how these are created, communicated and understood and on the value assigned to these ideas.

¹⁵ See, for example, the work being done by Architecture for Humanity (www.architectureforhumanity.org).

Fricker (2007) identifies two forms of epistemic injustice and also highlights a form of social injustice which might, mistakenly, be considered to be epistemic injustice. All three are of relevance here. First, she points out that examples of distributive unfairness with respect to epistemic goods, such as information, are examples of social and not epistemic injustice. For example, when respeaking is not provided at events, certain audience members will not be able to (fully) access the content; in terms of distributive access alone, this is an example of social injustice.

The forms of epistemic injustice she outlines involve additional subtleties within the injustice seen. In the first, named **testimonial injustice**, prejudice causes a hearer to give an incorrect level of credibility to a speaker's word (p.1). Most often, injustice will result when a person's credibility is deflated, but the reverse is also possible. Testimonial injustice relates to identity, identity power and the conception of what it is to hold a certain identity. The injustice might be that the police do not believe a person because they are Black (p.1). A person's accent or professional role in a particular situation may lead another to give them more or less credibility than they deserve (p.17-18). The power dynamic between people may also lead to testimonial injustice occurring. In this instance, identity power relates to the power and influence that different groups in society hold. If, for example, there is a situation where a woman disagrees with a man who is present, but feels that it is socially unacceptable for women to speak out against or contradict men, and consequently remains silent, then, according to Fricker, identity power is at play, even if the man has done nothing to instigate or exercise it (p.15). This form of injustice can be active or passive and its cause lies in the collective social imagination – "it can control our actions even despite our beliefs" (*ibid.*). In this sense, it is closely related to the stereotypes and resulting discriminatory norms and biases which we saw earlier in the chapter, which often take conscious interruptions to overcome.

The second form of epistemic injustice is **hermeneutical injustice** and it occurs at a prior stage to testimonial injustice; here, "a gap in collective interpretive resources puts someone at an unfair disadvantage when it comes to making sense of their social experiences" (p.1). This might occur if a woman suffers sexual harassment in a culture

that still lacks the critical concept of what sexual harassment is (*ibid.*); the fact that there is no recognised sense of what it means to be sexually harassed makes it almost impossible to name and respond to the reality of what has been done and why behaviour that, for lack of a name, seems normal and acceptable should not in fact be accepted.

Within this account of epistemic injustice, Fricker focuses on the transactional nature of many epistemic exchanges, meaning exchanges which occur between individuals, and accordingly emphasises the role that the virtue of each person plays in countering both testimonial and hermeneutical injustice. She suggests that an individual's virtue is sufficient to counteract any prejudices they encounter: if they have not taken up the prejudices of the society they live in, this virtue may function naively (Anderson, 2012: 167), but where prejudice is present, anyone who possesses this virtue will be able to overcome their prejudice through critical reflection, a practice which may become more habitual as virtue is trained (p.168).

However, some have challenged Fricker's positioning of this responsibility with the individual (Alcoff, 2010; Anderson, 2012). One reason is the inherent burden that this overcoming of cognitive biases may demand, as a result of the automatic and rapid way in which they function that we saw in section 2.1.1 above. The second reason, which Anderson highlights (2012: 171), is the structural impact that even exchanges between individuals may have:

But in the face of massive structural injustice, individual epistemic virtue plays a comparable role to the practice of individual charity in the context of massive structural poverty. Just as it would be better and more effective to redesign economic institutions so as to prevent mass poverty in the first place, it would be better to reconfigure epistemic institutions so as to prevent epistemic injustice from arising. Structural injustice calls for structural remedies.

As with intersectionality, structural change is called for. Whilst no complete solution is presented, a key step forward seems to lie in the potential of shared inquiry. When social groups are educated together and have the opportunity to co-engage equally in discussions on equality and injustice, the hope is a new shared reality may be created (*ibid.*).

These more subtle layers of epistemic injustice, in particular when considered at a structural level, are particularly relevant to the current study in the context of both accessibility studies and within the sociological turn of audiovisual translation. They call into question the power, authority and credibility that are given to every individual and, within the particular framing of these two disciplines, to the user who is potentially more vulnerable to this injustice. Testimonial injustice potentially questions the credibility of the user as expert. Hermeneutical injustice might occur when organisers of events, through a lack of knowledge or awareness, do not even register the need for access and consequently fail either to provide respeaking or to hold the event in an accessible venue. Here, distributive access to subtitles would be underscored by the more subtle layer of hermeneutical injustice. The need for a mutual respect between respeaker and audience was mentioned above; in the case of epistemic injustice, this respect would be lacking¹⁶. The deeper these injustices are embedded into society, the more widespread they are. For this reason, the social model of accessibility is suggested as a framework for achieving a more inclusive approach to access.

2.2.3 Social model of accessibility

The concepts of access and disability are intrinsically linked. The very symbol we commonly associate with disability is, in fact, the International Symbol of Access (Guffey, 2018) (Fig. 2.5).



Fig. 2.5: International Symbol of Access (*International Symbol of Access*, 2012)

¹⁶ This idea of mutual respect will be presented in 2.3.1.2.

The simplest definition of *access* relates to the act of being able to physically enter a space; from there comes the idea of engagement, participation and being included. *Accessibility* is the related quality and concept and is often used to refer to how products and services are designed so that they will enable this access. Conceptually, the terms *access* and *accessibility* suggest a right or a freedom to take part in all areas of life, yet, accessibility is not a human rights per se. Rather, it is a tool or “proactive principle” (Greco, 2016: 23) through which the entitlement accorded to us through human rights is fulfilled. Access to health (UN General Assembly, 1948: Article 25), does not simply mean that the health service exists; it also means that any individual is able to make use of it and to receive the required treatment and aftercare. Enjoying access to education (Article 26) means being able to join in activities on a linguistic, sensorial, social, emotional and academic level¹⁷.

The social model of accessibility, first introduced by Greco in 2013 and written about more recently (Greco, 2019a; 2019b) is a relatively new proposal for how questions of access can be framed. It was developed with awareness of and in response to two particular models from disability studies; limitations have been raised against both, which Greco (2019a:25) hoped to address through the new perspective and orientation of accessibility studies¹⁸.

2.2.3.1. From disability to accessibility

The first of these two models is the medical model of disability, referred to briefly in 2.1.1 above. Here, it is the individual who is considered to be the problem and the solution or fix for their disability comes through it being “prevented, cured or contained” (Inclusion London, 2015: 6). A normative view of how a person *should* be is therefore strongly embedded within this model and whilst it has now largely been rejected, in favour of the

¹⁷ In the context of this thesis, I am proposing live subtitles created through respeaking initially as a tool for providing sensorial access at live events. However, as we will see, for the event as a whole to be accessible, a range of tools will be required and once provided, the access that respeaking provides is likely to go beyond the sensorial.

¹⁸ As a new model, the amount of detail in print about the social model of accessibility remains limited; however, I have had a number of personal discussions with Gian Maria Greco where he has confirmed details about how the model was developed and the intentions behind it, which I also draw on in this section.

social model of disability, traces of medical-model thinking still remain in the nuances of the language we use and behind decisions taken. As we will see in section 2.3.1.1, this is something that practitioners, researchers and officials must remain alert to, if it is to be avoided.

The social model of disability emerged in the UK in the mid-1970s and remains a model that is frequently adopted for accessibility provision today. It makes a fundamental distinction between impairment, which is defined as physical limitation(s), and disability, which is considered as social exclusion (Shakespeare, 2013: 215). As a result, society becomes the disabling agent, which must change and actively remove the barriers that disabled people, who are now considered as an oppressed group, face. The definition of disability becomes “the disadvantage or restriction of activity caused by a contemporary social organisation which takes little or no account of people who have physical impairments and thus excludes them from participation in the mainstream of social activities” (UPIAS, 1975:4).

This model has been successful politically, instrumentally and psychologically (Shakespeare, 2013: 216-217) as it helped to build a social movement of disabled people, create a positive sense of collective identity, and achieve political and social change which resulted in improved access. This understanding of disability also shaped the *International Classification of Impairments, Disabilities and Handicaps* (ICIDH) published by the World Health Organisation in 1980 (Bickenbach *et al.*, 1999). Yet, the simplicity within social-model thinking has also been its “fatal flaw” (Shakespeare, 2013: 217)¹⁹.

The focus on society causing disability has resulted in the (often very serious) impact of impairment in people’s lives being neglected within the model; in its simplest or most extreme reading, this model could be understood as rejecting the need for any medical intervention. In addition, it is not only this neglect of impairment that is an issue; so, too,

¹⁹ I draw on Shakespeare’s account (2013) of the four weaknesses he identifies with the Social Model of Disability in the account that follows, however, my interpretation of the consequences of these potential weaknesses differs from his.

is the attempt to distinguish between the impact of the impairment and disability, since all too often, both are inextricably linked (p.218). The circularity of this model is also problematic: it functions by assuming, rather than proving that disabled people are oppressed, which by implication frames disability as something negative. However, many people who society identifies as being disabled would not necessarily self-identify this way, a point raised in section 2.1.2.2 above. Finally, the ultimate concept enshrined within the social model of disability, that of a “barrier-free utopia” (p.219), is an ideal, rather than a goal that can ultimately be achieved in practical or resource-related terms. One final model from disability studies is of relevance: the human rights model. As the name suggests, the human rights model takes a legislative perspective and could be considered a model of disability policy rather than disability per se (Lawson and Beckett, 2020: 22). This model states that impairments cannot be used as a reason for denying disabled people any human right, something which the *Convention on Rights of Persons with Disabilities* (2007) legislates as a form of discrimination (Maastricht University, 2019). Some see the human rights model as independent from the social model of disability (Degener, 2014; Maastricht University, 2019), whereas others have focused on how it is complementary to it (Lawson and Beckett, 2020: 22), especially as both situate any problems relating to disability with society rather than the individual. From the perspective of the social model of accessibility, the human rights model shares two notable similarities in its conception: the focus on the “inherent dignity of the human being” (Quinn *et al.*, 2002: 14) that is implicit within human rights and the explicit mention of how impairment is valued as part of human diversity (Degener, 2014: 12).

2.2.3.2 New discipline, new approach

Grounded as it is within the proposed new discipline of accessibility studies, the social model of accessibility reflects three key shifts in thinking at the “ontological, epistemological and methodological level” (Greco, 2019b: 27), shifts that have been noted within the diverse areas that form this larger discipline. These shifts are from

particularist accounts to a universalist account of access, from maker-centred to user-centred approaches, and from reactive to proactive approaches (ibid.).

These three shifts, alongside poetic design, which Greco (2019a:15) proposes as a method for accessibility studies and which I explain below, form the social or atimic model of accessibility. This notion of *atimia*, by which the model is also known, captures the fundamental concept behind it. In ancient Athens, an *atimos* was a person who had been excluded from society, and who had consequently lost the privilege of being able to participate and move freely within it (Manville, 1980). As the concept of what it meant to be a citizen developed in Athens, so, too, were the *atimos* deprived of this status. Embedded within the social model of accessibility is this right to participate and to be an active and embodied citizen. This model functions on a social rather than political level, yet the principle of human dignity seen in the human rights model of disability is present within it. Since accessibility remains an issue of society (Greco, 2018: 211), the driving principle of the social model of disability which requires societal change is also captured within it. The three shifts and principles of poetic design help to illustrate how this model brings a new perspective to the previous models from disability studies and offers a solution to the limitations that have been raised against them.

The first shift is from a *particularist* to a *universalist account* of access. Where the particularist account concerns either exclusively or mainly specific groups of people, usually disabled people, the universalist account concerns all human beings (*ibid.*) and recognises that everyone will have specific needs in specific contexts. There are a number of implications to this shift. Firstly, where disabilities were treated as 'special needs' in the particularist account that needed to be addressed, in a way that was reminiscent of the medical model, in the universalist account, specific needs are situation dependent²⁰. For example, when I lived in Japan, I needed linguistic access, in a way that I do not when I am in an English- or French-speaking country. Within this account, disability is placed within a paradigm of human variation, whereby difference is a natural

²⁰ In this respect, the social model of accessibility resembles Lubet's model of social confluence (Lubet, 2010).

part of life (Stiker, 1999: 12; Greco, 2019b: 28). On a theoretical level, all differences can be accommodated within this paradigm, and the issue of whether the difference stems from impairment, disability or culture, seen in the social model of disability is overcome; barriers to access which relate to any one of them – or more than one in combination – may be imposed by society, and will result in equal status being eroded or negated (Greco, 2019a: 25)²¹. This reframing of how access is accommodated has another advantage. Whereas in particularist accounts, any benefits for people who were not in the targeted access group were a “spill over effect” rather than a planned result (Committee on Disability in America and Institute of Medicine, 2007: 179; Greco, 2018: 218), the universalist account paves the way for benefits among the wider audience to become a research focus, as is the case in this study. In addition, as we will see in section 2.3.1.1, this account also allows research in media accessibility to step beyond the bounds of audiovisual translation and embrace a new range of services (Greco, 2019b: 27). I would also add that this universalist approach is a way of pre-empting instances of hermeneutical injustice from occurring. By adopting this principle of fostering access for all (Greco, 2018: 220; Moores, 2020a), a variety of needs are met; a group that might otherwise themselves have become a ‘particular’ group now find that their specific needs are already being addressed.

The idea that a universal approach to disability and accessibility is needed is not a new one. Zola (1989) talked of the need to universalise a disability policy; the principles of Universal Design, defined by Mace (1988: 1) as the “design of products and environments to be usable by all people, to the greatest extent possible, without the need for adaptation or specialized design”, attempted to embed the concept of universal access more widely by creating built environments that would be accessible for both

²¹ One criticism made of the universalist account is that by speaking of the benefits that accessibility can bring to all, including in the example I gave about my own experience in Japan, there is a risk that the lived experiences of disability and impairment are denied, as the focus returns once more to normative able-bodied people (Elcissor, 2015). Doing that is not my intention. Rather, it is through listening to d/Deaf, deafened and hard of hearing audience members, those with expertise in this area, that I hope to establish a service which first suits their needs and expectations.

disabled and non-disabled users (Hamraie, 2017: xiii). However, objections to this idea have also been raised. For the most part, these relate to a lack of clear criticality and explanation of how the concept of universalism is being operationalised: if taken to the extreme of providing a single solution for all, regardless of individual differences, any good intent in providing access would be negated since a single solution, in all settings, is unlikely to be suitable for all. Both Imrie (2012) and Hamraie (2017) raise such issues and these would also be applicable to the social model of accessibility if it were not for its coupling with poietic design which, as we will see, makes it very clear that responding to individual specificities (Greco, 2019a: 24) is a vital part of the model. A universalist approach does not equate to a single, universal form of access being provided.

The second shift, from maker- to user-centred approaches relates to who is considered as holding the knowledge (Greco, 2018: 212) in the quest for accessible solutions. Accessibility has challenged the attribution of expertise that was previously accorded to makers alone, and foregrounded the knowledge and experience that users bring, which must be taken account of in the design process (*ibid.*). This is strongly allied to the testimonial value and justice referred to in section 2.2.2 above and to a sense of composite expertise, which is fundamental to this research, and built through knowledge that the multiple focus groups in this study bring. As Greco explains, this shift does not negate the expertise that makers bring, but instead encourages a “co-construction” (2019a: 24), where multiple parties work together to shape the design process.

The third shift, from *reactive* to *proactive* approaches is for the most part a procedural one, yet it also allows the universalist account to be more fully embedded within the access provision offered, since consideration can be given to how different groups will make use of the services offered. Here, rather than accessibility being actioned at the end stages of production, almost as an afterthought, it is moved to an earlier stage in the process, ideally the initial design phase (Emiliani, 2009: 2.6; Greco, 2018: 213). As well as enabling more thought to be given to how best to embed the access, this shift also encourages a critical reflexivity which is in line with the other theoretical and methodological approaches adopted in this study and central to this model.

Poietic design is presented by Greco (2019a: 24) as an access-oriented methodology which embeds the fundamental features of accessibility within the design process. In particular, it enshrines the individual and collective responsibility that everyone holds within the process, since it has become one of co-construction (*ibid.*; Greco, 2019b: 31-32) within this process. Six basic principles of Poietic Design are presented (p.24-25). The first two principles relate to the first shift: they are the principles of universality and personalisation. Despite the fact that access concerns us all (1), one size is not expected to fit all (2). Instead, following principle two, it is expected that the design should respond to the “specificities of individual users” (p.24). Principles three, four and five relate to the maker-user shift. The principle of user-centrality (3) is that design focuses on users and their specificities. The principle of epistemic inclusivity (4) states that all stakeholders bear valuable knowledge and the principle of participation (5) reinforces the idea that stakeholders should be actively involved in the design. The final principle (6), that of proactivism, restates the third shift.

This reconfigured approach to access and design inherent within the social model of accessibility offers novel solutions to the criticisms raised against models which preceded it. Whilst it is a model of inclusion, as the principle of universality suggests, it encapsulates an approach where there is both an equality and equity of access (Moore, 2020a: 180-181). Everyone has an equal right to access, yet, via the principle of personalisation (2), equity ensures that the way the access is operationalised is fair and fit for purpose (Mann, 2014), hence the principle of personalisation (2). Difference, in need, experience, knowledge and context is at the heart of this model and it is something that is valued. As with intersectionality and social justice, this sets the social model of accessibility at odds with diversity when it is used simply as a “happy point for inclusion” (Ahmed 2012: 14), or as a tool for concealing systematic inequalities at play (p.53). This model embeds the responsibility and agency of access that we each have within the process of access provision and demands that a critical stance is taken to ensure that all efforts are made to avoid its principles being violated. It is also understood that there is an inherent complexity within access provision that must be addressed if accessibility

is to be successfully achieved. Greco *et al.* (2012)²² visualised this as the chains that run through the chain mail that a knight might wear, where the strength and protection this armour provides comes from the integrity of its structure. Once a link breaks, damage may be done. I have adopted this as the idea of the chain of access, which must run unbroken through every step of the process.

2.3 Media accessibility as a bridge between audiovisual translation and accessibility studies

2.3.1 The scope of access and translation

In the previous sections, much of the discussion has surrounded issues relating to access and accessibility, yet even within this, examples have been given of language- and translation-related issues. Whilst I see audiovisual translation and accessibility studies as two distinct academic disciplines, there are also many areas of overlap between them. Just as access is “a central concept in human life” (Greco, 2019a: 16), so, too, is language and by extension translation.

As I define it, translation involves both a product and a process (Munday, 2012: 8). The product may take various forms, including written text, spoken words, visual images; in many instances, and certainly in audiovisual translation, the product is likely to be a multimodal combination of these elements. The process of translation is similarly diverse. Whether intralingual, where the source and target language are the same, or interlingual, where they differ, it is very often intersemiotic, traversing different modes of communication. Respeaking, which moves from the spoken to written word is a key example of this. Whilst I do see translation as a linguistic process, this linguistic element cannot function in isolation. Translation must necessarily move beyond language if it is to fully capture the vital and complex whole that is created when different media and modalities interact within specific geographical, social, cultural and historical contexts.

²² Greco *et al.*, 2012 is the first published account of the chain of access and I have also had numerous discussions with Gian Maria Greco about its concept and application.

Content and meaning cannot be transferred without questions being raised about power, privilege and access and, over the years, these debates have shaped the ideologies and approaches that translators have adopted.

Considering translation as “ubiquitous” (Blumczynski, 2016:168), at once a process and a product of its own situation, captures this broader sense of translation and aligns the concepts of translation and access: “Translation has a philosophical, theological, linguistic, anthropological, political, historical, social and ethical dimension (as well as many others); by the same token, all these (and numerous other) fields involve and vitally depend on translation, which is a matter of the HOW rather than just the WHAT”.

In different terms, Marais (2014:11) conceptualises translation as a complex phenomenon that both emerges from and is involved in the emergence of other social phenomena.

Translation and access are not one and the same, yet areas of overlap clearly exist between them and each can usefully inform the other. Nevertheless, the distinctions that exist between these two disciplines are as important to note. Just as the social model of accessibility offered new insights on issues that were previously tackled from within disability studies, accessibility studies, with its distinct perspective, offers a new orientation towards research, which is not usually found in audiovisual translation, or translation studies, alone.

2.3.1.1 Meeting points and divergences between media accessibility and audiovisual translation

Greco (2019a) aligns the shift from particularist accounts to a universalist account, not just to a new understanding of access and accessibility, but also to the emergence accessibility studies in its own right (Fig 2.6.).

The Three Accounts of Media Accessibility

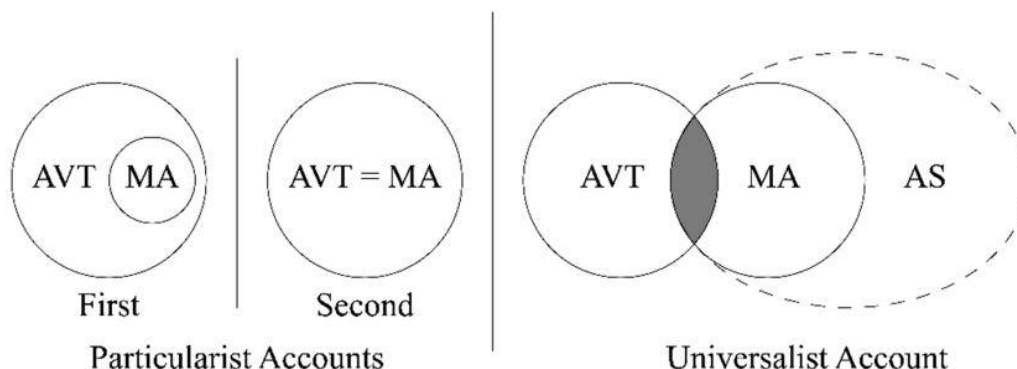


Fig. 2.6: The three accounts of media accessibility (Greco, 2019a: 19)

SOURCE: Reprinted by permission from Springer Nature Customer Service Centre GmbH : Springer International Publishing, Accessibility Studies: [Abuses, Misuses And The Method Of Poietic Design](#) by Gian Maria Greco, in Stephanidis, C. (eds) HCI International 2019 – Late Breaking Papers. HCII 2019. Lecture Notes in Computer Science, vol 11786. Springer, Cham © 2019

From the perspective of translation studies, media accessibility, with its initial focus on providing sensorial access through subtitles and audio description, was born as a subdomain of audiovisual translation (Greco, 2018: 211). Gradually, as the diverse ways in which these modalities could be used was recognised, for example as a means of providing linguistic access to visual and aural content, media accessibility often came to be considered as synonymous with audiovisual translation. However, this view conceptually restricts the range of accessible services that can be offered to those related to the media, and more specifically, services that offer some form of translation (Greco and Jankowska, 2020). Yet, from an access perspective, especially one which is proactive and user-centred, it should be a very easy step to move from creating audio description (translation-based) to writing and including an audio introduction (non-translation based).

In naming accessibility studies as a discipline in its own right, Greco proposes an alternative scenario whereby contributions from a range of fields, of which one is media accessibility, meet and create a diverse, yet centralised whole (Fig. 2.7).

The Formation Process of Accessibility Studies

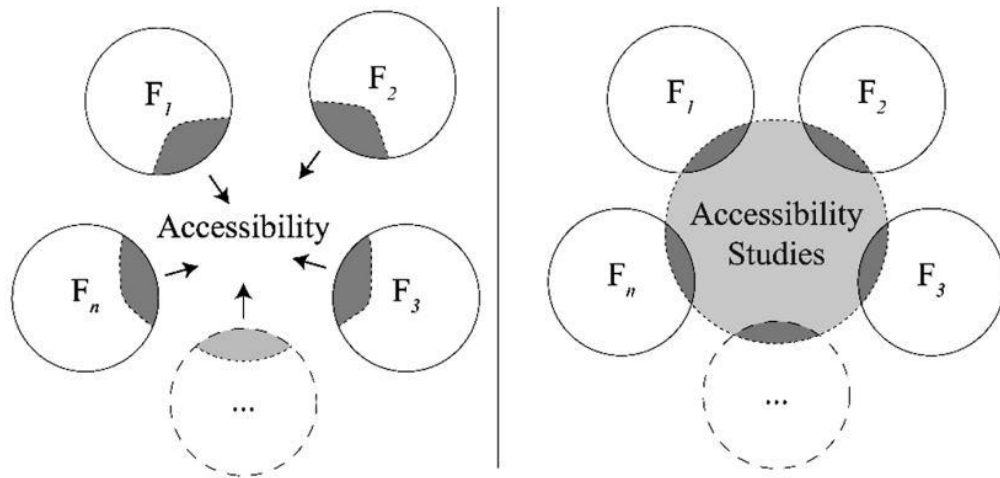


Fig. 2.7: The formation of the field of accessibility studies (Greco, 2019a: 22)

SOURCE: Reprinted by permission from Springer Nature Customer Service Centre GmbH : Springer International Publishing, Accessibility Studies: [Abuses, Misuses And The Method Of Poietic Design](#) by Gian Maria Greco, in Stephanidis, C. (eds) HCI International 2019 – Late Breaking Papers. HCII 2019. Lecture Notes in Computer Science, vol 11786. Springer, Cham © 2019

This conception allows media accessibility to be redefined as concerning “access to media and non-media objects, services and environments through media solutions, for any person who cannot or would not be able to, either partially or completely, access them in their original form” (Greco, 2019a: 18).

Whilst this distinction is a subtle one, it is nevertheless significant. My study is one that begins in and is informed by translation studies, and more specifically audiovisual translation, yet extends beyond it. The theoretical stances it adopts, the way that both media and non-media objects and translation- and non-translation-based solutions are implemented take it beyond into the realm of accessibility studies. Rather than focusing too closely on the boundaries of either discipline, I will instead focus on how drawing from both disciplines can enhance this study.

Within audiovisual translation, four key turns have been noted: the descriptive, the cultural/ideological, the sociological and the cognitive/empirical. The final two are of particular relevance to this study and relate closely to the three-fold shift noted in

accessibility studies (Chaume, 2018). Within the sociological turn, increased consideration was given to the role of the translator and to where power lies in the process of creating audiovisual texts. Alongside this, increased awareness was placed on the active role of audiences and, in line with the maker-user shift within accessibility studies, producers and distributors had to take account of the “social, co-participating and unrestricted nature of this new scenario” (p.52). The expectation was that contents were produced with rather than for the audience, who were elevated from consumers to ‘prosumers’. Within audiovisual translation, the sociological turn captured a spirit of activism, as this new audience understood that translation was no longer simply a product they received, but that it could be a force and tool of power, activism and even of revolution (Boéri and Maier, 2010; Tymoczko, 2010). Many of the methodological approaches I adopt in this study also entered audiovisual translation through this turn, including focus groups, questionnaires and mixed methodologies.

The current cognitive/empirical turn has led to a closer consideration of the mental processes of the translator and the audience’s response (Chaume, 2018: 53-54) and this is reflected in the experimental design of this study. Reception studies are used to uncover the likes and dislikes of the audience, and action research, which was the precursor to this turn (as in Neves, 2005), drives the study.

2.3.1.2 The notions of vulnerability, respect and trust

Alongside these turns, Díaz Cintas and Remael's framing (2007: 55) of subtitles as a form of vulnerable translation is highly relevant to this study. This vulnerability stems from the difference that exists between the spoken original and written translation as a result of the time and space constraints inherent in the process of subtitling, which may be highly visible to certain, if not all, members of the audience - depending on the range of their access to the original content.

In respeaking, this situation is compounded; as a comparatively new modality within audiovisual translation and media accessibility, its production method is not always known or understood; further, the subtitles produced are subject to the added constraint

of real-time production; for this reason, respeaking is even more susceptible to audience scrutiny. At the same time, just as it is a potentially vulnerable form of translation, its target users, d/Deaf, deafened and hard of hearing people, could also be considered 'vulnerable' in the sense that many depend on the subtitles for access; their vulnerability might come from the idea that the access they receive differs from the original content, either as a result of the limitations imposed by the time and space constraints or because of the inaccuracies, be they obvious or potentially hidden, that are associated with live subtitling in particular.

All too often, vulnerability may lead to increased marginalisation. In her study of *Music and Translation*, Desblache (2019: 326) reflects on the pathway that media accessibility usually takes, using a musical metaphor of major to minor:

Media companies hence translate programmes for the Deaf and Hard of Hearing and for the Blind and the Visually Impaired [sic], but very few events created or enacted by the latter are offered for the hearing and the sighted. Voices that are considered to be minor have gained the right to see and hear, but not to be seen or heard.

Dangerfield (2018, 2022) takes this further, exploring through her work on accessible filmmaking, the different forms of access that are required for disabled people to have access to content creation, and highlights the variations in modalities of media accessibility that may be required for this to be achieved.

While most of the presenters at the research events in the current study are those of the 'major', in the dynamic setting of live events, the voices of everyone present must be heard; communication is a two-way process. The model for participatory engagement, presented at the end of this study, explores how the diverse needs of all who attend can be considered.

With this in mind, where translation studies has often drawn on discussions of fidelity and loyalty when discussing the relationship of the translator between source and target texts and sides (Nord, 1997: 125), in the current study, the forging of a principle of mutual respect between the translator (respeaker) and audience seems far more appropriate. A shared communication and shared awareness of the others' needs, understanding and

expectation is vital for respeaking and access to be provided and to be made use of. If mutual respect is present then it is hoped that this will also lead to an underlying trust that the respeaker, and indeed any access provider, is doing their best work, given the inherent constraints that exist within live subtitling, something explored further in Chapter Three.

2.4 Drawing together the theoretical threads

In the current chapter, I have sought to draw attention to the inequalities that exist in society, in particular with respect to questions of diversity, access and inclusion, and I have attempted to highlight the complexity that exists within our use of labels. Such a simple and useful tool so often conceals the diversity and fluidity that exists within so-called groups, yet as I write I cannot avoid their use. This conundrum remains throughout the chapters that follow.

The triple theory approach that underpins this research has also been presented, an approach which combines intersectionality, social and epistemic justice and the social model of accessibility which together seek to address these societal inequalities on a practical, provision-based level, and also with regard to more ethical considerations. These different strands mean that my work is interdisciplinary in its nature, as it straddles the fields of AVT, MA and AS.

At the heart of the work is a principle of mutual respect, for those who provide the service of respeaking and those who seek to use it. By listening to the voices of all involved, and seeking to meet the expectations of diverse audience groups, it is hoped that such a service can be created for use in the UK and further afield. Since the live event setting is a largely unregulated area, there might be a flexibility here that is not present on television, to explore how a respeaking service could be embedded in a way that contributes to wider issues of accessibility, whilst complementing this triple theory approach. At the same time, the technical considerations inherent in respeaking mean that, especially in such a dynamic situation as a live event, a suitable implementation of the practice must also be negotiated.

With this in mind, in Chapter Three, the modality of respeaking is examined in more detail and the gap that this research is intended to fill is presented.

Chapter 3: Respeaking in a new setting²³

“Respeakers... write with their voices
so that deaf and hard of hearing viewers
can listen with their eyes.”

Romero-Fresco, 2011: 178

In Chapter Two, the theoretical background and drive behind this study was discussed and the approaches of intersectionality, social and epistemic justice and the social model of accessibility were outlined. I situated this study within both audiovisual translation and media accessibility, and the larger domain of accessibility studies, which is fundamental to the work this study involves.

In the current chapter, I explore the gap that exists within AVT and MA that this study begins to fill and explore exactly what respeaking is and what it involves, so that the challenges involved in its introduction into the live event setting in the UK can be better understood. In doing so, I consider the traditional audience that respeaking was originally conceived for and the wider audience whom I believe may also benefit from it. I present a number of issues relating to the reception of subtitles and situate this study within existing research that has been carried out into respeaking.

3.1 What is respeaking?

‘Respeaking’ is the production of subtitles in real time by a human using speech recognition (SR) software and it was introduced to UK television in the early 2000s (Lambourne, 2006; Romero-Fresco, 2011, 2018). Intralingual respeaking, the focus of this study, is used to provide access for live television programmes, including the news, sports and special events. In the UK, it is done by a single respeaker, as per the mono-live subtitling model of Remael *et al.* (2016), and the process is divided into three key stages. First, the respeaker listens to the broadcast content and speaks the aural content of the programme, voicing in punctuation, sound labels and any additional content that

²³ Some passages in this chapter draw on Moores, Z. (2014).

needs to appear in the subtitle. As they do this, they may edit the original spoken content slightly, perhaps adjusting the order or punctuation of the original, or omitting or paraphrasing certain words (Eugeni, 2006; Lambourne, 2006; Romero-Fresco, 2011). Next, the speech recognition software processes the input. Finally, the recognised utterances pass through subtitling software (SS) and the respeaker is able to make further, slight, adjustments to the subtitles as or after they appear on screen (McIntyre *et al.*, 2018) either by voice or by typing in a correction. Even while the second and third stages are happening, the respeaker must continue with step one, as the audio of the broadcast continues, determining the speed of respeaking required. Although the eventual output will be read by a viewer, the intermediary of speech recognition software throughout the process demands that the words spoken conform to the capabilities and limitations of the speech recognition tool in use, rather than the human ear. A carefully prepared voice model and good audio are essential for respeaking and setting these up could be considered as part of the process of respeaking (Moores, 2014: 20). If accurate subtitles are to be produced, intense concentration is also required. Where sections of a programme are scripted and the respeaker has access to them in advance, for example in the news, the respeaker is able to prepare subtitles to cue out as-live, or semi-live (Romero-Fresco, 2011: 12) rather than voicing them in.

On television, a distinction is made between the live subtitles that are produced through respeaking as the programme is broadcast, including the quickly prepared semi-live subtitles, and the pre-recorded or pre-prepared subtitles that are created for programmes such as documentaries or dramas, which are prepared in advance of the programme going to air. At live events, the distinction is slightly different and focuses more on the type of content in question. Stagertext, the key provider of such access and the charity with whom I have been working most closely, use the term 'captions' to refer to the access they provide at theatre performances. These captions differ from 'surtitles' which offer translations of the dialogue and lyrics at the opera, musical performances and theatre plays as captions are intralingual and contain added information for DH audiences. Both are prepared in advance (Díaz Cintas and Remael, 2007: 253), but the

way they are displayed may also vary (Mele, 2018). Stagertext's captions are cued out live by a captioner who is usually present onsite. 'Live subtitles' is the term used to refer to the access to unscripted speech that Stagertext offers at live talks and tours, a service that for the most part is provided through speech-to-text reporting (STTR)²⁴.

In this thesis, as a result of my close work with Stagertext, I have followed this same use of terminology and I use 'live subtitles' to refer to the subtitles used for unscripted content at live events, as well as those created in real-time on television. I chose not to adopt the term 'speech to text interpreting' (STTI) to describe live subtitles in non-television contexts, even though it is a term that is often used (Stinson, 2015: 399-400; section 3.3 below).

Similarly, whilst on other occasions I have chosen to use 'titles' rather than 'subtitles' to reflect the varied positions and display units where respoken output may be displayed (see, for example, Moores, 2020a), on this occasion I have not. I prefer the simplicity and clarity that 'live subtitles' offers, drawing on the generic usage of 'subtitle' to refer to text in any position on or off screen that is true in academia as well as in popular usage. As figure 3.1 below illustrates, these decisions also allow 'speech-to-text' and 'live titling' to act in wider translational contexts as umbrella terms for a range of different modalities including respeaking, keyboard-based methods such as stenotyping and velotyping, automatic speech recognition (ASR), and a host of interlingual live subtitling (ILS) modalities (Romero-Fresco and Alonso-Bacigalupe, 2022), which range from human-centric to fully automated (Davitti *et al.*, 2020; Romero-Fresco and Alonso-Bacigalupe, 2022)²⁵.

²⁴ When this access was first introduced at talks and tours, it was often advertised as, for example, "a lecture with speech-to-text transcription", as a Google search for 'Stagertext speech to text transcription' will reveal. On their new website, launched in November 2021, the terms 'live subtitles' and 'live subtitling' are used, with the explanation that they are "produced live by a speech-to-text reporter" (Stagertext, 2021c). They make one further distinction and use the term 'digital subtitling' to refer to subtitling they create for pre-recorded videos.

²⁵ Use of the QWERTY keyboard also falls within STT. Whilst use of this regular keyboard alone is usually too slow to achieve the speeds required for real-time access, it is often used in combination with speech recognition software during respeaking.

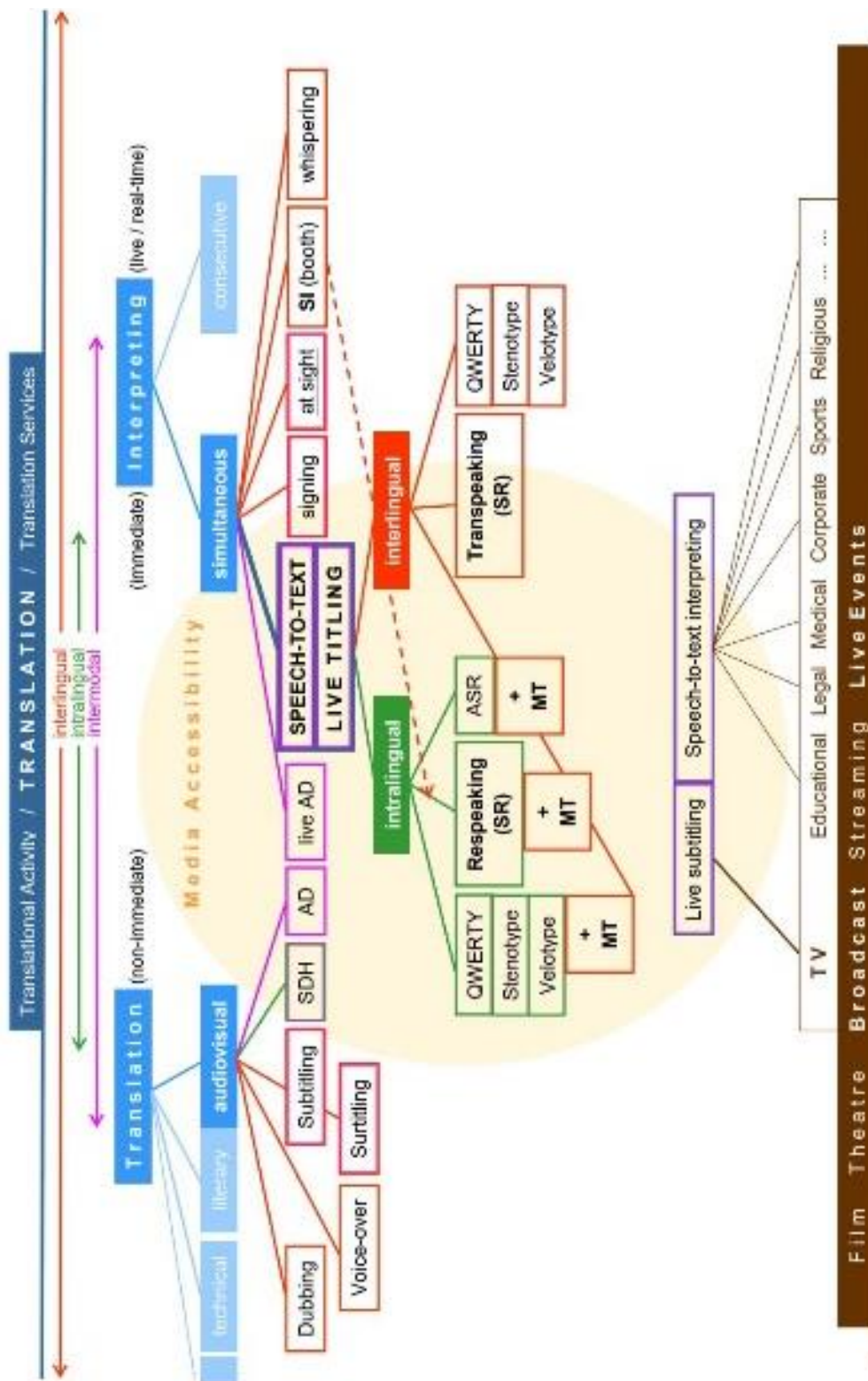


Fig. 3.1: The ILSA (Interlingual Live Subtitling for Access) project's terminological map of translational activity²⁶

²⁶ This map has not been published yet, but forms part of the Interlingual Live Subtitling for Access (ILSA) project reports and an extract of it appeared in Pöchhacker, 2020. ILSA is an Erasmus+ project, 2018-1-DE01-KA203-00; more information can be found at ilsaproject.eu.

One further clarification is required: whilst both ASR and respeaking now depend on automatic, speaker-independent speech engines, respeaking relies on the combination of a human working with the speech engine, whereas ASR does not²⁷. For this reason, ASR is a fully automated form of machine-based translation, whereas respeaking is more accurately described as computer-aided translation (Romero-Fresco, 2018: 96), since the role of the human translator is central to it.

3.2 Working with speech recognition software

All forms of subtitling involve the diamesic transfer of spoken language into written form (Gottlieb, 2012: 37). However, in respeaking this transfer is two-fold. In addition to thinking about how the spoken input of the programme or event must appear as written text, the respeaker must also be aware of their own (spoken) interaction with the SR software and subtitling or presentation software through which the recognised text passes before appearing on screen for the viewer to read.

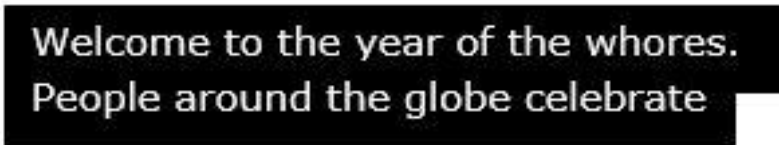
3.2.1 Converting spoken to written English

Speech involves the production of a continuous stream of sounds (Roach, 2009: 31), which for practical, descriptive purposes are divided into segments, or phonemes. In English, 49 individual phonemes are recognised (Wells, 2008; Roach, 2009), yet the reality of speech production is that a “practically infinite” number of variations exist in the way that each phoneme is produced by different speakers, and by a single speaker in different situations (Roach, 2009: 31). When words are spoken, slight alterations occur in each phoneme according to how the surrounding sounds are articulated. Individual speakers are likely to articulate in a particular or idiosyncratic way and voices vary according to accent, age and gender. During connected speech, when words are spoken in a continuous, or semi-continuous flow, there are many processes at work and the way

²⁷ In speaker-dependent SR systems, users train the system to their own generic speech pattern, whilst in speaker-independent systems, the user’s voice is matched to generic voice patterns (Davis, 2018). Older versions of Dragon were speaker-dependent, but from version 15, a speaker-independent system based on Deep Neural Networks is now in use. This is explored more fully in section 3.2.2 below.

a word is articulated varies considerably from its so-called dictionary or citation form. All of these impact on the ability of SR software to recognise the word being uttered (Bowker, 2002: 31). Elision or liaison occurs when words run together and the position in a word where stress is placed, the rhythm of speech and the intonation or pitch used all affect pronunciation (Roach, 2009). The degree to which this affects different words will depend on the part of speech they perform (Shockey, 2003: 15). The respeaker must be alert to all of these processes in order to interact with the SR software successfully.

The written form of certain words in English also poses a problem for the respeaker to work around. Homophones, such as 'to', 'too', 'two', sound the same, but differ orthographically; near homophones can be equally problematic, and harder for the respeaker to spot: 'horse' and 'whores' provides a well-known example which reached headline status (Fig. 3.2).



Welcome to the year of the whores.
People around the globe celebrate

Fig. 3.2: Recognition error captured on BBC News, adapted from the screenshot cited in Jivanda, 2014.

Similarly, in certain circumstances particular words or phrases need to be capitalised. The respeaker must also understand how to manipulate the SR and any subtitling software being used to achieve this.

3.2.2 Dragon by Nuance

Dragon is currently the speech recognition software of choice for intralingual respeaking in English²⁸. When the first version, named DragonNaturallySpeaking (DNS), appeared in 1997 with its 23,000-word vocabulary (Schuman, 2014: 30), the fact that it was able

²⁸ As confirmed through personal conversations with a number of respeakers and managers at the main access service language providers in the UK.

to recognise continuous, connected speech (if not completely natural speech), marked a real breakthrough, especially since attaining continuous speech recognition had been an elusive goal since the 1970s (Juang and Rabiner, 2005: 9). Fifteen different versions have been released since then, and the software was rebranded as Dragon Professional Individual v15 (DPI) in 2016 to mark the changes which have occurred.

In previous versions of Dragon, a respeaker had to select their age, gender, the variety of English they spoke and their general accent before setting up their profile (Moores, 2014: 28). In the latest version, it is simply necessary to choose your region and accent²⁹. Where earlier versions of Dragon required users to complete a brief training programme, in DPI, this has been reduced to a short text of just over 100 words so users can achieve very high out of the box accuracy. The more it is used, the better accuracy becomes (Nuance Communications, no date).

Dragon, like all modern SR software, employs three basic speech models (Juang and Rabiner, 2005; Romero-Fresco, 2011: 57-63). In the first **acoustic model**, analogue sound is converted into phonemes. Here, noise, hesitations and changes in volume are removed and a digital representation of the phonemes uttered is created.

The second model is the **grammar or vocabulary model**, and it is here that phonemes are converted into words. It is vital that the respeaker is familiar with the words that appear in Dragon's vocabulary, since only these words can be recognised. Any out-of-vocabulary word uttered, no matter how clearly it is articulated, will be misrecognised for one or more words in the list. New words must be added into the vocabulary and trained. Even with in-vocabulary items, initial recognition is not clear-cut and possible words are determined and given a probability-based confidence score (Moores, 2014; Romero-Fresco, 2011:58). Before the acoustic model makes its choice, a final layer of processing is used to further refine what is recognised.

²⁹ Rather than regional accents from within the UK, the options in DPIv15 are standard, Australian-accented English, Indian-accented English, Southeast Asian-accented English and Accent not otherwise specified, which reflect the larger pools of speech patterns that Dragon is built on. For this reason, close to standard pronunciation in any of these accents is likely to be most easily recognised by Dragon, which may mean that respeakers need to alter their natural accent to use it effectively.

This final layer is the **language model**. Here, consideration is given to the context present within entire phrases to determine whether a word recognised by the acoustic model is plausible. In DragonNaturallySpeaking, n-gram Hidden Markov Modelling (HMM) (Romero-Fresco, 2011: 59) was used to refine the probability calculation. Within each utterance, every word is in turn analysed as the focus, and words that come before and after this focus word contextualise recognition. BestMatch IV and V were possible within DNS, providing quad- or quintgram capability³⁰. This means that if a nine-word phrase were to be uttered, each word could be compared to determine recognition, increasing the accuracy potential. According to Romero-Fresco (2011: 59), where language models use quadgram-HMM, there is the potential to reduce the error rate by a factor of four, raising accuracy from 80% to 95%. As we will see in Chapter Eight below, recognition is a vital factor when assessing respeaking quality and any within-software improvements such as these have the potential to lead to improved user reception of live subtitles.

DPI marked a new step in the processing power of the Dragon software, with its Deep Learning speech engine (Orcutt, 2016; Nuance Communications, no date). Here, HMM technology is replaced with Deep Neural Networks, which are able to work at faster speeds, process information across multiple layers (Chandra, 2018) – in this case, the acoustic and language models specifically (Orcutt, 2016: 1) – and learn continuously, even during dictation (p.3). As Orcutt explains (p.1), these networks allow information about word frequency and combinations to be processed at the same time as the acoustic information is deciphered. In addition, during each session the engine performs a “fast adaptation” of its acoustic model based on only a few seconds of speech, enabling it to adapt to how a user’s voice sounds in the moment and any variations which may be caused by them, for example, having a cold, using a different microphone or by any changes in environment” (*ibid.*)³¹.

³⁰ The exact algorithm employed depended on the length of the utterance, the version of Dragon in use, and the processing power of the machine (Orcutt, 2014: 3).

³¹ Further offline adaptation occurs and texts which the user has created can also be inspected to expand the active vocabulary in use (Orcutt, 2016: 1.).

These changes are likely to mean “noticeably improved accuracy” for anyone with “an accent, speech impediment or specialised job (like court reporting with the steno mask)” (p.3). In the case of this study, when respeakers will be working in changing environments, using the masks that Orcutt refers to and have limited time to train their voice profile, this new version of Dragon is particularly timely³².

Nevertheless, despite these improvements in processing and recognition that version 15 brings, respeakers must continue to adapt their voices to Dragon, speaking in Respoken English (Moore, 2014) or software-adapted speech³³, rather than truly naturally-spoken English. In addition to changes in articulation, this way of speaking also involves inserting additional pauses during speech to allow chunks of recognised output to be released and developing listening techniques which are conducive to respeaking (Moore, 2014)³⁴.

3.2.3 Tools available to the respeaker

A number of tools are also available to respeakers to support them in their interaction with Dragon and the software used to display their respoken output to the viewers. A range of display software exists for this purpose, very often developed in-house to meet the individual specificities of company demands and workflows. In this account, I make a general distinction between **Presentation Software**, which provides an interim screen

³² However, as Orcutt suggested in a personal conversation, enabling Deep Learning may have limited the ability to dictate directly with Dragon into nearly any non-Dragon friendly environment, so the use of Dragon with applications other than DragonPad and Word may be affected. It is possible that this shift accounts for the technical errors seen during respeaking at some of the live events in this study, discussed in section 6.6.2.1 and Chapter Eight below. KnowBrainer’s Dragon Capture may offer a solution to this and is worthy of future research.

³³ Software-adapted speech is term adopted in the SMART Project at the University of Surrey. SMART stands for Shaping Multilingual Access through Respeaking Technology and this ESRC-funded project, ES/T002530/1, is led by Dr Elena Davitti. More information can be found at smartproject.surrey.ac.uk.

³⁴ It was only after the inherent latency within earlier versions of Dragon was overcome that it became the software of choice for respeaking (Romero-Fresco, 2011: 65-66). Nevertheless, a pause is still required before words are released. Some in-house subtitling software pulls the input from Dragon directly before it would naturally be released, reducing the frequency with which respeakers need to pause and thereby reducing latency by enabling subtitles to be transmitted more quickly (Moore, 2014). Since this interrupts the functioning of Dragon’s language model, it is likely that this does impact on the potential recognition accuracy that Dragon can achieve, but, to my knowledge, no published data on this is available to the public.

where respeakers can view the SR output and make direct edits to it before it is displayed on screen, and **Subtitling Software**, which includes increased functionality, for example, the ability to apply additional corrective layers to the output from the SR. As will be explained, this distinction is important when respeakers are asked to transfer their skills to the live event setting, as they may have to adapt familiar techniques to the functionality and capability of a new piece of software.

3.2.3.1 Editing Dragon's vocabulary

As explained in 3.2.2 above, it is vital that the respeaker is familiar with the words that appear in Dragon's vocabulary. They are able to edit this vocabulary, adding and deleting words as needed and training individual words to support the recognition of them as they are uttered. This is likely to form a key part of a respeaker's preparation.

3.2.3.2 House styles

Some subtitling software will allow the respeakers to use house styles, which function in a similar way to Word's 'Find and Replace'. When used correctly, they can significantly improve the accuracy of a respeaker. A specific word or phrase from Dragon's output can be entered as a match term (Find) and replaced with a different form. House styles may only need to be applied in certain contexts, so it is likely that a respeaker would set up a number of different house-style filters for particular genres, which can be activated and deactivated as needed. Examples of house styles applied during the weather might include:

Match	Replace
Apache	patchy
Chile	chilly
chilli	chilly
son	sun

Similarly, house styles might be used to ensure that the capitalised form of words appear when needed:

Match	Replace
north	North

3.2.3.3 Macros

Spoken voice or customised commands, called macros, are another tool that is available to the respeaker. Here, a new spoken form may be assigned to a word to improve the likelihood that it is recognised correctly by Dragon; very often this spoken form includes the word 'macro' to give it a distinct acoustic pattern.

Written form	Spoken form
!	ex-macro
--	dash-macro
...	dot-macro

Macros may be used to ensure that punctuation is recognised correctly, since it is often spoken at speed. A macro may also be created for words which are frequently misrecognised. For example, I set up the following macro, as I could not rely on 'rural' being recognised accurately:

Written form	Spoken form
rural	country-macro

I chose 'country-' as a tag as it was easy to pronounce and contextually linked to 'rural', making it easy to remember.

Generic macro codes can also be created, which are easily adapted to the specific task a respeaker is working on. For example:

Written form	Spoken form
	man-macro
	woman-macro
	place-macro

Here, written forms would be inserted and replaced task-by-task, but the spoken-form system would remain the same, making it easy for the respeakers to recall and apply. If the respeaker is using subtitling software, it is possible that this list could be updated within it and applied as a house style; if the respeaker is using presentation software alone, the spoken forms would need to be inserted directly into Dragon's vocabulary.

3.2.3.4 Live correction

Whilst house styles and macros allow the respeaker to pre-empt many errors which might occur and make pre-emptive corrections (Moore, 2014: 32), some live correction will also be needed. These corrections may be spoken or typed and, in addition to the delay within which the error is realised, the functionality of the subtitling or presentation software will determine whether these corrections are made before or after the text is displayed to viewers. The software may also incorporate additional corrective tools, such as temporary macros³⁵. A double dash (--) is the usual notation to flag corrections that appear on the viewer's screens, a few words after the error itself³⁶. Any correction a respeaker makes will increase the delay with which the viewer receives the subtitle; for this reason, respeakers aim to produce as few errors as possible (Moore, 2014) and when errors do occur, it is likely that the respeaker will focus on correcting those errors which are likely to impact most on the viewer's comprehension of the subtitled content.

³⁵ Temporary macros may be used if an out-of-vocabulary term is spoken which is likely to be used many times while the respeaker is live. The respeaker would enter it into the subtitling software to use for the duration of the broadcast as it would not be possible to enter a permanent macro in Dragon while live.

³⁶ During live events, especially when many lines of subtitles are being displayed, delayed corrections may occur. The ability to do this may depend on the subtitling or presentation software in use. In this instance, the erroneous text would simply be replaced with the correct tense, and the – would most likely not be used.

3.2.3 The effort involved in respeaking

Despite all the tools available to the respeaker, the process of respeaking remains a challenging one, especially when the respeaker is considering at once what is being said by the original speaker and how to communicate this to the software in an appropriate way and to the audience whilst respecting constraints related to speed, content and supportive display and segmentation.

A number of studies indicate that the genre of the programme, and pace of its delivery, will impact both on the nature of the editing required (Van Waes *et al.*, 2013: 39-40) and the cognitive load of completing this task.

Szarkowska *et al.* (2016) defined cognitive load through a series of parameters – the difficulty of the respeaking task, the occurrence of temporal and mental demands, and frustration and engagement with the task - and explored this in intra- and interlingual respeaking. They found that it was the parameter of difficulty that related most closely to the notion of mental load, which itself could be regarded as an estimation of what cognitive load a task required (Paas *et al.*, 2003; Szarkowska *et al.*, 2016: 227). Whilst interlingual respeaking was for the most part found to be more demanding, one parameter overrode this finding, the pace of the programme. In particular, frustration levels seemed to correlate with this parameter, with higher speech rates leading to a higher level of frustration being experienced by the participants (p.229).

Szarkowska *et al.*'s study also highlighted moments of “respeaking crisis points” (RCPs) (p.231), which were defined as “moments when respeakers’ performance was at its lowest and where frustration was the highest” (*ibid.*) or “potentially problematic moments in the respeaking process, resulting from the difficulty of the source material and/or cognitive overload on the part of respeakers” (Szarkowska *et al.*, 2017: 179). To determine what these RCPs were, the focus was placed on concentration and frustration as indicators of cognitive load (p.182). Electroencephalograms (EEGs) were used to capture moments where peaks in both measures occurred to identify the RCPs. A number of features were found to trigger these crisis points including both very slow and very fast speech rates, the visual complexity of the material, overlapping speech,

numbers and proper names, speaker changes, word play, syntactic complexity, and implied meaning (p.197).

They also noted, however, that what triggers a crisis point may vary between individuals and that rather than leading to a drop in respeaking performance or the arising of actual problems, such points may equally trigger the adoption of a particular strategy so that the respeaker can continue performing at the same level.

What draws all of these findings regarding challenge and crisis together is the potential implication that they have for the training of respeakers. While few have yet been integrated into training programmes, recommendations for this to be done are made by each of the authors.

3.3 Questions of quality and quantity

The *Communications Act* of 2003 required Ofcom to publish a code stipulating the obligations of UK television channels with respect to the provision of access services on television so that the “understanding and enjoyment” of television services for all viewers could be promoted (Ofcom, 2021: 2). Published in 2004, this code initially specified the minimum percentage of programmes that each channel with a significant viewing audience should be subtitled by the end of a ten-year period. For the majority, this was 80%, for ITV Channel 3 and Channel 4 it was 90% and for the BBC, 100% (p.5). In order to meet these higher percentages, the subtitled of live programmes was required and attention was placed on expanding this provision. Once these targets were met, for example, with the BBC achieving their quota in 2008 on its main channels (BBC, no date), followed by Channel 4 in 2012 (Channel 4, 2013: 2), attention turned once again to quality. Certainly, in the UK, despite the additional complexities in the production of live subtitles which relate to how fast they are displayed, the latency with which they appear on screen, possible technical issues (including frozen, lost or badly-positioned subtitles) and recognition errors, especially when the limits of speech recognition are pushed due to the fast-paced speech encountered on television, the expectations regarding the quality of live subtitles are high. Viewers expect full access to the

programme being broadcast so the quality of the subtitles, both in terms of grammatical accuracy and coverage of content, is under close scrutiny (for example, Ofcom, 2013). This shift in attention was reflected in a two-year study, led by Ofcom, into the quality of live subtitles on television (Ofcom, no date). Despite concerns raised by some viewers over the accuracy of the subtitles seen (Ofcom, 2013: 6) and these production challenges, the study revealed that, in addition to ensuring that subtitling quotas set by Ofcom were met, the quality seen in respoken subtitles has continued to improve (Romero-Fresco, 2016)³⁷. Further, one of its outcomes was an increased understanding among broadcasters of the demands of live subtitling, which led to improved pathways for collaboration and communication. In the case of the news, in particular, this facilitated the sharing of scripts and running orders with respeakers for them to use in their preparation and consequently led to the presence of more semi-live subtitles in that genre.

Respeaking is also a growing field within audiovisual translation and media accessibility internationally. It is used in many different countries (Romero-Fresco, 2018) and current pan-European research projects into the application of respeaking in interlingual settings³⁸ and into the professionalisation of the industry³⁹ demonstrate its growing and future potential.

In many countries, there is already a tradition of using respeaking to provide access in live settings outside of the television sector. For example, in Flanders, both intralingual and interlingual respeaking are being introduced at conferences; in Poland, respeaking has been used at meetings of the Polish parliament (Szczygielska and Dutka, 2017). In Germany, Austria and Switzerland, speech recognition software is used in combination with a specialised set of shortcuts that are programmed on regular keyboards. Referred to as speech-to-text interpreting (STTI), this service provides access in educational,

³⁷ Further discussion of how quality is defined in the Ofcom study, and in this thesis, follows in Chapter Six.

³⁸ For example, the ILSA project – see footnote 26 for more information.

³⁹ For example, the Live Text Access (LTA) project. This is an Erasmus+ project, 2018-1-DE01-KA203-00, and more information can be found at ltaproject.eu.

political, social and medical settings (Eichmeyer, 2017). In the USA, voice writers, the term used for respeakers, have joined stenographers, who traditionally provided Communication Access Realtime Translation (CART) or Live Event Captioning at conferences, events, classes and discussions (CCAC, 2016a, 2016b). Yet, in the UK, the uptake of respeaking in the live event sector has been slow. Ai-Media have explored its use in the educational setting (Nesta, 2014) and it is used in some university lectures (Ai-Media, 2021), and it is used occasionally at Stagertext events, when a captioner already has this skill (section 5.2.2 below), but it had got no further than this, until the pandemic began. Now, some live events are made accessible through respeaking (section 9.5.1. below). Despite these developments, respeaking remains a method of subtitle production that is little understood and frequently criticised. The increasing presence of respoken subtitles on public screens in airports, pubs and waiting rooms means that most people, and not only those who choose to view them on individual screens, have been exposed to them. Nevertheless, few understand how they are actually produced and the general perception is that these subtitles have been typed, and typed badly at that. Whilst respeaking is undoubtedly a profession, the lack of professional status for respeakers (Romero-Fresco *et al.*, 2019), together with these public perceptions, mean that attempts at introducing respeaking outside television have had limited success.

Nevertheless, there is a need for increased sensorial access outside of broadcast environments so that everyone can benefit from it. Currently, as explained in Chapter One, excellent sensorial access does exist at live events, however, without the regulation of an equivalent body to Ofcom, the quantity of the access provided in the cultural sector does not compare to that seen on television. The *Equality Act 2010* states that “reasonable adjustments” to services are expected. Whilst I understand that these adjustments may relate to physical features, auxiliary aids and/or the provision of information (Section 20), the details of exactly what these entail remain vague. The *European Accessibility Act 2015* and directive on the accessibility requirements for products and services which followed in 2019, *Directive (EU) 2019/882* (EUR-LEX,

2019), specify a range of devices and technologies which are to be made accessible, and include many which might be used in the context of live events, for example, when booking tickets or accessing content on mobile devices. However, as with the *Equality Act*, precise requirements for the cultural sector are not set.

Given that we know respeaking has successfully been used in live settings internationally, the purpose of this research is to explore how the number of events that are accessible in the UK can be increased and how respeaking can play a role, by complementing the access provision which is already in place. This might involve introducing access for d/Deaf, deafened and hard of hearing people into venues where this service is not yet offered, or by offering respeaking as an alternative means of access provision in venues that already cater for this audience. In the case of the latter, it must be stressed that the aim is not to replace current access provision, but instead to extend it.

3.4 Defining live events in the context of this study

Within Event Studies, planned events, the events we are concerned with here, are defined as:

(being) created to achieve specific outcomes, including those related to the economy, culture, society and environment. Event planning involves the design and implementation of themes, settings, consumables, services and programmes that suggest, facilitate or constrain experiences for participants, guests, spectators and other stakeholders. Every event experience is personal and unique, arising from the interactions of setting, program and people.

(Getz, 2007: 21)

They therefore incorporate a range of settings and activities. In this study, the focus will be on cultural events, though it is hoped that the findings will be more widely applicable. At a time when the arts are too often forgotten or overlooked (for example, Hewison, 2020; Weale, 2021) an added benefit of this research is to be able to highlight the interest, enjoyment and fulfilment that such access might bring (Crossick and Kaszynska, 2016).

In the context of this study, 'live events' has a more specific meaning. 'Live' is used to refer to an event happening in real-time, where the audience attends in person, and which is not watched in its entirety through a screen, although parts of it (the subtitles and certain visual elements such as PowerPoint slides or video clips) will necessarily be displayed on (at least) one screen⁴⁰. Live-subtitled events on television therefore fall outside the scope of this study; if, however, additional respeaking were to be provided for the audience present at the recording, then that would fall within the definition of a live event used here.

It is also important to note that the use of respeaking is only being suggested at unscripted or partially scripted events. Where a full script exists, such as at the theatre or opera, preparing captions or surtitles in advance and cueing them out live is the better option for access. However, where the event is for the most part unscripted, for example Q&As and discussions, or, perhaps, semi-scripted, where detailed notes exist but the exact words spoken may vary, for example at presentations, talks and tours, the suggestion is that respeaking could be used⁴¹.

Early on in the project, the possibility of attending live events remotely was raised, whereby an audience member might attend in real-time, but watch the entire proceedings via a screen. Even then, it was clear that such a service would be in demand, and the pandemic made it a reality for many of us. Since the scope of a live event that is attended in person was already a broad one, with many variables that a respeaker must be prepared for to explore, I decided to keep non-remote events as the primary focus of this project, so that I could consider them with sufficient rigour. Nevertheless, from my own experience of respeaking through platforms such as Zoom, I can say that many of the

⁴⁰ At events where audience members attend in person, there is also the possibility that additional audience members attend remotely, for example, through conference call software. Given the broad scope of live events defined above, the primary focus of this project has remained on non-remote events.

⁴¹ In this research, partially scripted events, like unscripted events, were respoken without semi-live cueing, despite the fact that the Text on Top software had this capability and the respeakers practiced this function during training. This is, however, a feature that could be used at future events.

findings from this study do apply when the audience attend remotely, and this is likely to be true for hybrid scenarios as well⁴².

Live, unscripted and partially-scripted events could therefore include museum tours, Q&A panels after cinema screenings and theatre shows, conferences, lectures and many more besides. A broad selection are tested within this research project.

3.5 Who are subtitles for?

3.5.1 The traditional audience

According to the first account of MA that we saw in section 2.1.6.2 above, it is d/Deaf, deafened and hard of hearing viewers who are the traditional or target audience for subtitles, be they live or respoken. While access to television content through BSL interpretation is also provided across the main television channels, the quotas set for this are far lower, so Deaf BSL users are likely to use both subtitles and signed content to watch television.

According to statistics compiled by the Royal National Institute for Deaf People (RNID)⁴³, it is estimated that in the UK in 2020, there were 12 million adults, the equivalent of one in five of the UK population, with hearing loss greater than 25 dBHL. Considered as mild, this number is likely to rise to around 14.2 million by 2035. Out of those 12 million, 1.2 million were estimated to have hearing loss greater than 65 dBHL considered moderate to severe (RNID, no date)⁴⁴. Approximately 151,000 people use British Sign Language, of whom 87,000 are Deaf (*ibid.*)⁴⁵.

As I intimated in Chapter Two, behind the term ‘d/Deaf, deafened and hard of hearing’ (DH), lies a very diverse group of people. Each person’s experience of hearing loss

⁴² When providing access in scenarios where some or all of the presenters or audience are attending remotely, the key considerations will relate to how the live subtitle feed is incorporated into the conference-call software, and in the physical event if applicable, and what options exist for the audience to position and view it. Different procedures may be needed to manage turn-taking both between respeakers, but also when there are audience questions.

⁴³ The RNID (rnid.org.uk) was previously known as AOHL.

⁴⁴ These figures are the current estimate displayed on the *Facts and Figures* page of the RNID website and come from RNID prevalence estimates which use Office for National Statistics population data (2018).

⁴⁵ Cited on the same *Facts and Figures* page, the source of this figure is the British Deaf Association (BDA, no date), who state it is based on extrapolated data from the Scottish Census in 2011.

differs, be it through the onset time and degree, the communication style a person uses, how easily they are able to cope in different situations, and according to the effects of more traditional factors of diversity including native language, gender, sexual orientation, age, disability, ethnicity and race. During any discussion of how to meet sensory needs through subtitles, it is vital that this diversity in lived experience is remembered and that no assumption is made that people who share one of the labels that follows have either identical hearing or would require the same access to meet their sensorial needs.

From a medical point of view, hearing loss occurs when the sounds around us aren't communicated properly to the brain (Specsavers, 2021). **Sensorineural** hearing loss results from "damage to the cells that pick up sound or to the nerves that carry information from the ear to the brain" (*ibid.*). This might include damage in the inner ear, auditory nerve or to the brain itself. In contrast, **conductive** hearing loss is caused by blockages in the ear canal or middle ear; here, the level of sound that can pass through to the inner ear is reduced (*ibid.*). A person may also experience **mixed hearing** loss, if both occur at the same time.

Many people experience a combination of hearing loss and other sensorial conditions. Since the vestibular system which is responsible for our balance is also located in the inner ear, nausea, dizziness, noises in the ear and balance problems may accompany hearing loss, and **deafblindness** refers to people with dual sensory loss (Sense, 2021). Some people may experience sensitivity towards noise, **hyperacusis** (AOHL, 2018), and a large number of people experience **tinnitus**, which is the perception of a sound, such as a ringing, buzzing, hissing, or whistling, which has no external source (British Tinnitus Association, no date).

Hearing loss may occur at any age, though it is often a feature of ageing. One important distinction that is often made is whether the onset was pre-, peri- or post-lingual as this may affect how speech develops (Kiversal, 2019). It may be experienced in one or both ears and could result from many different causes including ear wax, an infection, a burst ear drum, various diseases, loud noises or foreign bodies or growths in the ear (NHS, 2017).

Hearing loss is generally classified by degree, and four different levels can be identified (AOHL, 2015: 23). If a person has **mild hearing loss**, it is likely that they will be able to distinguish some sounds in speech, but there may be a general difficulty in following what is said, either in conversation or on television. **Moderate** hearing loss means fewer than half the necessary sounds in an utterance are accessed and other noises, like doorbells or the telephone, are not heard. If hearing loss is **severe** or **profound**, spoken words appear as only soft or loud mumbles, and following television programmes, even with the volume turned up, is difficult or impossible (Bricker, 2015; Romero-Fresco, 2019: 102). At the point of severe hearing loss, people will usually need to lipread or use sign language, even if they have hearing aids; with profound deafness, hearing aids may no longer be helpful, though cochlear implants may be of benefit (AOHL, 2015: 23).

The most common way of talking about people with hearing loss uses terms that refer to the range of sounds that can/cannot be heard and figure 3.3 illustrates the decibel range where some of these sounds sit.

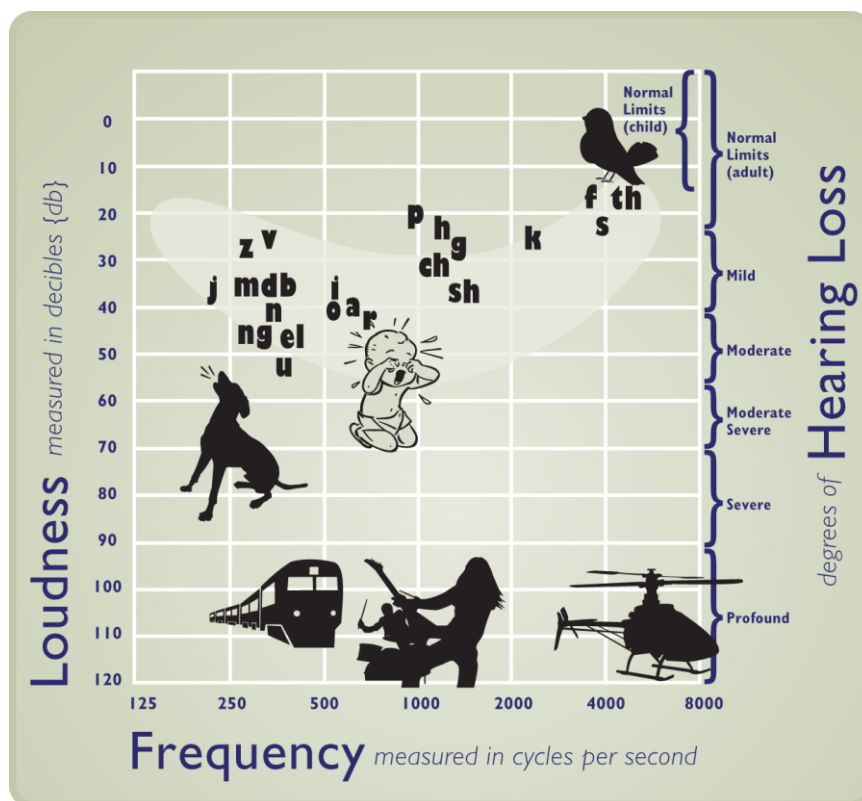


Fig. 3.3: Familiar sounds audiogram (KDH Research & Communication, 2012)

The term '**deaf**' is often used in a generic way to refer to "the full, broad range of people who have some level of deafness" (Stagetext, 2017b: 4), but it also refers more specifically to severe or profound hearing loss (SignHealth, no date), where people are unable to hear anything below 70dB. A further distinction within this term is often indicated by its (non) capitalisation: written as 'deaf', the term usually refers to people whose native language is English, whereas '**Deaf**' refers to people whose preferred language is British Sign Language (BSL) and who are often prelingually deaf (SignHealth, no date). Used in this way, the capitalisation emphasises "the strong cultural identity among the deaf community, often referred to as 'capital D' Deaf" (Stagetext, 2017b: 4).

This d/Deaf distinction is one that has been used widely in AVT and MA, and which I adopt in this thesis. I am, however, also aware that it risks being an "often unhelpful and unintentionally oversimplifying distinction" (Kusters *et al.*, 2017 in Lewis, 2017); some academics, such as Kusters *et al.*, consequently suggest that the term 'deaf' should be used throughout. Whilst I agree with the point they make and the inherent risk of simplification behind any of the labels surrounding hearing loss, in the case of this thesis, I have decided to retain the d/Deaf distinction. Rather than simplification, I see potentially very different experiences among deaf and Deaf audience members in their use and expectations for all forms of subtitling on television and at live events and would like to explore this more fully; the fact that BSL provision also exists in both settings, albeit to different degrees, makes it all the more important to explore. Where possible, in the chapters that follow, I try to explore the experiences of individuals and the groups to which they belong as I examine any patterns in the opinions that are expressed.

'**Deafened**' refers to people who "were born hearing and became severely or profoundly deaf after learning to speak" (Stagetext, 2017b: 4). These three terms, deaf, Deaf and deafened, are all concerned with the most profound and severe levels of hearing loss, although as already explained, the circumstances surrounding the onset of the hearing loss and the communication styles individuals use differ.

In contrast, the term '**hard of hearing**' is used to describe people who don't hear well either because they were born without hearing or because they, perhaps gradually, lost some or all of their hearing later in life (*ibid.*). The term may refer to people with mild hearing loss, from 25dB, up to severe (WHO, 2021) or profound levels (IFHOH and Blau, 2021: 4).

3.5.2 The wider audience

Whilst SDH subtitle provision on television may have been intended for d/Deaf and hard of hearing audience members, a far broader audience in fact makes use of it. According to figures published by Stagertext (2017c), out of those people who use subtitles on television, 80% use them for reasons other than hearing loss and it is now widely recognised that adding captions to online content increases the views it receives. There are many different situations where people without hearing loss may choose to use subtitles, and the number of screens in public places where subtitled content is displayed is increasing⁴⁶.

By adopting a wider audience approach of the universalist account of MA, outlined in Chapter Two, I am also able to address this trend in subtitling use and practice.

As an example of a wider audience group, I decided to explore how non-native English speakers (NNE) living in the UK might make use of live subtitles. Whilst potentially as diverse as DH audience members, I felt that this was another group who might benefit from using subtitles and thought it was a good opportunity to see how their needs might overlap.

For many years, the idea of using television to support learning was treated with suspicion, as it was considered a "shallow and easy medium" (Vanderplank, 2016: 34) more appropriate for entertainment and information than language learning. This opinion held, despite many early studies which demonstrated the learning potential of captioned video (Price, 1983; Vanderplank, 1988; Montero Perez *et al.*, 2013). Over time, this view

⁴⁶ The fact that subtitles may be used by people with the television sound turned down is noted in Ofcom's Code on Television Access Services (Ofcom, 2015: Annex 4), where the non-homogeneity of users of access services is also commented on.

has evolved and much research has now been carried out into how same-language subtitles, or captions as they are called in this context, can benefit NNE, or L2 users⁴⁷. The findings of Montero Perez *et al.* (2013: 733), who used a meta-analysis to review primary research of the previous 30 years, support the claim that “captioning helps learners to improve comprehension and fosters vocabulary learning”. In particular, these authors believed that the presence of captioning allowed speech to be monitored more consciously, as proposed by Vanderplank (1988: 277), and offered a solution to typical problems that L2 learners might encounter by providing them with a real-time script which reduced the effort needed for decoding (Montero Perez *et al.*, 2013: 731). This allowed L2 learners to focus in on individual items of vocabulary, thereby supporting the acquisition of new terms.

In the current study, I am less concerned with formal language learning, but rather interested in how people attending might make use of the intralingual subtitles that are available to them in less formal settings. This shift in setting may have a number of implications on how well NNE participants engage with the captions.

Whilst the audiovisual material used in lessons may be carefully tailored to the listening comprehension level of a student, this may not be the case for content found on television and at live events. Some have argued that “exposure to authentic audiovisual material which has not been carefully enough selected or made accessible to non-native viewers can be a very inefficient pedagogical approach” (Danan, 2004: 71), while others offer (anecdotal) evidence that authentic material is beneficial (Vanderplank, 2016: 249-250). For this reason, the degree to which NNE audience members find the audiovisual material at the events complex or challenging will be something to monitor during this study.

On the other hand, in the EURECAP Study, Vanderplank (2016: 221) found that it may be in informal settings where the prime value of captioned viewing lies, since it is here

⁴⁷ Interlingual subtitles can also be used to support L2 users, though the benefits they bring differ from those that intralingual captions offer (Danan, 2004; Mitterer and McQueen, 2009).

that learners have choice and control over what they view. His cognitive-affective model of language learning (p.240-241) suggests that when they have an interest and purpose in viewing particular content, learner-viewers will attend to the captioned input “systematically and reflectively”, recalling prior knowledge and synthesising it in new ways; the language each person attends to will relate to their individual purpose in watching. A similar sense of agency and choice may be seen at live events, as audience members select and decide which events to attend.

To the best of my knowledge, this marks the first study into how L2/NNE perceive respoken subtitles and would wish them to be created, and whilst this is only a very small contribution towards this large body of work, it may mark a step in a new direction within it⁴⁸.

The findings from Vanderplank (2016) also suggest an immediate similarity between DH audience members and NNE in the way they may approach a live event, and indeed use the subtitles provided. Where DH audience members may find that subtitles help them fill in gaps in the sounds they hear to make what is said more comprehensible (*Stagetext*, 2011), audience members without English as a native language, may find gaps in comprehension being supplied by the subtitles. For both, effort is required.

In even broader terms, live subtitles may bring benefit to anyone in the audience. Anyone, hearing or not, native speaker or otherwise, may encounter new ideas and new terminology at a live event and may benefit from having this reinforced in script. Where respoken subtitles began as providing sensorial access, and then came to be considered as linguistic access, from the perspective of NNE, it may evolve once again into an access modality, where, in line with the Universalist account, the access it provides is non-specific. Considering respoken subtitles in this way allows it to be framed from the both/and perspective noted in Chapter Two. What will need to be explored in this study is the

⁴⁸ Some exploration of user preferences over live subtitling techniques has been conducted by Eichmeyer-Hell (2021), though it does not appear that the audience were non-native speakers.

impact this has on how content is captured and on how this content is received by the audience, given the scrolling form of the subtitles, which appear with a slight delay.

3.6 Subtitle reception: Factors affecting the speed, content and presentation of subtitles

3.6.1 The question of speed

By its very nature, subtitling is a form of translation which is constrained (Bogucki, 2004: 72). Both the visual context and factors related to time and space impose constraints on what content the subtitler is able to include. In most forms of subtitling, the aim of the subtitler is to achieve an optimum balance of speed, timing and editing, whereby the written word is as closely in sync with the spoken word as possible (timing), whilst capturing the original content of what was said as closely as possible (editing) and presenting it in a way that the viewer can comfortably read and follow (speed)⁴⁹. However, determining what this comfortable speed is presents a number of challenges: speech and reading rates differ from each other and between people and different expectations exist around whether and how subtitles might be edited. This is particularly true in the case of intralingual subtitles, where any editions may be more apparent to the audience.

Since the real-time nature of live subtitling means there is an inherent delay between a word being uttered and its written form appearing on screen (latency), a further time-related challenge is introduced. The viewer is prevented from processing the written subtitle at the same time as the spoken word and original time-matched visual image; the respeaker must attempt to minimise this delay as far as possible. This delay is common to all forms of live translation and interpreting but becomes particularly

⁴⁹ Certain types of subtitles are created with different intentions in mind. For example, the EASIT project (EASIT, 2019: 29-35) examined easy-to-use subtitles, where the visual load of the subtitle in general would be reduced (p.32). As well as using short words, an active tense and simpler syntax, this might involve providing additional content, such as glossaries, and offering options to control the legibility of the subtitles (p.30). Ai-Media have developed Simple Text for students on the autism spectrum. Here, when subtitling lessons and lectures, the respeaker removes figurative language and metaphor from their subtitles and focuses on presenting key concepts in a (visually) clear way. This allows the students to have a single point of focus and leads to “less distraction, reduced anxiety and improved attention” (Ai-Media, no date b).

noticeable when captured in intralingual subtitles. Already a vulnerable form of translation (Díaz Cintas and Remael, 2007: 55) since the combined presence of the visual target text alongside the spoken source means that editions may be obvious to the audience, leading the way to potential comparison and critique, the visibility of this latency increases the vulnerability of respoken subtitles. This lack of synchronicity is frequently commented on in discussions about the quality of access respoken can provide (Ofcom, 2013, 2014a, 2014b, 2015a, 2015b; Romero-Fresco, 2015: 141-142 and Chapter Five of the current study).

For this reason, speed has always occupied a “privileged position” among the commonly debated topics in subtitling literature (Romero-Fresco, 2015: 335). As well as being a technical concern, it raises cognitive, economic, political and ideological questions about the nature of subtitling (*ibid.*). It is one of the few constants in discussions of respoken across country and context (Romero-Fresco and Eugeni, 2020: 286) and as the broadcaster regulator Ofcom suggests, it is “arguably the key underlying issue behind nearly every important issue” in SDH (Ofcom, 2005: 11). It is a topic that came up frequently within the discussion with key parties in this research (see Chapter Five).

3.6.2 Three measures

In the UK, speed is measured in words per minute (wpm)⁵⁰. In traditional subtitles, two rates must be considered, whilst in live subtitling an additional third rate is also of consequence.

⁵⁰ Subtitle speed can be measured in words per minute (wpm) or characters per second (cps). Studies on English-to-English SDH have traditionally used wpm (Szarkowska & Gerber-Morón, 2018: 2) and I will follow this tradition to enable comparisons to be made. The following conversion table from Díaz Cintas & Remael (2020: 112) provides a useful guide to how wpm and cps compare when blank spaces are included within the calculation:

12cps	13cps	14cps	15cps	16cps	17cps	18cps	19cps	20cps
150wpm	160wpm	170wpm	180wpm	190wpm	200wpm	215wpm	225wpm	240wpm

and when they are not:

12cps	13cps	14cps	15cps	16cps	17cps	18cps	19cps	20cps
130wpm	140wpm	150wpm	160wpm	170wpm	180wpm	190wpm	200wpm	215wpm

It must be noted that the way average subtitling speeds are calculated vary (Fresno and Sepielak, 2020) and different equivalences will be seen within the literature.

The first is the speech rate of the original source text, the original speech rate (OSR) (Romero-Fresco, 2010: 177). It is this rate that sets the pace for the subtitles that follow. Much variation in this rate is seen across programme genres (table 3.1). On the BBC, sports programmes were found to have an average OSR of 160wpm (ranging from 124-182wpm); the news had an average of 180wpm (ranging from 161-198wpm) and some interviews and weather reports reached even higher speeds, averaging at 230wpm (ranging from 211-245wpm) (*ibid.*).

Table 3.1: Speech rates on BBC programmes, based on Romero-Fresco, 2010: 177

BBC	Original speech rate (OSR)	
	Average	Range
Sports programmes	160wpm	124-182wpm
News	180wpm	161-198wpm
Interviews; weather reports	230wpm	211-245wpm

The second rate is the respeaker’s speech rate (RSR). Studies by Eugeni (2009) and Romero-Fresco (2009) reveal that this rate depends largely on the OSR, ranging between 106-190wpm, yet always falling below the OSR by up to 20wpm at OSRs of up to 180wpm, and by up to 40wpm at OSRs above this. Whilst the respeakers in fact uttered the same number of words as the original speaker, the wpm fell once spoken punctuation and macros had been converted to text. As Romero-Fresco (2010: 178) notes, to match the OSR and produce verbatim subtitles, “respeakers would have to speak faster than their original speakers...which is probably against the grain in this shadowing-like type of translation”.

The final rate is the speed of the subtitles themselves, calculated as the rate they are presented to viewers and referred to by various terms including *subtitling speed*, *reading speed*, *subtitle presentation rate* and *subtitle display rate* (Szarkowska, 2018: 7). Traditionally, a *six-second rule* has been applied to the speed of subtitles on television

(Díaz Cintas and Remael, 2007: 23), based on the time it takes the average viewer “to read and assimilate the information contained in two lines of a subtitle, where each line consists of some 35 or 37 characters”. Displaying such a subtitle for less than six seconds means too little time is allowed for it to be processed; displaying it for longer than six seconds risks the viewer beginning to re-read it (Szarkowska, 2018: 7). This six-second rule translates into approximately 140-150wpm or 12cps (Díaz Cintas and Remael, 2020: 109)⁵¹ and allows the amount of text to include in subtitles of shorter lengths to be calculated from it⁵². As Díaz Cintas and Remael (2020) explain, “these rather slow subtitle rates are the approximate values traditionally applied in the subtitling of audiovisual productions that are aimed at viewers who may find it challenging to read text on screen, like young children”.

3.6.3 Validity of the six-second rule

Although this six-second rule is the traditional way that subtitle speeds have been calculated, many have questioned its validity and continued relevance, especially as viewers become more exposed to reading text on screen and the actual speed used in different settings vary, with Díaz Cintas and Remael (2020: 110) noting higher speeds of 180wpm in DVDs, and more recently of 200wpm on Netflix programming for adults. Stagertext’s policy is to provide full and equal access and in the case of their digital work, the priority is to subtitle the content in full rather than to edit to a particular speed (Stagertext, 2017a), since the audience has the option of replaying content when needed as the viewing is on-demand.

Closely linked to the question of speed is the question of content. Slower subtitles involve editing, which is a costlier process for broadcasters (Ofcom, 2005: 6), whereas verbatim subtitles, with increased text, have higher subtitle speeds but can be produced more

⁵¹ As may be expected given the variation in how subtitling speeds are calculated, the exact wpm associated with the six-second rule also varies. Some sources calculate it as 144wpm (Díaz Cintas, 2003, cited in Romero-Fresco, 2009: 114; Martí Ferriol, 2013: 208); Díaz Cintas and Remael listed it as 140-150wpm (2007: 97) and later as 150wpm (2020: 109).

⁵² When testing the suitability of this rule through eye tracking, d’Ydewalle *et al.* (1987: 321) noted that one-line subtitles should be kept on screen for longer than the time derived from it. They referred to the two lines as being 32 characters each plus spaces (p.313); Díaz Cintas and Remael (2007: 96) specify a maximum of 37 spaces per line, giving a total of 74 characters.

quickly, and therefore more cheaply. The opinion of the audience is harder to determine; while many react “with considerable sensitivity and antagonism towards the idea of editing” (Ofcom, 2005: 3), opinions about which are easier to read are divided, which is to be expected in such a diverse audience group (see Chapter Five below).

On television, official recommendations match those of the six-second rule, approximately 140-150wpm, but actual display rates are far higher. In 1999, the ITC Guidance on Standards for Subtitling (ITC, 1999: 11) recommended that “the subtitle presentation rate for pre-recorded programmes should not normally exceed 140wpm”, but acknowledged that “in exceptional circumstances, for example in the case of add-ons⁵³, the higher rate of 180wpm is permitted”. Ofcom (2005: 6) noted that average speeds of pre-recorded programmes were 160wpm for pre-recorded programmes, and suggested this guidance was in need of revision. In 2015, a higher reading speed was proposed, where 160-180wpm was suggested for pre-recorded programmes (Ofcom, 2015c: A4.19, p.19). In the case of live subtitles, it was acknowledged that setting such a limit was more complicated: “Although it may not be practicable to restrict the speed of subtitles for all live programmes, commissioning editors and producers should be aware that dialogue which would require subtitles faster than 200wpm would be difficult for many viewers to follow” (*ibid.*).

As a solution to this, the use of three-line subtitles was suggested, since they allowed longer for the subtitles to be read (*ibid.*). Despite this updated guidance, the maximum speed that is actually used on television is far higher than this, with a limit of 250wpm for pre-recorded programmes and of 320wpm for live programmes, on at least some channels (Ofcom, 2015a: 25). Audience opinions gained in focus groups for the DVT4ALL study conducted in 2010-2013, while these speeds were increasing, suggest that pre-recorded television speeds in general were “about right” (Romero-Fresco, 2015:

⁵³ Add-ons, or cumulative subtitles, involve the use of dynamic text (ITC, 1999: 14-15). In this technique, additional words are added to a subtitle which is already present – and remains – on screen. This is a technique that is used in certain circumstances in pre-recorded subtitling, for example to provide the punchline of a joke (*ibid.*). However, the add-on nature of this technique may be likened to the way that new text appears in live subtitles, although when the scrolling format is used, the original text moves along as the new words appear.

156). While participants noted that they sometimes struggled with the scrolling mode of live subtitles (p.160; and an effect seen in Sancho-Aldridge, 1996), it was more the accuracy and latency of the subtitles that caused them to complain. However, it is possible that exposure to this scrolling mode has acclimatised viewers to faster speeds in pre-recorded subtitles.

Szarkowska and Gerber-Morón (2018) highlight the mixed findings on the effect of subtitle speed in academic research, while Romero-Fresco (2015) moves away from the notion of subtitle speed towards that of viewing speed “to account for the audiovisual nature of the programme” (p.337-338). He defines viewing speed as “the speed at which a given viewer watches a piece of audiovisual material, which in the case of subtitling includes accessing the subtitle, the accompanying images and the sound, if available.” Using eye-tracking technology, he illustrates that as viewing speed increases, the time spent on the subtitles increases at the expense of time spent on the images, as illustrated in table 3.2⁵⁴:

Table 3.2: Viewing speed and distribution of gaze between subtitles and images (Romero-Fresco, 2015: 338)

Viewing speed	Time on subtitles	Time on images
120wpm	±40%	±60%
150wpm	±50%	±50%
180wpm	±60%-70%	±40%-30%
200wpm	±80%	±20%

Whilst viewing speed and subtitling speed share the same value, viewing speed encourages a greater focus to be placed on the viewer’s experience when processing the subtitles, and increased consideration of the image, and its duration, when determining what speed the subtitle should be (p.339). In contrast, Szarkowska and

⁵⁴ The ± indicates that the figures included in the table are average values for the distribution of gaze, calculated across all the participants on the basis of their viewing speeds (Romero-Fresco, 2015: 338).

Gerber-Morón (2018), through a mixture of eye-tracking, questionnaires and semi-structured interviews (p.5), found that “fast subtitle speeds do not necessarily hold viewers back from watching the filmic image” (p.25).

Ultimately, it seems that how viewers respond to the speed of subtitles will depend on a range of factors, but what is important is that the subtitle content and speed should reflect the programme and genre in question. Where the subtitle content differs noticeably from what might be expected from the visual image, for example when a person on screen appears to be speaking at length, but only a short subtitle appears, the viewer is more likely to sense the dissonance between them. In cases where a subtitle is displayed at a particularly slow speed, re-reading may also be triggered (*ibid*; p.1). As Sanchez (2014) writes:

Sometimes [determining reading speed] may mean choosing to have subtitles with a slightly higher reading speed rather than lose valuable information. At other times, it may mean opting to edit down because the rhythm of the last few minutes of subtitled dialogue has been frenetic, or to sacrifice audio synchronicity for the sake of maintaining word order and verbatim text.

3.6.4 The question of live subtitles

In the case of live subtitles, the discussion on speed and content is a little different. The subtitling software will determine how fast subtitles can be respoken or cued out in semi-live mode in programmes such as the news where partial scripts are provided (Romero-Fresco, 2011: 12). Much will also depend on what speech rates/respoken rates individual respeakers can achieve and on the degree and type of editing that a particular genre of programme will demand, as a result of their typology or the content and features within them (Eugeni, 2008; Sandrelli, 2013, 2019, 2020; Szarkowska *et al.*, 2017). At live events, no such limits will be set by the software. Rather, there will potentially be a freedom for the presenters to respond to the audience and the respoken subtitles and to adjust the pace of their delivery accordingly. This is something I explore further in Chapters Seven, Eight and Nine.

In an early analysis of respoken subtitles from sports, news and interviews/weather reports on the BBC, Romero-Fresco (2009: 109) suggested that the editing carried out

by respeakers “led to a minimal loss of information, especially as compared to the potential loss of information for viewers reading respoken subtitles at the current speeds”. The more recent Ofcom (2015a: 27-28) study on the quality of live subtitles, the largest study on a corpus of respoken subtitles to date, showed similarly low editing rates on content such as the news, where the degree ranged from 12-17% across the four rounds of the study. In other genres, however, the degree was far higher, with a median edition rate of 27.5% seen in chat shows and 32.1% in entertainment programmes in the final round (Ofcom, 2015a: 27-28).

Eye-tracking has also been used to explore the reception of live subtitles. Romero-Fresco (2010: 175) found that the word-for-word scrolling display mode resulted in viewers spending 90% of their time looking at the subtitles and only 10% on the images, whilst Martínez *et al.* (2010), cited in Pérez (2012), reported on the effect of punctuation-based text segmentation. They displayed subtitles in three modes – add-on phrases and add-on sentences, which in both cases showed up one-by-one, and as blocks with no segmentation, where the subtitle area was “filled with the maximum amount of text regardless of any semantic or grammatical segmentation” (p.139). The best results in terms of both reception and comprehension were displayed in the add-on sentence mode. In this mode, the number of fixations in the subtitle area was at its lowest, allowing the maximum amount of time to be spent on the image (*ibid.*). According to Rajendran *et al.* (2013: 9), who referred to this study:

Viewers of scrolling subtitles thus spent a larger proportion of their time (88%) processing text rather than the visual scene, while viewers of blocked subtitles could devote a smaller proportion (67%) of their time to doing the same. Scanpath visualisations of eye movement patterns suggested that fast readers read ahead of the scrolling subtitles and cast their gaze (astray fixations) on gaps where they expected to find the next word while slow readers lagged behind and needed to re-read words (regressions).

Although based on blocked cinema subtitles, Perego *et al.*'s (2010: 243) contribution on the impact of “syntactically incoherent” segmentation, and whether this would have “a disruptive effect on information processing and recognition performance”, is also of relevance to how live subtitles are processed. Knowing that print reading, which is

sequential and holistic, occurs in chunks, they wanted to find out whether subtitle lines ending on unnatural breaks would make processing more difficult as a result of greater integrative effort being required, or whether no disruption would occur, as reading is an automated process (p.249). They found that the quality of segmentation in the subtitles did not have a significant impact on processing (p.263). In addition, they noted that “participants seem to deploy efficient and selective attention allocation strategies which enable them to reach a good understanding and a good recognition of presented information” (p.264).

Rajendran *et al.* (2013:11) evaluated four subtitling styles in live, respoken subtitles. The first was blocked, equivalent to Perego *et al.* (2010), then word for word scrolling, chunked (scrolled) by phrase and chunked (scrolled) by sentence among hearing participants, recognizing that their reading patterns may differ from those of DH audience members. They found that different styles of segmentation did elicit different viewing behaviours, and that text chunking either by phrase or sentence reduced the amount of time spent on the subtitles, as the text was presented in a way that was more easily processed (*ibid.*).

All of these findings highlight the challenges involved in the creation of real-time subtitles; not only must the respeaker maintain a good pace, and respeak accurately, they must also be aware – as far as possible – of how the way in which what they say appears on screen may impact on audience comprehension. This will be considered further in the 3.7 below.

3.6.4.1 Display

A few studies have been conducted on the effect of subtitles being displayed on different screens, which is certainly of relevance in the move to the live event setting. In the UK, theatre captions and opera surtitles have long been displayed in different positions around the stage and more recently new devices have been adopted, such as screens built into the seat in front of you, as on an aeroplane, and caption glasses, where the audience see the captions displayed on the lenses of the glasses (National Theatre,

2018). At live events, tablets may be provided for subtitle streaming, and on occasion, personal devices can be used as well⁵⁵. Miquel-Iriarte *et al.* (2012: 262) considered the use of smartphones for viewing semi-recorded subtitles, which were “prepared in advance but launched live”. They found that participants’ visual attention flowed regularly between the presentation image and smartphone device (p.273). In addition, they suggest that the inclusion of subtitles in university lectures, where this was tested, might support a more efficient use of the time spent on the image, since the act of reading the subtitle content directed participants’ gazes to the relevant part of the image on screen. Whilst this may not be applicable to all live events, it is certainly of interest in the current study. Romero-Fresco and Fryer (2016) explored audience preferences for different display units for displaying theatre captions; they found that LED screens positioned at the side of the stage (side LEDs) were the preferred choice, followed by integrated LEDs which were positioned as part of the set or more centrally above it, and then tablets. The objective measures matched these findings and indicated that side LEDs allowed more time to be spent viewing the action. Tablets seemed to work well for certain content, and when speech rates were lower, but not for fast-paced dialogue.

3.7 Respeaker training

3.7.1 Skills for intralingual respeaking

Given the complexity of the respeaking process, much attention in academia has been paid to what training for respeakers should look like and which professionals might be suited to this profession. Although in-house training began in the early 2000s, with the BBC offering a live respeaking service in 2001 (Lambourne, 2006), the first academic proposals appeared in 2007-8.

Van der Veer (2007: 1) began by posing the question of whether a good interpreter would automatically be a good respeaker and proposed that respeaker training would be best included in an education programme for conference interpreters. This began the trend

⁵⁵ To date in the UK, aside from when caption glasses are used (Stagetext, 2021b), the captions and subtitles on these screens are displayed by a captioner; in the US, I Caption is a device in use at theatres which displays captions based on lighting cues (Romero-Fresco and Fryer, 2016).

for the development of training programmes for respeakers which drew on skills from both interpreting and subtitling (Arumí Ribas and Romero-Fresco, 2008; Romero-Fresco, 2011) and the consideration over which backgrounds new trainees might come from, for both its intra- and interlingual forms (Szarkowska *et al.*, 2018; Dawson, 2019; Davitti and Sandrelli, 2020 and the ILSA and SMART⁵⁶ projects)⁵⁷.

In 2008, Arumí Ribas and Romero-Fresco provided a comprehensive taxonomy of the skills involved in intralingual respeaking and a practical proposal for how this training could be taught in academic institutions, thereby expanding the range of training settings from in-house training alone.

The skills were divided into three key sections; those drawn from subtitling, those from interpreting, and those which were unique to respeaking. Further distinctions were made between the skills that needed to be activated before the process began and during the respeaking task, and, for the latter, whether they related to the source or target text, or a cross over between them (p.113).

Prior to the process, the range of skills included software-related skills, both drawn from SDH subtitling and specific to respeaking, and preparation and strategic skills, including teamwork, that most resembled those required from simultaneous interpreting (SI).

During the respeaking task, source-text-based skills were drawn entirely from SDH and SI and involved the ability to analyse, synthesise and reformulate the "communicative intention of the source message" (p.115). Listening comprehension skills were essential for this. The crossover tasks involved the synchronisation toolkit of subtitling, with the added ability, from interpreting, to multitask, receiving an incoming message and monitoring the outgoing message, whilst coping with the *décalage* or "distance between the original speaker and respeaker" (Romero-Fresco, 2011: 99). Live skills of managing stress, correcting errors and coping with frustration were now combined with respeaking-

⁵⁶ See footnote 33 for further information about the SMART project.

⁵⁷ In this account, I focus on the training of the skills involved in intralingual respeaking, in line with the use of intralingual respeaking in this study. Further information and resources on the training of interlingual respeakers can be found in Dawson, 2020 and on the ILSA project website <http://ka2-ilsa.webs.uvigo.es/course-and-training-materials/>.

specific skills of coping with "four types of simultaneous intersemiotic tasks: listening while speaking, writing and reading" and live skills such as changing the colour and position of subtitles, and pre-editing whilst respeaking. Target-text specific skills included production skills and awareness of the target audience from subtitling, the delivery skills of interpreting and respeaking-specific delivery skills, namely those outlined above, including accurate oral punctuation, maintaining a consistent delivery that is adapted to the SR software (Arumí Ribas and Romero-Fresco, 2008: 116). This remains a core taxonomy on respeaking courses.

Pražák *et al.* (2020) provide an outline of how these skills might be realised in a training programme, detailing four steps required in developing the required competencies. The first trains the respeaker's skill in listening and speaking simultaneously, the second focuses in on optimising delivery so that it is attuned to the SR software and the highest possible accuracy can be achieved. The third is focused on rephrasing and condensing; in order to do this well, the respeaker must also be able to monitor the output that appears on screen and perform corrections. At the same time the authors suggest respeakers refine their voice models and add and adjust vocabulary as needed. They envisage feedback from an instructor as being particularly important in this stage. Sandrelli (2019: 163) also highlights the importance of reviewing work and refers to the use of the NER model to support assessment in the courses she leads. In the final phase in the training programme, the respeaker works with the whole subtitling system but in an as-live form, so no subtitles are broadcast, the idea being that only once all four phases have been mastered, will the respeaker begin working live. Conducting training in this way means that respeakers can progress at their own pace, though an average estimated time for an intensive training plan like this is two to three months, which equates to a minimum of 100 hours training.

My own experience of being trained in-house, and the way I approach delivering respeaking training at the University of Roehampton, differs slightly from this. My preference is for there to be a focus on dictation in the initial weeks of the course, so that students have the opportunity to learn what the software can do, as they begin to

gradually work with it at increased speeds, and incorporate the split attention required to perform respeaking effectively. At the same time, voice care features prominently in this work to support respeaking for increased durations.

Ultimately, what is key is that any trainee has the time to develop the core skills whilst focusing on the accuracy of the output they are achieving. In that way, they can draw together the practical understanding required to work as respeakers.

In the training offered in this research, the training outlined above is considered as foundational. All those taking part in the training will have already mastered these skills and worked professionally for a number of years. The bespoke training offered will focus more on providing opportunities for these professionals to adapt these skills that they hold for work in a live event environment and to develop the wider skills that they will require to enable them to fulfil the role of a live event subtitler.

3.7.2 Skills for a live event subtitler

When respeakers move beyond the threshold of their in-house workplace, they are still likely to share assignments with a fellow respeaker, but the wider team that supported them in-house will no longer be present. This means that additional responsibilities will fall onto the respeakers themselves, ranging from increased technological know-how to wider discussions around access and organisation. Far less research has been conducted into what skills and competences are required in these scenarios.

In the LTA project⁵⁸, Oncins *et al.* (2019: 130) envisage the role of a subtitler at live events as one of a linguistic and cultural “mediator”, where the live subtitler becomes “a meaning-making agent”. In this context, creating meaning as the source language is transferred into the target text draws on far more than linguistic and technological skills; an understanding of accessibility and the ways it can be implemented (p.143) as well as entrepreneurship and service competence (p.145-6), which include management and interpersonal skills, stress management and business strategies, are also required.

⁵⁸ See footnote 39 for further information about the LTA project.

Within this skillset, they also reference Katan's (2009: 89-90) logical levels, where a mediator must determine what is going on within the context of culture and in the context of the moment at the level of environment, behaviour, strategies, values/beliefs, identity, role and mission in society (p.132). Oncins *et al.* note that this role, and the training required for it, are still being shaped (p.150).

The ACT Project (Remael *et al.*, 2019) presented an alternative role, that of an “access manager”. Those involved considered the training that would be required in a variety of live event scenarios, and the wider organisational and instructional co-ordination that access professionals would require. The focus was less on the cultural aspect involved in translation or mediation, and more on carrying out access provision. There was an implicit understanding that many benefiting from this training would be complementing knowledge and understanding that they already held, so a flexible approach would be needed, both in how the course was structured and in the way that the skills learned could be deployed (p.145).

In this thesis, I adopt the term ‘designated access co-ordinator’ and envisage it as a role that the respeaker may share with other members of the event team.

Greco (2019b: 41) has written of the importance of a critical learning pedagogy for accessibility studies, where an understanding of access is embedded into training so that any access provision develops alongside an understanding of the role that the access plays in social inclusion. In line with this, many recent courses on access service modalities, such as ILSA and EASIT, the Easy Access for Social Inclusion project⁵⁹, and also some university-based access modules, now begin with modules on what accessibility is before the technical skills of the modality, in this case subtitling and respeaking, are introduced.

⁵⁹ EASIT is a project funded by Erasmus+ Strategic Partnership programme, 2018-1- ES01-KA203-05275; information about the project can be found at <https://pagines.uab.cat/easit/en>. See footnote 26 for further information about ILSA.

3.8 Joining the dots: Respeaking and accessibility⁶⁰

As I explained in Chapter Two, I want this study to be one that pursues accessibility, and one that goes further than examining respeaking simply as a modality of access services. Eugeni (2020b: 21) highlights the fact that such studies around respeaking are uncommon: “Despite its numerous applications, training courses, conferences and research projects, studies on respeaking as a tool to promote social inclusion are scarce, especially when it comes to consider the latter in the widest meaning of Universal Design as promoted by the United Nations.”

Eugeni (*ibid.*) himself has provided examples of this more accessibility-oriented form of research through his work on respeaking and plain language, where the “main aim is satisfying the needs of people with sensory or intellectual disabilities” (*ibid.*), as tools for providing access to a wider audience, or, as he termed it, people with a “linguistic disability”. His study was in the context of interlingual translation, where a respeaker subtitled intralingually into plain language, which was then machine translated into ten target languages. A similar exploration of plain language was noted in the case of Simple Text respeaking for autistic children in Australia (see footnote 49). More recently, the EASIT project explored how audiovisual information could be made easy to understand. Their recommendations include suggestions for how this practice might be shaped in both live and pre-recorded programmes (EASIT, 2020: 29-35).

Eugeni (2015) has also taken a different direction in his quest for accessibility by exploring how different user groups can approach the practice of respeaking, in his work with blind respeakers. This broader view of accessibility, where the traditional roles of user and maker, at least from a media accessibility perspective, fall away, captures the essence of Desblache's (2019) major to minor metaphor, referred to at the end of Chapter Two. Other academics have expressed similar desires, exploring how ‘users’

⁶⁰ *Joining the Dots* is a documentary about blindness made as per the accessible film model (Romero-Fresco, 2013, 2019: 12).

can be more involved in the provision of access and also its instruction (Szarkowska, 2020).

3.9 The research gap this study aims to fill

Over the course of this chapter, I have presented the gap in the research that this study is intended to fill. Respeaking is a challenging process wherever it is conducted, yet when being introduced into a new setting it is likely to throw up many more challenges, which must be resolved if respeaking is to be considered viable for use. In addition, within the backdrop of the UK, there are many misconceptions and prejudices about the use of this modality which must also be addressed if potential audience members are to warm to its use in a setting where the use of STTRs is well established. The live event setting is one where more access is needed, and respeaking could be a solution to complement and expand the access that is currently available. Failing to expand its use in this way in the UK may be a missed opportunity, especially as it is a modality that is expanding and growing internationally.

In addition to entering the new territory of live events, there are a number of other aspects to the originality of this research: the consideration of the wider audience, the complementary aspects of training that will be provided that go past the skills required for respeaking alone, and the novelty in the focus on respeaking and accessibility alongside each other.

In order to achieve this, and to discover how to embed respeaking into this new setting, there is an inherent complexity within this research. Access provision itself is complex, and in order to embed respeaking into it fully, it is expected that as well as building on prior research, as illustrated in this chapter, much practical application of this technique will also be required to ensure a respeaking service for live events will meet the needs of potential users. Approaching this study through an action research methodology provides a framework that supports the complexity and practical aspects of this study, and at once draws together the principles of the triple theories on which this research is

built. This methodological framework underpins the research design, and is the topic of the next chapter.

Chapter 4: Methodology and research design

“No research without action,
No action without research.”

Lewin, as cited in Marrow, 1972: 90

In Chapter One, I described my journey as both a person and researcher and illustrated how certain life experiences have shaped the way I approached this study into the provision of respeaking and access at live events. Chapter Two developed this further as social and epistemic justice, along with intersectionality, and the social model of accessibility, were presented as the theoretical frameworks driving this research. In the current chapter, I return to the research questions posed at the start of the study, explain the methodology and give an overview of the research design developed in response to them.

Action research is the methodological framework adopted in this study. I begin the chapter by presenting the specific way in which I have used it, as I explain the three different cycles of my research design, and situate it within a broader commentary on action research. In doing this, I will demonstrate why action research is such a good fit to the work being carried out in this study and show how I have attempted to overcome certain criticisms which have been raised against it as a methodology. Next, the individual stages of the research and the specific research methods and tools used in each will be presented. In the first research cycle, focus group work was conducted using semi-structured interviews and questionnaires⁶¹. Alongside this, I took the opportunity to observe what access was currently provided at live events, and in what way it was implemented. The second and third cycles, involved two rounds of respeaker training and research events. Individually, the research events acted as case studies into how different types of events were run and together they formed a large audience reception

⁶¹ In section 4.2.1 I explain how I use the term focus group more fully to include the group sessions people attend, individual interviews and also questionnaires which many participants completed remotely.

study. During both cycles, data collection involved a NER analysis of the respeaking at the events and further questionnaires and focus group discussions.

Across the study as a whole, both qualitative and quantitative data were collected; the majority was qualitative, and the reasons for this will also be discussed as the chapter progresses, and in particular in section 4.1.6 below. More detailed explanations of the processes that were adopted will also follow in later chapters, immediately preceding the analysis of the findings which resulted during each cycle of research.

4.1 Combining theory and action: development of a methodology and research design

4.1.1 Research questions

The ultimate aim of this study was to find out whether respeaking could be introduced into the live event setting as a means of providing access for the audience. At no point was any attempt made to compare respeaking to other forms of provision, for example, BSL interpretation or stenography; rather, I sought to examine what would be needed to implement respeaking as a service and explore the benefits it could bring, including and extending beyond sensorial access, which was the original access it was intended to provide to DH audience members. To do this, three key research questions were posed:

1. What training and technical set-up is needed to allow experienced television respeakers to transfer their skills to the live event setting?
2. How can high quality respeaking and access be ensured?
3. Can the findings from this UK study be applied to the provision of access across borders and in society more broadly?

The first two questions took a more practical slant and directly shaped the methodology and the research design. Responding to them necessitated both the process and product of respeaking to be examined and one of the expected outcomes from these questions

was the creation of materials which could be used by others to establish this service, once quality expectations had been reached.

As explained in Chapter Three, whilst respeaking is commonly used in the UK on television, its use outside this sphere is very limited. In addition, it is a skilled profession. Before going live and having their subtitles broadcast for viewers to see, respeakers must understand the specificities of the software in order to be able to use it effectively; they must also have mastered numerous techniques to ensure that their respeaking is accurate and so that their stamina can be maintained. Examining the processes involved in respeaking and the technical support and set-up required for its effective transmission would be essential in order to establish it as a service at live events. Since respeakers regularly have additional training before beginning work on new programmes, it was assumed that some degree of training or upskilling would also be required before working at live events. Opportunities for the respeakers to provide access at live events would also be required to test how effective the training and set-up had been and these live events would also allow the quality of the provision, the product, to be assessed.

To fully answer the first and second research questions, it was necessary to test respeaking across the spectrum of live events; focusing on a single genre of live events would not be sufficient. This added greater complexity to the research design, as a wider range of features within live events needed to be examined; it was also reflected in the diversity of the audience, which included DH and wider audience groups.

Action research seemed to be an immediate fit to these research requirements and to the challenges that were implicit within the questions posed, as I will now explain.

4.1.2 Action research as a methodology

In the simplest terms, action research is “learning by doing” (O’Brien, 2001). A problem is identified and action is taken to resolve it; periods of reflection are built in allowing the effectiveness of the action to be assessed and further changes to be made as required. By doing this, the gap between ‘knowing’ and ‘doing’, which often arises during other types of applied research (Reason and Bradbury, 2008: 1) is removed. Whilst these other

applied studies might end by raising questions which remain unanswered, in action research, the aim is to answer these questions as the process unfolds and to do this in a way that is meaningful (Meyer, 2000).

Action research is an approach that is commonly used in settings such as education, healthcare and many areas within social sciences. Whilst it is not one that has been widely used in translation studies, it is one I was personally familiar with through my reflective practice as a teacher, the two being closely related (Leitch and Day, 2000). For this reason, I was confident in applying it to a new field, even one where few precedents existed.

It was Neves' doctoral study in 2005 in audiovisual translation that first presented the field with a substantial piece of action research. Prior to this, action research had been named as a possible research method (Hatim, 2001; Williams and Chesterman, 2002) or implied by reference to the need for an opportunity to bridge the theoretical and applied (for example, Díaz Cintas, 2004). All too often, it remains a research method that is listed without accompanying translation-related examples of its use or implementation being given (Bogucki, 2010; Saldhana and O'Brien, 2013). Where action research has been widely reported is within teaching and training contexts in translation studies (Cravo, 1999; Kiraly, 2000; Cravo and Neves, 2007) and more recently in audiovisual translation and media accessibility in the context of training for interlingual respeaking (Dawson, 2020). Aside from these, few other published examples exist of action research studies, at least not in name⁶². It is possible that other translation researchers "might have, at some point, used or enacted AR projects quite unaware of the scope or nature of their venture" (Neves, 2005: 49).

Neves (2012, 2016) herself has continued to implement action research widely in audiovisual and accessibility-related contexts, for example exploring how, in the latter, a museum could be made more accessible for all who visited; her work shows how action

⁶² In a personal conversation, Agnieszka Szarkowska said that many of her early projects were action-research-based, though few refer to the term directly; she cited Szarkowska, 2011 as an example.

research is both a way of thinking and project design; once adopted, it can be a very natural way of working, as has been in the case in this study as well.

What is interesting with action research is the way the approaches adopted within it can vary so greatly; as a methodology it can be crafted to the research in question. Both Neves' study (2005) and my own explore professional practice with a view to effecting change through the creation and future implementation of professional subtitling guidelines. The participatory nature of the research and professional collaboration are similar in both, yet differences exist, too, for example, in the theoretical approaches adopted and the contexts studied. Whilst we both sought to establish guidelines in new areas, prescriptive guidelines were highly appropriate in the case of SDH in Portugal, yet guidelines which encourage questioning and application are more suitable for my own study.

This is indicative of the inherent fluidity and flexibility of action research. It can be applied to new and shifting scenarios and there is a great degree of variation between different action research approaches. At the same time, there is an openness towards different practices and solutions being possible. Rather than being a prescribed methodology per se, action research can more usefully be understood as the orientation taken towards a particular inquiry (Reason and Bradbury, 2008:1). This flexibility means that the precise cycles included within the research design can be shaped to suit the specific questions posed. In this study, it meant that it could be adapted to fit a research project situated across TS, AVT and MA. In particular, it meant that the complexity needed to test different event types and to talk to different focus groups could be incorporated within the natural cycles that emerged.

However, it is not only because of this practical and flexible approach that action research was well-suited to this study. The collaborative and participatory nature of the research that it offers made it an even better choice because of the guiding principles behind the study. Social and epistemic justice and intersectionality are the overarching theoretical notions which underpin this research and the social model of accessibility has been presented as a framework whereby barriers created by society can be removed,

generating increased opportunities for this justice to occur. In line with this, and at the heart of this research, is the opportunity to create an access service that is universalist, user-centred, and proactive (Greco, 2018; Chapter Two above) in response to meeting the demands of the wider audience who attend live events. Whilst d/Deaf, deafened and hard of hearing audience members may be the initial or primary users of the service and consequently the people whose needs must absolutely be met, others will also use it. When I refer to the 'wider audience', I include all potential users of the service – audience members who both are and are not d/Deaf, deafened and hard of hearing. By considering the fullest possible range of users and expectations from the outset, the greater the likelihood is that this extended range of needs can be met by the eventual service suggested. Respect on many levels is fundamental to achieving such a service: respect for information, expertise and opinions that come from a range of people and respect for both the users of the service and the professionalism of those who provide it. From this perspective, every person is a holder of knowledge and expertise is considered as composite; it is sought from many sources and many groups of people. The third research question, which asks whether the findings from this study can be applied to the provision of access in society more broadly, reminds us of this theoretical approach and demands that these core principles drive the direction of the research and that they are both considered and respected in the choices made, acting as an ethical framework behind this research (Greco, 2018: 220).

The earliest application of action research, as set out by Lewin in the late 1930s and tested in the 1940s, was to "raise the self-esteem of minority groups, to help them seek independence, equality, and co-operation through action research and other means" (Lewin, 1946 in Adelman, 1993: 8). Lewin's aim was to solve social conflicts and overcome exploitation and colonisation which they had been exposed to (*ibid.*). Whilst some of these notions may not be relevant when discussing the experience of d/Deaf, deafened or hard of hearing people in society today, the quest for equality, or, as I argue, equity, certainly is. The activism inherent within Lewin's approach continues to resonate today and is an appropriate fit for social and epistemic justice (Given, 2012; Griffiths,

2012). This has fed through into many different branches within action research, including participatory action research (Jordan, 2012) and community action research (Ozanne and Anderson, 2010). Both have informed the current study, though it does not fall within either branch.

Therefore, action research presented a framework which allowed these core principles to be embedded within the research design, whilst offering the cycles of action and reflection required to answer the research questions and develop an effective respeaking service at live events. In this study it functions as both a philosophical approach and overarching methodology and, for this reason, it is the central feature of the project design.

4.1.3 The cyclic nature of action research

The three research questions posed at the start of this study were broad, in particular since they focused on the live event setting as a whole. Whilst the outcome was specific, how to establish a respeaking service, there was flexibility in the route that could be taken to get there. Three cycles of action research were identified which enabled the necessary data to be collected and allowed the questioning inherent in each cycle to become increasingly focused, as unsuccessful actions and strategies were improved or rejected. In this sense, action research can clearly be seen as a process of refinement and more precise questioning, as unknown territory becomes familiar. Cycle One is exploratory, seeking to determine the key issues which need to be investigated. By Cycle Two, with these issues identified, the focus shifts to establishing this in a practical way and designing the respeaker training programme and research events and by Cycle Three, enough is known for the emphasis to be placed on refining the process and giving consideration to the finer details involved in running live events.

These cycles and the research design as a whole were developed to suit the questions raised. Out of the many action research models that exist, two in particular were useful in shaping the design of the current study (figures 4.1 and 4.2).

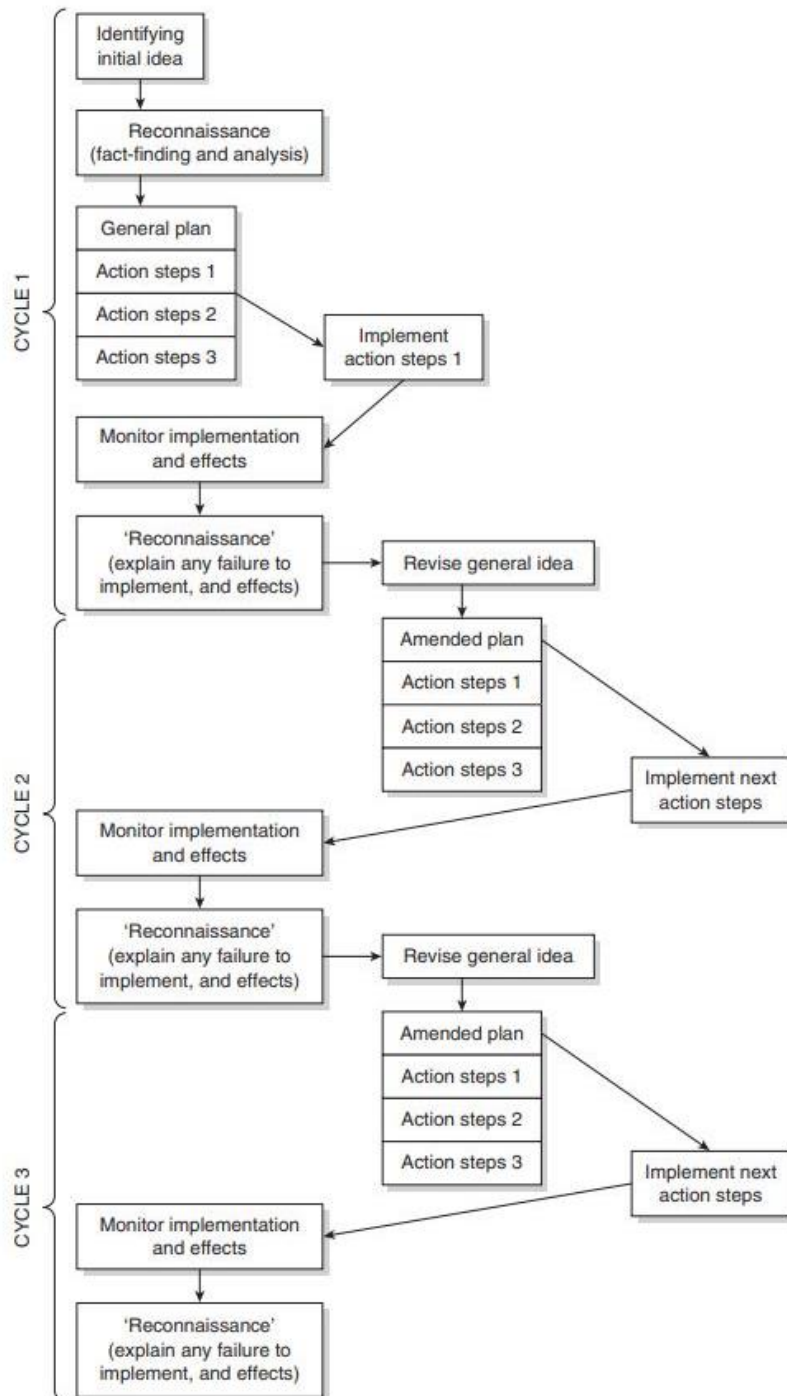


Fig. 4.1: Elliot's action research model, entitled 'What is action research?' (Elliot, 1991: 71 cited in Koshy *et al.*, 2010: 7)

SOURCE: 'Figure 1.2 Elliot's action research model' is from Elizabeth Koshy, Valsa Koshy and Heather Waterman, (2010) *Action Research In Healthcare* p.7, SAGE Publications, Ltd. It was adapted from 'Fig 6.2' from John Elliot, (1991) *Action Research For Educational Change* © 1991 Buckingham: Open University Press. Reproduced with the permission of McGraw Hill.

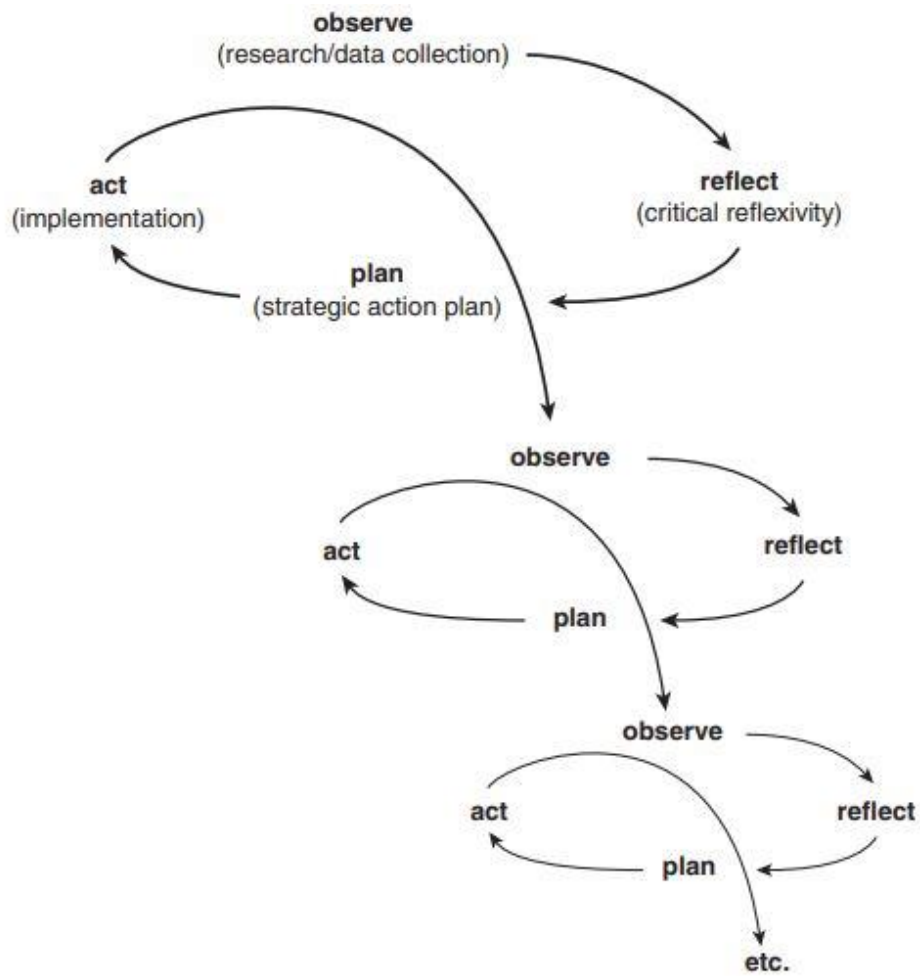


Fig. 4.2: O'Leary's cycles of research, (Koshy *et al.*, 2010: 8)

In my opinion, Elliot's model (figure 4.1) accurately captures the intricacies of the different steps involved in each cycle and highlights, in particular, the need for reconnaissance at regular intervals, including at the end of the study. As knowledge shifts, this fact-checking opportunity and the chance to align new knowledge with what was previously seen was particularly useful. O'Leary's (2004: 141) model in figure 4.2, on the other hand, visually demonstrates the importance of the cyclic process in directing the research that follows. Over the course of the different cycles, the focus of the research becomes more and more precise, as enough knowledge has been gained to reject certain options.

4.1.4 Research design

These three cycles of data collection were planned across a period of two years, and each with a period of reflection on the action taken and results collected (table 4.1). Full Ethics Committee approval through the University of Roehampton was obtained for all stages of the research. A sample consent form can be found in appendix 4.1.

4.1.4.1 Cycle One

The first cycle was primarily one of reconnaissance and consultation and involved exploring the access already in place at live events and determining what the expectations of each of the key parties involved in a respeaking service, both users and providers, would be. It was designed to complement and build on my previous experience of respeaking, since it took me directly to access in the live event setting. Consulting Stagertext, the leading charity who provides captioning at theatre performances and live subtitles at talks and tours, was the first exploratory action taken. This was followed by a series of visits to a performance, talk and tour, which enabled me to observe the process behind access being provided and to experience it directly from within the audience. Towards the end of the first cycle, I also attended two events and respoke as-live. Whilst no subtitles were broadcast to the audience, this provided an opportunity to experience respeaking at live events first-hand.

Throughout the research, I also attended many live events in the UK and further afield. Some were social events and some were professional, for example, conferences where I presented my own research. Some were captioned or subtitled, many were not. All contributed to my own experience and knowledge of live events, as a presenter, audience member and researcher, and gave me an opportunity to reflect on the key questions explored in this research. Action research travelled with me, and took place in more and less formal settings. As well as being a recognised methodology, it has become my personal approach taken to my research and one that pervades many areas of my life. Just as I had been a reflexive teacher, and had become a critical researcher, I have a sense of curiosity in all I do.

Table 4.1: Research Design

Cycle One – Initial reconnaissance (2016)				
Questions posed	Actions taken	Format of data collected	End of cycle Reflection	Outcomes
<ul style="list-style-type: none"> • What is the nature of the access currently provided at live events? 	<ul style="list-style-type: none"> • Consultation with Stagertext • Observations of captioned and subtitled live events 	<ul style="list-style-type: none"> • Observations • First-hand experience 	<ul style="list-style-type: none"> • What overlap is there between the expectations of the different parties? • What will the likely technical set-up at live events be? • What equipment will be needed? • What training will professional TV respeakers need to begin working at live events? 	<ul style="list-style-type: none"> • Technical set-up determined • Familiarisation with equipment already purchased • Version one of the training programme for respeakers
<ul style="list-style-type: none"> • What expectations do the key parties have for a respeaking service at live events? 	<ul style="list-style-type: none"> • Talking to key parties (Venues, respeakers, users) through focus groups and interviews 	<ul style="list-style-type: none"> • Interviews • Questionnaires 		
<ul style="list-style-type: none"> • What are the practicalities of respeaking at a live event? 	<ul style="list-style-type: none"> • As-live respeaking at events (not broadcast to the audience) 	<ul style="list-style-type: none"> • First-hand experience 		



Cycle Two – First round of respeaker training and research events (2017)				
Questions posed	Actions taken	Format of data collected	End of cycle Reflection	Outcomes
<ul style="list-style-type: none"> • How easily can the respeakers adapt to the new equipment being used? • What questions do they have about respeaking at live events? • What procedures need to be put in place? • Do the respeakers feel confident about working at live events? 	<ul style="list-style-type: none"> • Training provided for a group of respeakers 	<ul style="list-style-type: none"> • Observations • Photographs, audio and video footage of training • Discussions • Questionnaires • Participant reflections 	<ul style="list-style-type: none"> • What adjustments need to be made to the training programme? • Does the technical set-up need to be modified? • Are any other changes needed? 	<ul style="list-style-type: none"> • Technical set-up revised • New equipment purchased if needed • Supplementary training modules provided for the respeakers
<ul style="list-style-type: none"> • What information does everyone involved need? • Is everything in place? • What will the quality of respeaking be at the live events? 	<ul style="list-style-type: none"> • Four research events held as part of a large reception study • Focus group discussion held at the end of each event 	<ul style="list-style-type: none"> • Observations • Photographs, audio and video footage of each event • Transcripts of the respeaking • Discussions • Questionnaires • Participant reflections 		
<ul style="list-style-type: none"> • What is the nature of the access currently provided at live events? 	<ul style="list-style-type: none"> • Continued attendance and observation at other live events 	<ul style="list-style-type: none"> • Observations 		



Cycle Three – Second round of respeaker training and research events (2018)				
Questions posed	Actions taken	Format of data collected	End of cycle Reflection	Outcomes
<ul style="list-style-type: none"> • Are the respeakers independent in their work? 	<ul style="list-style-type: none"> • Supplementary training provided for a group of respeakers 	<ul style="list-style-type: none"> • Observations • Photographs, audio and video footage of training • Discussions • Questionnaires • Participant reflections 	<ul style="list-style-type: none"> • Are any further adjustments needed to the training programme? • Is the technical set-up effective? • Has suitable equipment been purchased? • Are any other changes needed? 	<ul style="list-style-type: none"> • Finalised version of the training programmes produced • Guidelines created for RLE • Information and training resources created for presenters and others involved in holding live events
<ul style="list-style-type: none"> • How effective is the technical set-up and equipment? • Will the events run smoothly? • Will the quality of the respeaking be of a high enough standard? • Is respeaking a feasible method for providing subtitles at live events? 	<ul style="list-style-type: none"> • Four research events held as part of a large reception study • Focus group discussion held at the end of each event 	<ul style="list-style-type: none"> • Observations • Photographs, audio, and video footage of each event • Transcripts of the respeaking • Discussions • Questionnaires • Participant reflections 		
<ul style="list-style-type: none"> • What is the nature of the access currently provided at live events? 	<ul style="list-style-type: none"> • Continued attendance and observation at other live events 	<ul style="list-style-type: none"> • Observations 		

As McNiff (2013:19) wrote, following Elliott (2007), “I believe that action research should be appreciated as a kind of practical philosophy, and even practical theology, inspired by an enduring sense of awe and wonder as we ask questions about what we should do and why we should do it”.

An initial round of focus group research then followed with user groups and providers. Since creating and watching subtitles at a live event is a different experience from doing so on television, it was essential that both were consulted. Further focus groups followed in the later cycles and so provided a structure to the research design and an opportunity to pause and reflect within the more practical aspects of the events (Fig. 4.3).

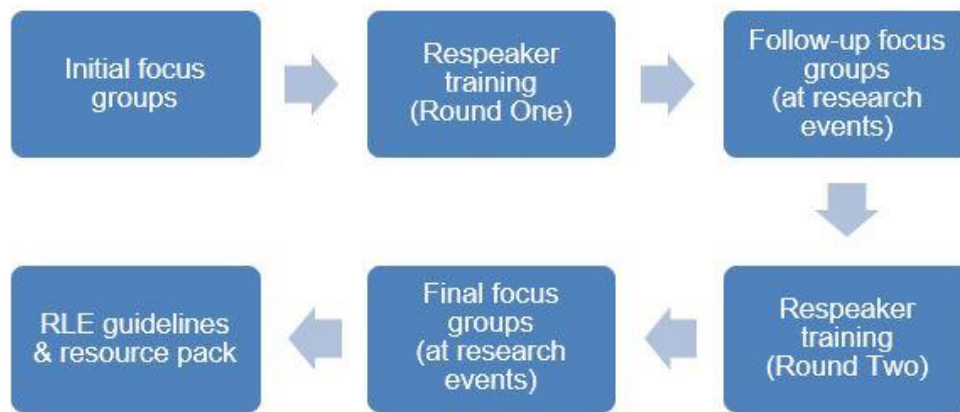


Fig. 4.3: Interweaving structure of focus groups and respeaker training (Moores, 2020a: 182)

In this way, the people who contributed to the process had a voice throughout and were as involved in the reflective stages as the periods of action.

A total of four different parties were invited to the initial round of focus groups. Two of the parties were user groups, made up of potential audience members. Respeaking at live events was initially targeted at DH audience members as it provides sensorial access to events and this focus group had the largest reach. As a representative group of the wider audience, interviews were also held with hearing non-native speakers of English (NNE) as there is evidence to suggest that intralingual subtitles can facilitate and support language learning (Vanderplank, 2016; section 3.5.2 above). In order to consider how to set up the service, two further focus groups consisted of the service providers, i.e. the venues that host events, and the respeakers themselves. Semi-structured interviews were held with members of all four groups. In the case of the audience users, they were held as groups sessions; the others were for the most part individual. The data collected from the DH users and respeakers was used to create a questionnaire which could be distributed for a greater reach.

The period of reflection which followed this first cycle was used to review what additional equipment would be needed to trial the respeaking during training and at the live events and to design the first version of the training programme for the respeakers.

4.1.4.2 Cycles Two and Three

The next two cycles of action research and data collection were comprised of two rounds of respeaker training followed by research events where the respeaking was tested in action. Both were concerned with determining whether respeaking could be used as a service for providing access at live events and began with a period of training to equip the respeakers with the skills and knowledge needed to do this. The second cycle of action research provided the opportunity to test what had already been created and the third cycle allowed this to be further refined.

Discussion was a central part of the initial respeaking training programme (Cycle Two), since a key purpose of it was to gauge the respeakers' responses to what they were being asked to do. It was an opportunity to determine how easy it was for the respeakers to adapt to the new equipment being used and to find out what questions they still had about respeaking at live events. The training was very collaborative in its nature, with procedures and the content of future modules evolving from all involved. In this sense, the respeakers were informed participants, which is very much a feature and advantage of action research, as will be illustrated further below. The second round of training, in Cycle Three, was an opportunity to fill in any gaps in knowledge revealed through research events 1-4 and therefore took the form of supplementary modules. It was also an opportunity for further reflection and discussion in a more relaxed setting, away from the intensity of a live event.

The research events provided an opportunity for live testing of the respeaking before an audience and were also an opportunity for other participants, including the venues and presenters, to experience respeaking in action. The first four events (Cycle Two) were exploratory. They focused on testing how well the equipment worked in different venues and contexts and how well received the subtitles were by the audience. The final four events (Cycle Three) were far closer to actual events, since the processes and routines had been established and the respeakers had received their supplementary training. The audience experienced subtitles that were far closer to a 'final product'.

The eight events functioned together as a large reception study and also acted individually as case studies, since there was diversity of content and location at each. Photographs, video and audio footage were recorded at each event and a respoken transcript was retrieved which, together, allowed a NER quality analysis to be performed. Questionnaires were given to all involved – the audience, respeakers, presenters and key venue staff. All audience members received the same questionnaire, regardless of their hearing level or native language. The demographic questionnaire nevertheless asked participants to identify their hearing and native language status, which meant that perspectives of these groups could be sought within the wider audience view. One additional focus group was added in Cycles Two and Three – the speakers and presenters who took part in the events. This group straddled the divide of user and provider. It was at the end of each event that the second and third layers of the focus group research took place in the form of a post-event chat. Everyone involved was invited to attend a semi-structured discussion at the end of the event. Unlike the focus groups of Cycle One, here the different parties, including the respeakers, joined in a single session.

The final period of reflection, which followed Cycle Three, was an opportunity to assess the quality of the respeaking seen across the two rounds of events and to compare this to other respeaking corpora. Final adjustments were made to the training programme for respeakers and the guidelines and resources for facilitating the Respeaking at Live Events (RLE) training and making live events accessible through respeaking were created. Also included within this pack were training resources for the different parties, since the cycles of action research had revealed this to be a necessary outcome of the project in ensuring the feasibility of respeaking at live events.

4.1.4.3 “Messiness”⁶³ within the action research process

Action research is a process that takes the researcher, and all who work with them, from their initial ideas and expectations, step-by-step, to a viable solution. However, this is not always a straightforward process. The action research models outlined in 4.1.3 above (figures 4.1 and 4.2) and my own (table 4.1), are likely to give the impression that the steps were clearly mapped out and taken in a progressive way. The reality of action research is that it is often a far more messy and convoluted process (Cook, 2009) than any table can sensibly capture, yet through this messiness, with its challenges and, perhaps, moments of desperation, a feasible service emerges.

Some adaptations to the research design became necessary as the project evolved. The design presented above is the one that was followed and key changes are outlined in the sections that follow. Certainly, the path taken was not as linear as table 4.1 suggests. Separate strands of research often happened concurrently. Cycle One had been designed to include only as-live respeaking; however, it became clear that the DH focus groups would need to be subtitled and these became the very first events where respoken subtitles were streamed to the audience. Scheduling the observations, interviews and events involved many people and was a complex process. The learning that occurred did not simply happen between cycles; it also happened from event to event, from module to module and even from moment to moment as the research progressed and everyone involved reflected on what they had been part of. Whilst I had envisaged the reflection period between Cycles Two and Three as the time when changes to the set-up and running of the events would be made, in practice, improvements to both the equipment used and procedures followed were implemented as they came up, between sessions and events. Added to this reflective process, live events themselves are dynamic and on-site, on-the-day peculiarities, ranging from participants getting lost to power cuts, meant that plans often needed to be changed.

⁶³ Cook (2009) writes of the purpose of mess in action research.

However, this is also the reality of working in live events, so gaining this experiential knowledge was also an essential element of the action research. It also served to shift the focus of the guidelines which resulted from the process from moving away from a series of steps that, if followed, would ensure access, to a series of questions that when considered, would lead to improved access.

4.1.5 Rationale for mixed methods

Action research studies frequently adopt eclectic methods of data collection (Given, 2012: 4) and this is certainly true here. The data collected across the cycles was a mixture of qualitative and quantitative and this mix seemed in keeping with the principles behind the study as a whole.

The qualitative data collected through interviews and discussions allowed a picture to be built of the experiences of each group and for individual voices to be heard. As expressed in Chapter Two, norms and labels only go so far as an identity and not everyone with one label will think, act, react to or experience an event, or the access surrounding it, in the same way. With this approach, there was space to see how accounts differed, as well as the opportunity to create some kind of representative voice. Action research, with its scheduled opportunities for pause and reflection, and its participatory nature, accommodated this well.

This qualitative approach also reflected the nature of the questions posed and the solutions sought. Tackling all genres of live events meant that guidelines would need to be applicable to different scenarios. As already outlined, a more fluid, descriptive approach to introducing and extending access would open the way for everyone to get involved and take ownership of the access provided; trying to apply prescriptive guidelines to very diverse events would be far more complex and potentially problematic⁶⁴. A good example of this would be where the subtitles should be located. Whilst the bottom of the screen is the position that most people associate with subtitles,

⁶⁴ The notion of flexibility within access provision, and need for a descriptive rather than prescriptive approach, is also strongly embedded within Lazard's *Accessibility in the Arts: A Promise and a Practice* (2019: 10).

on some occasions, this is not the most appropriate position for them, because of the lighting, the position of the stage, or even the height of the presenter.

Nevertheless, the quantitative data was an essential element of the research carried out. In Cycle One, the use of surveys meant that the opinions and experiences of a larger number of people, from an extended area, could be sought and added a certain weight to the experience shared, since trends could be followed in a different way. In Cycles Two and Three, the data collected in the audience questionnaires meant that responses could be compiled in a more succinct way. The accuracy score produced through NER analysis and related quantitative data also allowed the corpus of respeaking at live events to be compared to that seen in other corpora of respeaking in English (Chapter Eight below).

However, figures in isolation offer a very specific angle on a situation and can be open to (mis)interpretation. Even within the NER analysis, the qualitative comment that situates the accuracy score within the context of the respeaking is an essential element of the feedback it provides. Within this study, when interpreted alone, neither the qualitative or quantitative data is sufficient to respond to the three research questions posed. It is only the combination of what they both reveal, following the whole process of action research and periods of reflection, that provides the answers that were sought. Even within the quantitative data, much of the analysis remains a descriptive one. This is due in part to the numbers of participants involved in the study and also to the decision to compare different data strands, specifically the NER score and the audience score, to attempt to ascertain to what extent the audience view reflected the technical analysis of the respoken subtitles. Since the composition of the data differed, descriptive statistical analysis was the only method possible. This will be explored further in Chapter Seven below.

4.1.6 Criticisms raised against action research

Situated as it is in opposition to more traditional research methods, a number of criticisms have been made of action research since its conception in the 1940s; they are frequently presented as limitations and there are two in particular that I wish to address here.

The first relates to how objective the researcher is able to be when they are so actively involved in what is happening. In response to this, I would agree that a researcher is not fully impartial, but I do not see this as a negative; rather, it simply marks a different approach from a more scientific, detached approach. I readily acknowledge the extent to which my personal story and experience has shaped this research and celebrate the innovation and originality this brings. The fact that this is doctoral research makes my personal investment in the process even greater. However, by acknowledging this and developing a reflexive practice, I remain aware of this positionality. In addition, since so many different participants are involved, each bringing their own lived experience, any overly subjective views I might hold are balanced. In this way, the periods of reflection in the action research include personal observation as well as reflection on the action research 'proper'.

Another key criticism which has been raised against action research is that by being grounded in a single experience, it loses its potential to be replicated (Cohen and Manion, 1985: 216; Cravo and Neves, 2007). By holding two cycles of respeaker training and two series of research events, I have sought to ensure that a range of situations are tested and the observation and as-live testing period has further reinforced this. By including a diverse range of events, I have ensured that respeaking can be tested in a range of situations and different features which need examination can be revealed. Treating these events as a larger reception study also allows a strong audience response to be collated and compared to the more technical analysis gained from the NER model. This triangulation will add further weight to any conclusions drawn. Approaching the events as case studies ensures that all guidelines and recommendations are embedded within the context of an actual event and the non-prescriptive approach adopted in the final recommendations encourage inquiry and considered application in their

implementation. Whilst it is true that attempting to follow my exact research cycles and scenarios might not be possible, the outcomes have been investigated sufficiently to be robust and replicable.

Having situated my approach within the broader action research landscape, I will now look more closely at the steps within the process and choices behind the research methods and tools used.

4.2 Stage-by-stage – the selection of research methods and tools

4.2.1 Focus groups in Cycle One and beyond

The purpose of the focus groups was to establish what expectations around the use of providing access through respeaking were held by the different parties who would be involved in or affected by this service. It was hoped that the focus groups with users, in particular, would provide an opportunity for those involved to reveal personal accounts and experiences of accessibility on television and live events and that, within these discussions, the features that were most important for each group would be revealed. In the case of the providers, the respeakers and venues, the focus was more practical and sought to elicit the processes involved in providing access and organising events.

In its narrowest sense, a focus group refers to a session where multiple people come together, in a space of discussion and interaction, for the elicitation of ideas with key members of the groups (SAGE, no date). I use the term in an extended sense and also refer to the broader principle of creating a space for eliciting the views of key members of the group, as well as to the actual group sessions. For this reason, a combination of interviews, with both individuals and groups, and questionnaires were used to collect the focus group data.

The interviews with the user groups, both DH and NNE, were held as traditional focus groups, apart from one person who was interviewed individually as he was unable to attend the main group sessions. I hoped that the group sessions would create a space where different points of view could be shared (Puchta and Potter, 2004) and for that reason, I chose a semi-structured approach (Kvale and Brinkmann, 2009:130). While I

had a list of specific items that I wanted to cover, and resources with specific examples to help stimulate the discussion, I also wanted the participants to be able to direct where the discussion went and, as researcher, to feel that I had the freedom to follow this natural flow. Through this approach, I expected topics I had not envisaged to be addressed (Holliday, 2007). During these sessions, my role was to facilitate the discussion. Whilst the participants were aware of my own experience of subtitling, they were also told that I had no agenda beyond hearing individual points of view. The conversation that followed from topics I brought up was directed by the participants. (*ibid.*: 33). The sessions lasted 2.5 hours and included a short break, where everyone could chat informally. Access was provided through a BSL interpreter and two respeakers working in tandem. The sessions were recorded.

The points raised during these group sessions were incorporated into a wider survey, so that the views of a broader range of members of the DH community could also be sampled. The quantitative data provided by the surveys would also provide the means for a statistical account of the trends and patterns seen (Brannen, 1992), and help ensure the replicability referred to above. Care was taken when designing the survey to limit bias and contextual influence, potentially inherent in any survey (Foddy, 1993). To achieve this, the survey was modelled on one which had previously been used in a large reception study on subtitles for the DH audience (Romero-Fresco, 2015 and see section 3.2.4 below) and advice was sought from Stagertext to ensure that the terminology and language used were clear and appropriate.

The interviews with NNE followed a similar pattern to the DH groups, but were run on a smaller scale. Whilst running these sessions, it was recognised that the NNE focus group was so diverse that detailed qualitative accounts could prove more useful than trying to draw quantitative data. For this reason, no attempt was made at a broader survey of this group. Rather, NNE were invited to participate as audience members at the research events.

The interviews with providers could not be run as traditional focus group sessions, due to scheduling restrictions amongst the respeakers and the need for a site visit to be incorporated into the interviews with venues. Instead, they were held individually. As with the DH focus group, a questionnaire was also created for the respeakers to gain a wider range of views. I interviewed respeakers from three companies and I visited five venues in and around London, some of which became venues where the research events were held. The venues included the V&A (vam.ac.uk), the Natural History Museum (nhm.ac.uk), the Royal Society (royalsociety.org), the Wellcome Collection (wellcomecollection.org) and the Riverhouse Barn (riverhousebarn.co.uk). Once again, semi-structured interviews were used.

I used MAXQDA for the qualitative analysis and carried out the quantitative analysis in Qualtrics.

The focus groups held at the events were necessarily shorter, as they followed the events which were for the most part held in the evening. Here, running mixed focus groups, attended by all parties, was a practical approach and it allowed the cross-fertilisation and sharing of experiences between all involved. For the most part, the discussion was guided by the participants, but I also asked a few questions which related to the specific content of the event in question.

4.2.2 Observation and as-live respeaking

Observation played a central role within the action research process. Sometimes formal, sometimes less so, it gave me, as researcher, the opportunity to step into different roles and gain a fresh perspective on what had been evidenced to date.

The three observations organised through Stagertext were arranged more formally and were opportunities to see captioning and live subtitling in action, as an audience member, but also with the opportunity to explore behind the scenes, gaining insight into what I would need to implement at the research events and in the resulting guidelines.

As researcher, I chose not to respeak at the research events myself. I wanted to remain in a more neutral role. However, I did want to gain experience of respeaking at live events

myself. For this reason, I arranged two opportunities for as-live respeaking. For both, I was at the venue, but respeaking in a room that was separate from the main event, to avoid intruding upon it. Both allowed me to respeak in real-time and give consideration to where the subtitles would have been displayed if they were being transmitted live. In particular, I was able to test respeaking through a remote sound feed, without visual access to the event, perhaps the most 'extreme' of situations.

4.2.3 Respeaker training

Whilst not a research method, running the training programme for the respeakers was an essential element of the research design. The training offered the space and opportunity for the collaboration and discussion that is so vital to action research. Whilst I was training the respeakers and sharing information about the project, expectations of the different focus groups and likely scenarios at live events, they were sharing their knowledge and years of expertise of respeaking, ensuring the solutions we found respected those. In the same way as the semi-structured interviews allowed the participants to guide the direction of conversation, in the training, time was incorporated for discussions which could also be participant-led. Once this exchange of information was established, the respeakers became informed participants, which meant that the content covered in training and decisions made at the research events were critically examined and evaluated by both the respeakers and myself as researcher. The role of the respeakers in this study differed from other studies where attempts are made to record and analyse work within a natural environment (O'Brien, 2006; Teixeira and O'Brien, 2017); here, it is understood that the environment they will be working in is new and different.

In this instance, rather than using action research to explore what style of training should be used when training respeakers, the focus was two-fold: firstly, what content was required in the curriculum design so that future live event respeakers would gain the skills they needed; secondly, what features of respeaking would be trialled at the research events in order to inform the guidelines that would result from this study. The

latter led primarily to a discussion over whether verbatim, standard or edited respeaking would be tested⁶⁵. The actual style of study incorporated trainer-led content, opportunities for discussion and hands-on experience, all of which will be explored further in Chapter Six.

The original research design was to train two groups of eight respeakers, the first eight in Cycle Two and the second eight in Cycle Three. In this scenario, the revised training programme would have been introduced in Cycle Three. Due to the commitment involved, the training period and certain restrictions imposed by the access service providers, it was necessary to adapt the research design. Rather than training a second group of respeakers, the same respeakers took part in both rounds of testing and only four respeakers were trained. Reflecting back on this process, this modification was advantageous, since the respeakers could grow in their role as informed participants as the cycles of action research evolved and gain a more in-depth experience of working at a range of events. Further, it meant that the focus of Cycle Three was placed on the design of supplementary training modules and on refining the processes involved.

4.2.4 Reception study at the research events

4.2.4.1 *The use of reception studies*

A total of eight research events were held over the course of a year which together formed part of a large, multi-site, multi-method reception study (Hill, 2018) into the quality of respeaking at live events and audience opinion of them. The first four events ran in autumn 2017 and the second four in the summer of 2018, allowing a period for analysis, reflection and further training between each round. Planning the events with periods of two weeks between each, allowed time for reflection and any changes to be implemented.

Reception studies are used in many disciplines; they differ greatly in how they are run, the focus of the study and in the way findings are interpreted (*ibid.*). However, their

⁶⁵ This idea of a scale of subtitles from verbatim, through standard, to fully edited draws on the work of Szarkowska *et al.* (2011).

primary purpose is to explore how the audience responds to (aspects of) a specific product, be it a written text, broadcast material or public event. Whilst a consideration of the audience has always been part of translation studies, and audiovisual translation more specifically, it is only more recently that attention has been placed on them in this more active way (Chaume, 2018; Di Giovanni and Gambier, 2018) and the frequency of these studies has grown. At the same time, as technology and modes of consumption have evolved, audiences have become more actively involved in how they engage with new products (*ibid.*), making their reception of products all the more important. From the media accessibility perspective explored in Chapter Two, this mirrors the shift towards a user-centred approach, which is central to this research.

A wide audience were invited to attend these events, which, in line with the universalist approach, included people who were and were not DH. However, the question of who the audience were in this reception study requires further consideration on a conceptual level. Abercrombie and Longhurst (1998: 40) define the audience as “groups of people before whom a performance of one kind or another takes place”. In this sense, the audience were the people who attended each of the events. When analysing their responses, I explore individual responses and group responses, both event-by-event and for all the events combined. However, my concept of ‘the audience’ is a more active, and indeed interactive notion than this. Abercrombie and Longhurst go on to identify a kind of exchange effect, whereby the behaviour of the performers changes under scrutiny – whether intended by the audience and whether conscious or not for the performer. In this sense, there is a fluidity within the audience-performer relationship. They add that performance in an inclusive concept (Schechner, 1988; Abercrombie and Longhurst, 1998), taking place in any setting where there is contact between people. This fluid, interchangeable and interactive audience view is very applicable to this study, and is expected much more in communication in the present day. The talks, tours and presentations to which the audience were invited were opportunities for contact, communication and dialogue. Whilst people were present in a localised way at each of the venues they attended, they also brought with them their experiences of the world,

from the so-called performative society they live in (Abercrombie and Longhurst, 1998). Further, boundaries between audience members and presenters were blurred as each event unfolded and members of the audience took the microphone to speak and the presenter listened. Who was 'performing' to whom at this point?

Since this study explores the establishment of a future service, one further notion of the audience needs to be explored: the potential audience, the people who could attend these events made accessible through respeaking in the future and benefit from the input of the audience members who took part in this reception study. Here, the notion of the transnational audience (Althique, 2016) is very pertinent. Today, exchanges between local, national and global flows have become the norm for audiences around the world. I think this is as true for a live audience as it is for the sharing and screening of digital content. It is impossible to predict who will attend an event. The audience could include people who are and are not DH, who live locally or who are passing through while on holiday, people with different levels of interest in the event theme, even including one or two who have no interest at all but who are attending to accompany someone or simply because the event is accessible. People attend an event for all kinds of reasons. As Hill (2018: 18) writes, "it is important to set aside assumptions about audiences and conduct research with people, to make the human side of audience research centre stage." This is what I have tried to do here, by creating dynamic and social opportunities for people to come together, share an event and collaborate in this research study.

4.2.4.2 Event design: Case study approach

The scope of these events was broad in order to be able to respond to the research questions posed, and they included presentations, public speakers, film panels, museum tours and post-screening Q&As. This design allowed respeaking to be tested across single/multiple speakers, a seated/moving audience, diverse technical set-ups in and outside the event room and varied visual and spoken subject matter. Variation in the venue, event type and content were planned in advance; the variations in the working set-up were determined on-site according to the specifics of the location. In a similar set-

up to television work, the respeakers worked in pairs to provide the access at each event. Whenever individual schedules allowed, the pairings were alternated to facilitate discussion and reflection across events.

By approaching these events as case studies, the research also benefits from concrete, contextually situated examples of respeaking in action. As outlined in section 4.1.6, one criticism raised against action research is that it is harder to replicate. Whilst case-studies are also very specific in what they reveal, they do encourage an in-depth perspective. Given that an outcome of this research is to share resources for implementing respeaking in the live event setting, detailed examples in the form of case-studies will be particularly useful; the data collected here can be used to create them. Since individual aspects of the different events varied, understanding the context of the features in action will also be more useful than any attempt to isolate the exact cause of individual trends that are seen. Whilst the way in which certain aspects can be observed across all the different events, and even compared, it will not be possible to isolate the cause of every trend. Case studies will provide a useful point of reference for the interpretation needed.

4.2.4.3 Data from the reception study

Data was collected from all who attended the events – the audience, representatives from the venues who hosted the events, the presenters and the respeakers themselves, though the data collected from the audience was by far the most extensive. It came in the form of responses to short questionnaires, participation in the post-event focus group discussion and also through post-event reflections which some of those who attended chose to share. It therefore included qualitative and quantitative elements. The procedures followed changed slightly between rounds to make the research element of the events less obtrusive.

The questionnaires sought to determine the overall experience created by the inclusion of respeaking within the event and focused on features which were thought to relate to each group's experience of and interaction with them. For example, the questionnaire for the respeakers inquired into their working environment, whereas the presenters' was

guided towards how their behaviour might have changed as a result of the subtitles. Most questionnaires were given out at the end of the event, though the respeakers also had a pre-event questionnaire to try and track their experience of the event as a whole.

As explained in section 4.2.1 above, the second and third layers of focus group research were carried out as post-event discussions, usually following a short interval with drinks and snacks. Everyone involved in the event was invited to attend. The most usual make-up of these groups were for there to be many people from the audience and, sometimes a representative from the venue. The presenters attended on three occasions and the respeakers took part in all the discussions, sometimes respeaking their comments directly on screen. I used SPSS to analyse the quantitative data that was collected, whilst the qualitative responses were entered into MAXQDA.

The data collected from each group was also compared to the technical NER analysis of each event to explore the extent to which the experiences of those who attended and were involved aligned with the industry-style score assigned to each event.

4.2.4.4 NER analysis

The NER model (Romero-Fresco and Martínez, 2015) is the baseline for Ofcom's 2013-2015 (Ofcom, no date) review of quality in live subtitling and provides an assessment that has been proven as valid and which is user-focused (Romero-Fresco, 2016). During this analysis, recognition and edition errors are classified as serious, standard or minor according to the impact that each error has on the audience and penalties are deducted accordingly.

In line with the methods adopted by Ofcom, a sample of 10-16 minutes from each event was analysed using the NER model. Care was taken to ensure that the samples were representative of the different events, based on the known features of each and on my own observations. The trends seen within the live events corpus was compared to other known corpora of respeaking, notably the Ofcom corpus referred to above and the LiRICS corpus, gathered in the process of certifying professional respeakers (Romero-Fresco *et al.*, 2019).

Initially, the NER model was used to analyse the respeaking at live events exactly as it had been conceived for intralingual respeaking on television. However, as the analysis proceeded, it became clear that certain error types appeared within the live event setting that stretched the regular pathways of analysis present within the NER model and a modified version of the model, the NERLE, adapted to live event analysis, was proposed to capture the pathways I had effectively applied.

4.3 Analysis of the findings

The combination of the qualitative and quantitative data collected across these three research cycles allowed the three research questions posed at the start of the study (section 4.1.1 above) to be answered. This data will be analysed over the next four chapters.

In Chapter Five, the first cycle of action research will be presented and the data from the initial round of focus group testing will be analysed. In Chapter Six, the respeaker training programme and research events are presented, and the way in they evolved in the course of the study, following the action research approach, is explained. The findings from the research events of the reception study are split not by cycle, but by the nature of the analysis. In Chapter Seven, the analysis of the qualitative and quantitative data collected from the different focus groups during the reception study is presented, capturing the process of making the events accessible through respeaking; in Chapter Eight, it is the live subtitling itself, the product, that is the focus of the analysis as the corpus of respeaking at live events and the way in which it can be assessed is explored.

Chapter 5: Engaging with stakeholders

“Nothing about us without us.”

Charlton, 1998

In the previous four chapters, I have illustrated how my own journey as both an individual and researcher has shaped the direction of this project and the principles adopted within it and presented the theories that underpin this research. In the last chapter, as I outlined my research design, I explained how these theoretical stances combine so well with action research as a methodological framework and illustrated how they could be successfully incorporated into audiovisual translation and accessibility studies for the purpose of this research.

Chapter Five, the current chapter, is the first of four chapters which present and analyse the three cycles of action research. The data in this chapter comes from the first cycle, where I engage with various stakeholders for the first time.

5.1 Overview of Cycle One: Initial reconnaissance

Table 5.1: Cycle One: Initial reconnaissance

Cycle One – Initial reconnaissance (2016)				
Questions posed	Actions taken	Format of data collected	End of cycle Reflection	Outcomes
<ul style="list-style-type: none"> What is the nature of the access currently provided at live events? 	<ul style="list-style-type: none"> Consultation with Stagertext Observations of captioned and subtitled live events 	<ul style="list-style-type: none"> Observations First-hand experience 	<ul style="list-style-type: none"> What overlap is there between the expectations of the different parties? What will the likely technical set-up at live events be? What equipment will be needed? What training will professional TV respeakers need to begin working at live events? 	<ul style="list-style-type: none"> Technical set-up determined Familiarisation with equipment already purchased Version one of the training programme for respeakers
<ul style="list-style-type: none"> What expectations do the key parties have for a respeaking service at live events? 	<ul style="list-style-type: none"> Talking to key parties (Venues, respeakers, users) through focus groups and interviews 	<ul style="list-style-type: none"> Interviews Questionnaires 		
<ul style="list-style-type: none"> What are the practicalities of respeaking at a live event? 	<ul style="list-style-type: none"> As-live respeaking at events (not broadcast to the audience) 	<ul style="list-style-type: none"> First-hand experience 		

Three key questions were posed at the start of this cycle. Firstly, I wanted to determine the nature of the access currently provided at live events, which I did through consultation with Stagertext and observations of captioned and live subtitled events. Secondly,

through talking directly to key parties, made up of respeakers, venues and user groups, I wished to establish the expectations that each group held for a respeaking service at live events and ensure that various sources of expertise were embedded in this research from the outset. Finally, having gained this understanding, I wanted to experience respeaking at live events for myself. I attended two events where I respoke as-live, without my subtitles being broadcast to the audience. In addition, since the DH focus groups were made accessible through respeaking and BSL-interpretation, I also had the opportunity to experience being respoken myself and to gain early feedback from the respeakers about working at these live focus group sessions.

Cycle One was the longest of the three research cycles, and table 5.2 below shows its full timeline. The consultation with Stagertext began at the very start of my PhD and continued throughout its lifespan, though the amount of contact varied across the different phases of action and research. The focus group interviews took place within a more concentrated period, in the most part between February - November 2016. First came the DH focus groups, followed by interviews with venues and respeakers. The DH and respeaker surveys went out at the end of this period. The NNE focus groups came last and were staggered for logistical reasons. The first as-live respeaking opportunity coincided with the end of the focus group work and overlapped into Cycle Two, where the findings could be more directly applied to the research events.

5.2 What is the nature of the access currently provided at live events?

Whilst SDH on television, in particular pre-recorded subtitles, have long been considered a good benchmark for subtitling quality in the UK and further afield, theatre captions and live subtitles created through stenography have not had the same attention. My own experience as a subtitler was also limited to that of television and for this reason, it was essential that I began my research by exploring the access that was available at events and gained a better understanding of the current practice, especially as I knew that many participants would already be familiar with this setting.

5.2.1 Consultation with Stagertext and related organisations

As the leaders in providing captioning and subtitling at live events, Stagertext were a vital point of contact for this study and the advice they provided supported me throughout. They helped me realise gaps in my own knowledge and misconceptions I held, before they proved problematic or disruptive to the work being done.

At the start of the research, I met with Stagertext's General Manager (GM), Theatre Programme Manager (TPM) and Live Talks Manager (LTM) and also went to observe three of their events. This was an opportunity for me to explore Stagertext's work in more detail and to better understand the compatibility of respeaking with Stagertext's own ethos, procedures and practicalities at live events, and more broadly within this new setting.

It also helped me understand some of their concerns regarding the use of respoken subtitles. These related to the fact they are not verbatim, and that too much content might be lost if respeaking were used, something which would compromise their core principle of delivering full access through the captions and subtitles they offer. At the same time, concerns were raised over how long an audience member might want to follow a respoken event where there was necessarily heavy editing. They did, however, see occasions where using respeaking could be very appropriate. They wondered about its use for events attended by school groups, given the pace and editing they thought might be produced. Similarly, they wondered about events, such as concerts and pantomimes,

where unscripted content is interspersed into the main event. Here, the Stagertext software could allow respeaking in combination with cued material. The most challenging genre we discussed was comedy, since the latency might impact heavily on the experience for the audience member.

During our discussions, there was an understanding that while Stagertext were enthusiastic about exploring this new avenue for access, they were absolutely committed to upholding the principles embedded within their own work and there was no guarantee that they would adopt respeaking following this study. The tension surrounding the use of respeaking and stenography in the UK was also acknowledged. As the study unfolded, Stagertext took care to ensure that the stenographers who provide the live subtitles at their events were aware of my research and understood their position towards it, as well as my own: I ensured that Stagertext had informed the stenographers that my interest was in increasing the range of access provided and not in replacing or removing any form of existing access, stenography included.

During the course of the study, Stagertext advised me on the running of the focus groups, and commented on the DH questionnaire before it was released. As well as publicising my calls for volunteers and events, their representatives attended the DH focus groups and first round of events and they provided the tablets for the museum tours (Events 4 and 7). In line with the collaborative nature of action research, there were also opportunities within our interactions for me to feed information back to Stagertext, to promote their activities, and to introduce new venues and potential audience members to their work.

Whilst Stagertext were the main advisors for this research, I also worked closely with a number of other different organisations at different points in this study. The National Association of Deafened People (NADP) shared my calls for volunteers and the online questionnaire; their members attended the focus groups and research events and I was invited to share my findings in their newsletter and to present at the NADP Conference in June 2019. I visited the research department at Action on Hearing Loss (AOHL, now Royal National Institute for Deaf People, RNID) who also publicised the project. Through

this connection, I was recognised as a stakeholder by Ofcom and invited to attend their round table on live subtitling in June 2017, which followed on from their consultation into the quality of live subtitling referred to above.

5.2.2 Observations at live events

The three observations organised through Stagertext in November 2015 were opportunities for me to see captioning, STTR and respeaking, along with the presentation software I would be using, in action as an audience member; in addition, I had ‘behind the scenes’ access at the first and last event. The events I observed were a museum tour, *Dutch Genre Painting: Visual Proverbs*, at the Wallace Collection, London (www.wallacecollection.org/), a public talk at the Hunterian Museum, London (www.rcseng.ac.uk/museums-and-archives/hunterian-museum/), entitled *Wax Anatomies in the Medical Museum*, and a performance of *The Odyssey* at the Northcott Theatre, Exeter (www.exeternorthcott.co.uk/), which was followed by a Q&A with the actors and producers.

All three events demonstrated the preparation and co-ordination which was needed in advance of the event to ensure it ran smoothly. A great deal of information was shared before each event began including times, venue details, points of contact, information about the event content for the STTR and details of the technical set-up. The latter included specific details about where people would be positioned and how the audio feed would be established. For example, during the museum tour, the access was provided by a remote STTR, whose audio feed came via mobile phone and the live subtitling was streamed via Wi-Fi to individual tablets, whereas at the second event, the STTR was onsite positioned at the front of the lecture theatre. At the third event, the captioner would be positioned in the upper lighting box, with the captions displayed on LED screens, set up on tripods on either side of the stage, but positioned slightly further on stage than usual. The captioner would be using respeaking during the post-show discussion, and the chair was asked to repeat any audience questions for clarity to facilitate this process. Information was also provided on where the audience would be seated in order to view

these LED screens. Information about who would be attending was also shared. Different members of the Stagertext team attended each event, according to the expertise required and the nature of access being provided. A venue representative was also designated, and a BSL interpreter was present at the first event. The preparation start time varied for each event; for the museum tour it began around an hour before the event, while for the captioned performance three hours was required to set everything up.

The nature of this on-site preparation varied for each event. At the Wallace Collection, it was dedicated to checking the Wi-Fi signal along the tour route and preparing the tablets (checking their battery level and setting the optimum high contrast display and font size of 42). The introductions that each person would give were also checked. Fifteen minutes before the start of the event, Stagertext's LTM called the STTR to check the mobile signal, audio level and to ensure that the subtitles were streaming correctly; she then kept her informed of what was happening until the event began. During the public talk, the focus was on ensuring the lecture theatre was set up correctly and checking that the live subtitles did not obscure the PowerPoint presentation, as both were displayed on the same screen. A few slides were adjusted before the event began. The focus before the captioned theatre performance was on setting up the LED screens and testing the audio feed.

During each event, the Stagertext staff continued to play an important role and support the audience in making use of this access. The LTM monitored the live subtitles throughout the tour and talk and certain codes were in place, so that any issues could be quickly resolved. When the STTR was unable to capture what was said, they cued [INAUDIBLE], so that the presenter could be asked to repeat the missing content. At one point during the tour, [THE MICROPHONE IS COVERED] was cued, and the guide changed their grip and the position of the phone. During the Q&A at the end of the tour, [?] was cued when the STTR was unsure about the content, allowing it to be clarified. At the tour, the LTM had an 'Ask me about captioning' sign and audience members were instructed to wave if they had any trouble with the tablets. The venue staff were also involved in this process and had spare tablets to supply in case any needed resetting. At

the final event, captions were displayed before the performance began to explain that it would be captioned, and also at the end of the event to remind the audience that there would be a post-show discussion. The captioner knew the performance well and was very well prepared as she had captioned it at other locations around the country.

Care was taken to ensure each event finished on time, since the STTRs and captioners had been working alone and there was a brief review at the end of each event to address any problems that might have arisen, and to inform future events. For example, at the museum tour, the venue representative was interested in finding out whether the STTR would have liked more information about the content of the tour in advance, as it had varied a bit on the day; the LTM was interested in knowing how hearing people might have found words being spelled out on screen, which happened on a few occasions.

The third event allowed me to see respeaking in action at a live event and what this study could contribute to improving it as a live event service. The captioner had previously worked as a television subtitler.

Whilst the captions had provided very accurate access to the spoken content of the performance, it was much harder for the captioner as respeaker to capture the content of the post-show talk in the live subtitles. Q&A are naturally spontaneous in their content and for that reason challenging to respeak. However, other reasons seemed to contribute to this as well. The lighting box was quite far removed from the stage. This did mean that the audience were unable to hear the respeaking, however it also impacted on the quality of the audio the respeaker received, especially when the audience spoke. The chair did repeat some questions, but as no-one was formally monitoring the live subtitles, there was no one to remind them to do so, or to respond to cues from the respeaker so the cues in use by the STTR, referred to above, were not effective here. Together, these three factors meant that the subtitles at this event were not as good as the those at the previous two events I had attended and the respoken live subtitles were indeed more heavily edited.

These observations gave me a good understanding of the access provided by STTR and the challenges involved in the live event setting, and also of how respeaking could be

implemented there and, in particular, the questions that need to be addressed in doing so.

I was able to observe closely what conventions were already in use, ranging from punctuation (semi-colons, commas, speech marks in addition to sentence final punctuation) to the cued communication, conventions which respeaking would seek to replicate. In addition, I saw certain approaches that were adopted for more challenging content, which would also be useful for respeaking. For example, the museum tour included many foreign phrases. Handouts had been prepared and shared in advance containing some of the foreign phrases mentioned during the talk, and these were not transcribed.

The live subtitles created by STTR that I saw were very accurate in terms of the limited number of errors that appeared on screen and in the amount of content captured. Latency was also very low. Stagertext state that STTR is verbatim (Stagertext, 2021c) and it certainly seemed close to this. Given this standard, and Stagertext's own concerns about the possible editing used during respeaking, it became clear that this was something I would need to monitor during the study. Stagertext ask venues to provide preparation materials to ensure the process runs smoothly and to improve the speed and accuracy of the final subtitles that are produced (Stagertext, 2021a); the errors I observed here were certainly minor ones, such as 'double entendre' and 'stup endous', though on other occasions I have seen more complicated errors appearing. Nevertheless, it is clear that STTR sets a very high standard for access at live events, one which was likely to be reflected in the user and provider expectations for respeaking at live events.

5.3 What expectations do the key parties have for a respeaking service at live events?

Having observed access at live events for myself, I wanted to engage directly with potential stakeholders to understand what their perceptions, concerns and expectations were around the use of respeaking, and the provision of access more generally, at live

events. Whilst I knew that a variety of opinions were likely to be expressed by individuals within each group, I hoped to be able to identify commonalities between them that would help establish the core needs and wider expectations of both users and providers. In addition, I was interested in seeing whether there were any areas of overlap across the four groups.

5.3.1 Approach to data collection and analysis

The data from these key parties came in the form of interviews and questionnaires, enabling me to collect a mixture of qualitative and quantitative data, following a mixed methods approach.

Gough (2020) writes, “sometimes we are so hypnotized by [quantitative] data, we gaze past our own humanity. To get the whole picture, you need the story behind the data - the ‘so what?’, otherwise all you have is data”. In the analysis that follows, I have aimed at all times to capture the ‘so what’ that the participants expressed and to explore the qualitative comments that followed the closed quantitative questions. Whilst the quantitative data provided useful snapshots of the trends, it was the qualitative data that was often needed to convey the complexity of the opinions and experiences that were related by the audience and respeakers.

When reviewing this qualitative data, I used a light version of the reflexive thematic analysis approach first proposed by Braun and Clarke in 2006 and further refined in 2019. In this type of analysis, the researcher is encouraged to seek out the underlying themes in the data, themes which are present within participants’ responses yet go beyond the questions posed. I followed an inductive approach when doing this; the themes I discuss grew from what was said by participants, rather than by an attempt to impose a list of themes drawn from the literature. I was nonetheless aware that I would be bringing my own perspective and view as I carried out the analysis.

Thematic analysis provided a flexibility in approaching the data which was well-suited to this dataset. Considering the fact that multiple strands of data were collected across different-sized focus groups, the fact that *occurrence* of a theme or topic, is more

significant than *prevalence* was particularly relevant (2006: 82); in other words, the very fact that a topic appears gave it value, regardless of either the brevity of its appearance, or how many times it appeared. As a result, small samples, such as those seen in the NNE group, were not problematic and neither were samples where multiple responses were logged by individuals, for example in the audience and respeaker groups, where some people had completed interviews and questionnaires.

On the other hand, trying to perform a combined quantitative analysis on these varied datasets is more complex. The way I have decided to approach this is as follows: when I write about DH audience data, after presenting the participants (section 5.3.3.2), I then rely on quantitative data from the questionnaire alone, since its purpose was to identify broader quantitative trends, but draw on qualitative data from both the focus groups and questionnaires as appropriate in the discussion. In the case of the other groups, in particular the respeakers, the quantitative data is not as precise since, in many cases, different numbers of people responded to each question. Here, I do give percentages when I can, but often rely on the general trends in the data to illustrate the points made. The appendices to this chapter contain the questionnaire responses and an overview of what was discussed within the interviews. The questionnaires were created and analysed in Qualtrics and I used MAXQDA alongside a paper analysis when reviewing the qualitative data.

5.3.2 Data from the audience

I spoke to two audience focus groups - people who were d/Deaf, deafened and hard of hearing (DH), and who for the most part were native English speakers or British Sign Language users, and people who were non-native speakers of English (NNE) and for the most part hearing. I considered the first group as 'expert' users, since they use subtitles regularly and generally see them as an essential part of their viewing practice; the focus group contact with this group was the most extensive. The NNE focus group can be seen as a counterpoint to this first audience group and an example of how wider audience groups might use respeaking at live events.

5.3.3 Views from d/Deaf, deafened and hard of hearing (DH) people

The DH focus group data comes from thirteen people who, in February and March 2016, attended a focus group session in-person at the University of Roehampton and completed a short take-home questionnaire (appendices 5.3.1 - 5.3.2.2) and from the 53 complete responses to the longer online questionnaire (appendices 5.4.1 – 5.4.2.2), which was live in August and September 2016.

5.3.3.1 Development of the focus group sessions and online questionnaire

Two separate focus group sessions were held, each lasting around two and half hours with a break for refreshments. They took the form of a semi-structured group discussion: the order in which the content was covered was directed by what individual participants shared. A BSL interpreter and two respeakers provided the access (Fig. 5.1) and participants were asked to raise their hands before speaking to facilitate communication.



Fig. 5.1: Respeakers and BSL interpreter at the second DH focus group session

I briefly outlined the purpose of the research before the discussion began. The presence of the respeakers meant we were beginning the research at the place where we had expected to end it, by using respeaking – albeit in trial form – to make the sessions accessible; in addition, the respeakers were using new equipment at each session (see section 5.4 below). For that reason, I also commented on this and the action research approach.

The online questionnaire was expected to take 45 minutes to complete and could be done across more than one sitting. It also contained a brief introduction to the study and can be found in appendix 5.4.1.

The content of the focus groups and online questionnaire were closely linked, both in terms of the topics covered and because the feedback gained in the focus group sessions shaped the questions that were included in the questionnaire. Table 5.3 illustrates the topics covered.

Table 5.3: Topics covered in the DH focus groups and online questionnaire

Section	Topic	Content
A	Demographic Information	<ul style="list-style-type: none"> • Age • Gender • Hearing status • Education • Communication modes
B	Experience and use of subtitles/captions across four key settings	<ul style="list-style-type: none"> • Pre-recorded television • Live television • Theatre • Live arts and cultural events
C	Accessing events	<ul style="list-style-type: none"> • Experience at live events to date • How should captioned/subtitled events be chosen?
D	Subtitling/captioning conventions	<ul style="list-style-type: none"> • Legibility of subtitles and captions • Character identification • Position • Sound, music, manner • Visual presentation
E	Subtitling/captioning speed	<ul style="list-style-type: none"> • Pre-recorded television • Live television • Theatre • Live arts and cultural events
F	Live subtitling style	<ul style="list-style-type: none"> • Verbatim • Standard • Edited
G	Subtitling/captioning priorities	<ul style="list-style-type: none"> • Live television • Theatre • Live arts and cultural events
H	Opportunities for further comments	Participant-led

Much of this content, and especially sections A, B, D and E, was based on the long questionnaire used in the DTV4ALL study which explored the reception of SDH on television in seven European countries (Romero-Fresco, 2015) in preparation for the

advent of digital TV⁶⁶. I adapted the questions on these topics so that I could explore participants' responses to the conventions currently in use in television subtitles and theatre captions and determine what they would like to see at live events, whereas the focus of the DVT4ALL UK questionnaire had been television and DVD subtitling. In sections F and G, I explored the issue of live subtitling speed more closely and then asked participants what their priorities were for the content of subtitles and captions in each setting. Finally, section C explored participants' experience of access more broadly. Many of the responses seen in the online questionnaire reflected the answers given in the DVT4ALL study, which was completed by 434 Deaf and hard of hearing people in the UK, which suggests the wider applicability of the data and discussion that follows.

5.3.3.2 Who were the participants?⁶⁷

Out of the thirteen people who attended the focus group sessions, six were male and seven female. Most participants were aged above 36, but one selected the 26-35 category. Eight had completed studies at (technical) college or university and four at postgraduate level. Nine had attended mainstream schools, two Deaf schools and one had attended a school for disabled children. Nine were currently employed and five were retired or semi-retired.

When asked, "Are you...?", two selected 'Deaf (BSL)', four 'deaf', one 'deafened', five 'hard of hearing' and three selected 'other', adding that they were profoundly deaf, completely deaf without their CI and, finally, bilateral severe to profound deaf. The most common age of onset was from birth (n5) and then 5-19 years (n3). The other ages were evenly split between 2-4 years, 30-49 years and 50-64 years of age. One person used a

⁶⁶ Funded by the European Commission, this project ran from 2010 – 2013 with the aim of facilitating the availability of access services on digital TV and making recommendations to the relevant stakeholders on how these services could be improved (p.9). At the start of the project the hope was that it might lead to "a unique standard to help harmonise SDH practices across Europe" (p.10), but the reality was that the landscapes of each country differed too greatly, in economic, social, political and audiovisual terms, to name but a few. This meant that standardisation across Europe was unlikely to be reached.

⁶⁷ The take home questionnaire from the focus groups was completed on paper. On occasion, participants chose to give more than one answer to a question, or skipped a question entirely, so there is some variation in the total number of responses for each question. Twelve out of the thirteen participants completed the questionnaire.

cochlear implant, and eleven used a hearing aid. When asked how they communicated with hearing people, lipreading (n4) and hearing aids (n2) were the most common responses; BSL interpreters (n1) and Sign Supported English (SSE) (n1) were also used. Out of the people who completed the online questionnaire, eleven were male, 40 female, one queer and one person no gender. They were aged between 30 and 88 and all lived or worked in various locations around the UK.

Nineteen people selected 'deaf', sixteen 'hard of hearing', twelve 'deafened' and two 'Deaf (BSL)'. No-one selected 'hearing' or 'hearing, but I don't hear as well as I used to'. Three people selected 'other' and I assigned this answer to one person who had not replied, so that I could include their data and use the cross-tabulation feature in Qualtrics to sort the responses to later questions according to the hearing status selected. Out of those who selected 'other' one was "totally deaf without their cochlear implant, and hard of hearing with it on" and had become hard of hearing aged five and received their CI aged fifty-nine. The next person had acquired profound hearing loss, which began gradually as an adult and the third person had situational deafness, which began as a child. Whilst the exact degree or description of hearing loss for the final person is not clear, it did begin before they were 2 years old.

It must also be pointed out that even where participants had selected a hearing status, variation in the lived experiences behind the answer was vast, especially where people were deaf, deafened and hard of hearing. For example, among the sixteen people who said they were hard of hearing, nine became hard of hearing at birth, the others later in life. Most (n12), but not all used lipreading alongside hearing aids to communicate with hearing people and one used a BSL interpreter. Out of the nineteen people who said they were deaf, eight had been since birth, and a further seven by the age of eleven. Once again, the way they communicated with hearing people varied, and included BSL interpreters (n4), SSE (n7), lipreading (n17), hearing aids (n15) and cochlear implants (n4). In the case of the twelve people who said they were deafened, this had for the most part happened after the age of twelve (n11), with the most common age bracket being between 22-51 (n5). Some had hearing aids (n8) or CIs (n3) and almost all of them used

lipreading (n11). The two people who were Deaf BSL users shared a similar profile; both were Deaf from birth, had hearing aids and had attended Deaf schools; however, one had also attended a mainstream school and their age and occupations varied.

5.3.3.3 Viewing habits

Most people who responded to the online questionnaire watched television for 1-2 hours (n24, 45.3%) or 3-4 hours (n17, 32.1%) each day and they viewed a wide range of programmes, including news and documentaries, sports, series, films and comedies. Almost all the respondents (n48, 90.6%) said that they always used subtitles when watching television and 71.1% (n37) said that they chose the programmes they watched according to whether they were subtitled. If they discovered that a programme was not subtitled, most (n43, 46.7%) would switch to another channel to find a subtitled programme there. A few would continue to watch and either guess what was said (n3, 3.6%), lipread (n6, 6.5%), turn up the volume (n4, 4.4%) or ask someone else to tell them what was said (n7, 7.6%). It seemed that continuing without subtitles was more likely to happen if a participant was watching with family (participants FGB5 and FGB6), if they were watching sports (OL7 and OL50) or on programmes where there was a sign language interpreter (SLI) (OL20). One person said they would prefer to try to watch the programme later, on demand, rather than to persevere without subtitles (OL16). Whilst everyone had used subtitles to access television programmes, fewer people had used subtitles and captions to access other types of content, as displayed in Fig. 5.2.

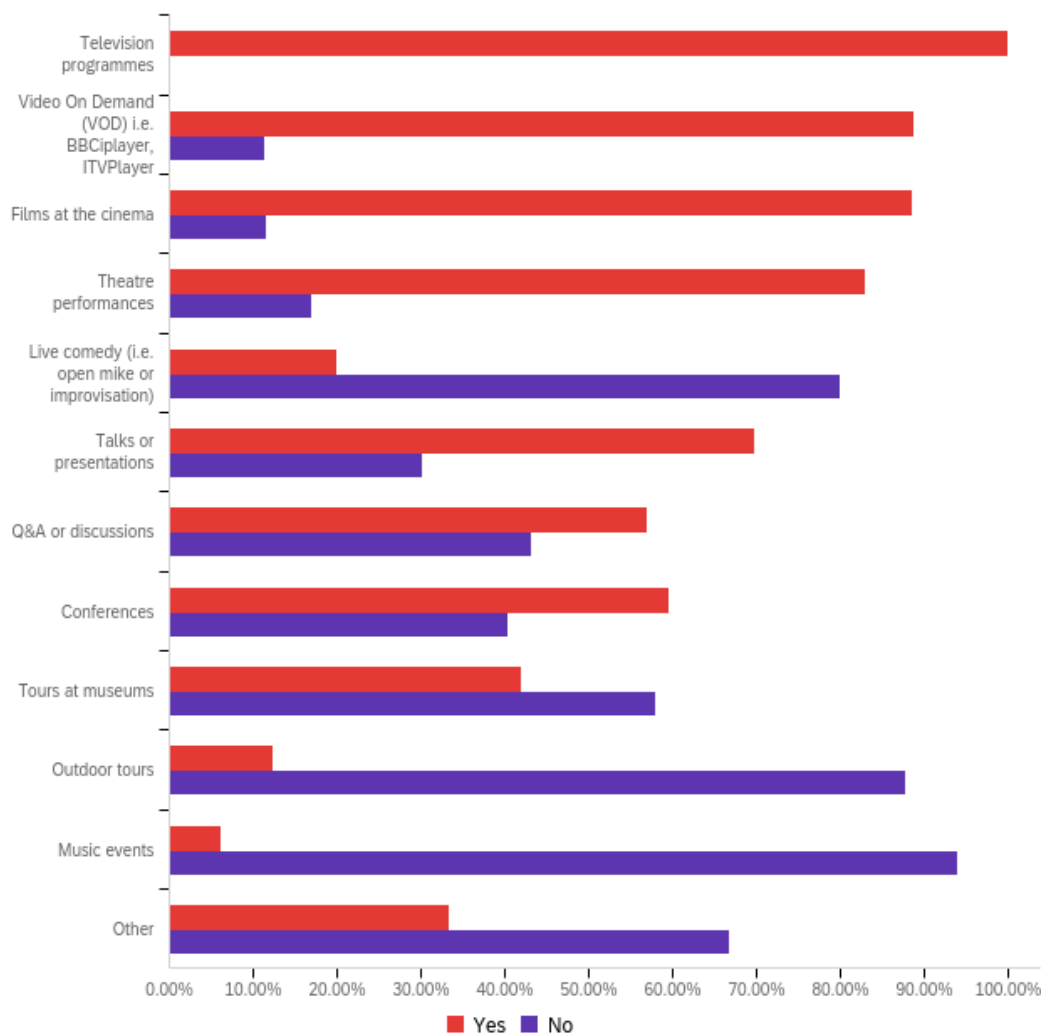


Fig. 5.2: B1: Have you ever used subtitles/captions to access the following?

Subsequent questions revealed that while 53.9% (n28) went to captioned performances 1-4 times a year and 21.1% (n11) went more frequently, 25.0% (n13) never went. When asked about performances without captions, most (n33, 62.3%) said they never went, and the number who went 1-4 times a year dropped to 26.4% (n14). Fewer people went to live events than to the theatre and a similar pattern was seen; subtitled events were more frequently attended than events without subtitles. 40.4% (n21) attended 1-4 subtitled events each year, dropping to 19.2% (n10) when there were no subtitles. While 38.5% (n20) had never attended a subtitled event, 59.6% (n31) had never attended an event without subtitles.

Although many possible reasons may lie behind this range of attendance, it is clear that a lack of access provision, in all settings, is likely to lead to either reduced or no attendance. Participants were asked about different situations in which they may have difficulty in using English and accessing content without subtitles; table 5.4 summarises the main results.

Table 5.4: Difficulties in using English and accessing content without subtitles/captions

	Yes	Sometimes	No
Do you have difficulties understanding written English?	0.0% n0	3.8% n2	94.3% n50
Do you have difficulties understanding spoken English?	26.4% n14	43.4% n23	18.9% n10
Do you have difficulties reading or understanding subtitles?	0.0% n0	28.3% n15	71.7% n38
Do you have difficulty watching TV without subtitles?	92.5% n49	5.7% n3	1.9% n1
Do you/would you expect to have difficulty accessing the theatre without captions?	92.5% n49	5.7% n3	1.9% n1
Do you/would you expect to have difficulty accessing other live arts and cultural events without subtitles?	92.5% n49	5.7% n3	1.9% n1

Whilst there was some variation in the difficulty experienced when understanding spoken English and when reading or understanding subtitles, what was striking was the similarity in responses to the questions on how difficult content would be to access without subtitles or captions. 60.6% (n43) of respondents stated that subtitles and captions were the only way they had to access the dialogue and 28.2% (n20) said they helped them to understand. An absence of subtitles and captions at events might be particularly problematic because speakers may be too far away to lipread effectively, they may move around the stage and the sound bounces around as well, making catching what was said more difficult (OL26). The impact of this lack of access is considered further in section 5.3.3.7.


5.3.3.4 Conventions

Much variation is seen within the conventions used in the UK, not only between the different settings subtitles and captions appear in (television, online players, cinemas, events) but even within individual settings; television provides a perfect example of this, with variations in style adopted across the different channels. For this reason, section D of the online questionnaire was particularly important as it sought to explore audience preferences for a range of different conventions. Although my eventual focus would be on provision in the live event setting, my aim was to use the well-established medium of pre-recorded subtitles, which most people were likely to be familiar with, to determine preferences and then to incorporate these into the guidelines for live event subtitling. Appendices 5.1.1-5.1.12 contain the resources I used to illustrate the different conventions at the focus group sessions and in the questionnaire.

Table 5.5 summarises the preferred conventions for each feature, which I will now discuss more closely.

Questions D1-6 related to font and typeface. 90.6% (n48) of participants stated that the fonts used on television were easy to read, and 75.0% (n39) said that the font used in captions was easy to read. At many live events, the presentation software may only allow a single colour of text to appear throughout. Since screens may sometimes be located at a distance, it is particularly important that a good colour contrast is chosen. Even on individual tablets, where people can often adjust the visual appearance of the subtitles, it is good practice to initially present text which has the optimum size and colour contrast. For these reasons, I wanted to explore participant's preferences for the colour combination of the main speaker, rather than discussing the range of font colours available. There was a clear preference to use a black background rather than an outline to enhance legibility. On television, there was a slight preference for white letters over yellow (42.5%, n22), but at live events, no preference was registered.

Table 5.5: Audience preferences on subtitling/captioning conventions

Feature / function	Scope	Combined responses	Deaf (BSL)	deaf	Deafened	Hard of hearing	Other
Easiest font to read 	All settings combined	A (Arial)	A (Arial)	A (Arial)	A (Arial)	A (Arial)	A (Arial)
	Television	White letters on a black background 42.3%, n22	White on black / No preference Both 50.0%, n1	White on black 47.4%, n9	White on black 50.0%, n6	White on black 37.5%, n6	No preference 75.0%, n3
	Theatre	No preference 35.9%, n19	White on black / No preference Both 50.0%, n1	Yellow on black 42.1%, n8	White on black 41.7%, n5	Yellow on black 43.8%, n7	No preference 50.0%, n2
	Tablet	No preference = White letters on a black background 35.3%, n18	White on black / No preference Both 50.0%, n1	White on black 36.8%, n7	White on black / no preference Both 41.7%, n5	White on black 31.2%, n5	No preference 75.0%, n3
	Screen (i.e. at a presentation)	No preference 40.4%, n21	White on black / No preference Both 50.0%, n1	White on black 36.8%, n7	White on black 50.0%, n6	No preference 50.0%, n8	No preference 50.0%, n2
Colour combination for main speaker <ul style="list-style-type: none"> • White letters on a black background • Yellow letters on a black background • Outlined white letters (no background block) • Outlined yellow letters (no background block) • No preference 							

		Combined responses	Deaf (BSL)	deaf	Deafened	Hard of hearing	Other
Character/speaker identification <ul style="list-style-type: none"> Colours Positioning subtitles next to or under characters (TV only) Name tags Combining colours and positions (TV only) Combining colours and name tags 	Television	Colours 66.0%, n31	Colours / Name tags Both 50.0%, n1	Colours 47.4%, n9	Colours 58.3%, n7	Colours 62.5%, n10	Colours 100.0%, n4
	Theatre	Name tags 52.3%, n23	Colours / Name tags Both 50.0%, n1	Name tags 47.4%, n9	Colours 26.5%, n5	Colours 58.3%, n7	Name tags 100.0%, n4
	Live events	Name tags 50.0%, n23	Colours / Name tags Both 50.0%, n1	Name tags 47.4%, n9	Colours 58.3%, n7	Name tags 58.3%, n7	Name tags 75.0%, n4
Position of main subtitle/caption <i>(options varied according to the setting)</i>	Television	Bottom of the screen only 80.8%, n42	Bottom of screen only 100.0%, n2	Bottom of screen only 72.2%, n13	Bottom of screen only 66.7%, n8	Bottom of screen only 93.8%, n15	Bottom of screen only 100.0%, n4
	Theatre	On both sides of the stage 20.0%, n10	Front of stage 100%, n1	On both sides / on smart glasses / other 16.7%, n3	On both sides 33.3%, n4	At the front / on both sides 20.0%, n3	Other 50.0%, n2
Position of sound related information <ul style="list-style-type: none"> Top-right side of screen Bottom of screen next to subtitles Next to the source of the sound 	Television	Bottom of the screen next to the subtitles 76.0%, n38	Bottom of the screen next to the subtitles 100.0%, n2	Bottom of the screen next to the subtitles 66.7%, n12	Bottom of the screen next to the subtitles 90.9%, n10	Bottom of the screen next to the subtitles 86.7%, n13	Top-right side of screen 50.0%, n2

		Deaf (BSL)	deaf	Deafened	Hard of hearing	Other
Descriptions of sounds <ul style="list-style-type: none"> Where the sound comes from (ALARM CLOCK RINGS) Using words to reproduce the sound (CRASH!) Describing the sound (HIGH-PITCHED WAIL) 	All settings combined	Where the sound comes from (ALARM CLOCK RINGS) 100.0%, n2	Where the sound comes from (ALARM CLOCK RINGS) 47.4%, n9	Where the sound comes from (ALARM CLOCK RINGS) 66.7%, n8	Where the sound comes from (ALARM CLOCK RINGS) 43.8%, n7	Other 50.0%, n2
	Combined responses	Where the sound comes from (ALARM CLOCK RINGS) 50.9%, n27	Where the sound comes from (ALARM CLOCK RINGS) 47.4%, n9	Where the sound comes from (ALARM CLOCK RINGS) 66.7%, n8	Where the sound comes from (ALARM CLOCK RINGS) 43.8%, n7	Other 50.0%, n2
Background music <ul style="list-style-type: none"> Song title Type of music ♯ to indicate 'music' # to indicate 'music' Nothing ♯ alongside the lyrics # alongside the lyrics 	All settings combined	♯ alongside the lyrics 100.0%, n2	♯ alongside the lyrics 52.6%, n10	♯ alongside the lyrics 50.0%, n6	♯ alongside the lyrics 56.3%, n9	♯ alongside the lyrics 100.0%, n4
	Combined responses	♯ alongside the lyrics 58.5%, n31	♯ alongside the lyrics 52.6%, n10	♯ alongside the lyrics 50.0%, n6	♯ alongside the lyrics 56.3%, n9	♯ alongside the lyrics 100.0%, n4
Information about the music <ul style="list-style-type: none"> Instrument or style (BAND MUSIC) When the music starts or stops (MUSIC FADES) Description of the music (LIVELY MUSIC) Combination (EERIE ORGAN MUSIC / BAND MUSIC FADES) Other 	All settings combined	Combination (EERIE ORGAN MUSIC / BAND MUSIC FADES) 100.0%, n2	Combination (EERIE ORGAN MUSIC / BAND MUSIC FADES) 52.6%, n10	Instrument or style (BAND MUSIC) 33.3%, n4	Combination (EERIE ORGAN MUSIC / BAND MUSIC FADES) 43.7%, n7	Combination (EERIE ORGAN MUSIC / BAND MUSIC FADES) 75.0%, n3
	Combined responses	Combination (EERIE ORGAN MUSIC / BAND MUSIC FADES) 47.2%, n25	Combination (EERIE ORGAN MUSIC / BAND MUSIC FADES) 52.6%, n10	Instrument or style (BAND MUSIC) 33.3%, n4	Combination (EERIE ORGAN MUSIC / BAND MUSIC FADES) 43.7%, n7	Combination (EERIE ORGAN MUSIC / BAND MUSIC FADES) 75.0%, n3

	Combined responses	Deaf (BSL)	deaf	Deafened	Hard of hearing	Other	
Manner (especially when not obvious from the visual image) <ul style="list-style-type: none"> • Symbols: 0, 1, ?, !?, SHOUTING • (AGGRESSIVELY) • AGGRESSIVELY: • Nothing 		Symbols: 0, 1, ?, !?, SHOUTING / (AGGRESSIVELY) Both 50.0%, n1	(AGGRESSIVELY) 52.6%, n10	Symbols: 0, 1, ?, !?, SHOUTING 41.7%, n5	Symbols: 0, 1, ?, !?, SHOUTING 43.8%, n7	Symbols: 0, 1, ?, !?, SHOUTING 50/0%, n2	
	Number of lines (options varied according to the setting)	Television 2 lines 67.3%, n35	2 lines 100.0%, n2	2 lines 68.4%, n13	2 lines 66.7%, n8	2 lines 62.5%, n10	2 lines / 3 lines Both 50.0%, n2
Presentation of sound labels <ul style="list-style-type: none"> • HE LAUGHS • he laughs • (HE LAUGHS) • (he laughs) • [HE LAUGHS] • [he laughs] • No preference • Other 	Theatre 2 lines or 3 lines Both 42.0%, n21	2 lines 100.0%, n2	3 lines 42.1%, n8	2 lines 41.7%, n5	2 lines / 3 lines 43.8%, n7	3 lines / 4+ lines Both 50.0%, n2	
	Live events 2 lines 39.6%, n19	2 lines 100.0%, n2	2 lines or 4 lines Both 31.6%, n6	2 lines 33.3%, n4	2 lines 58.3%, n7	2 lines 58.3%, n7	3 lines / 4+ lines Both 50.0%, n2
	All settings combined (he laughs) / no preference Both 23.1%, n12	(HE LAUGHS) [he laughs] Both 50.0%, n1	No preference 31.6%, n6	(HE LAUGHS) (he laughs) No preference All 25.0%, n3	(he laughs) [HE LAUGHS] Both 31.3%, n5	(he laughs) [HE LAUGHS] Both 31.3%, n5	[he laughs] 66.7%, n2

	Combined responses	Deaf (BSL)	deaf	Deafened	Hard of hearing	Other
Display mode <ul style="list-style-type: none"> • Word-by-word • Block 	Live events	Word-by-word/ Block Both 50.0%, n1	Word-by-word 68.4%, n13	Word-by-word 58.3%, n7	Word-by-word 56.3%, n9	Word-by-word 75.0%, n9

A few participants highlighted factors other than the font and colour contrast which might affect legibility at events. These included where the caption display was positioned in relation to the stage and wheelchair seating, and whether the lights shone directly onto the screen.

Questions D7-D10 related to character and speaker identification. Dialogue is an essential component of subtitles (Romero-Fresco, 2015: 157), and for good comprehension and access, it is essential that the words spoken can be attributed to the correct speaker. This may be all the more relevant at live events, where people attend in person and may have the opportunity to interact directly with those present (see Section 8.6.1 below for further discussion of this point).

In the online survey, this question was asked for each of the three settings. Whilst colours were the preferred choice for television (66.0%, n31), name tags were preferred at the theatre (52.3%, n23) and live events (50.0%, n23), though some variation was seen among the different hearing groups. These answers reflect the current practice seen in each setting.

The additional comments provided added insight into speaker identification. At events, the fact that it might be harder to see who is speaking than on television, and the distance between the subtitle display and speakers meant that name tags were especially important for some participants. Others pointed out that since colour allocation is not always consistent throughout, they weren't sure how that would work at the theatre or at live events. The feedback shared on positioning, suggested that it was a less popular choice because the moving subtitles were more likely to block critical content. However, one participant wondered whether this could work at the theatre, where there was a possibility of more "empty spaces" (OL15); they felt this was something worth exploring. A final comment was that "the captioner should use common sense in determining the best way to get the message across" (OL22).

Questions D11-D14 explored preferences regarding where subtitles, captions and sound labels should be positioned. The bottom of the screen was the preferred position for sound labels (76.0%, n38) and subtitles (80.8%) on television, with the understanding

that this would not preclude the subtitles from moving temporarily when necessary to avoid blocking any on-screen captions or strap lines. More variety was seen in the responses to where captions should be positioned; both sides of the stage was the most popular choice overall (20.0%, n10), but the qualitative comments revealed that, rather than selecting a permanent position, the most important factors when determining where they should be placed should be that action should not be obscured and the need for head movement should be limited. In addition, communication with the theatre about where the unit(s) would be positioned was essential so that seats could be allocated appropriately. Where visuals such as slides would be used, some thought that positioning the subtitles next to them would be the best choice. Many participants mentioned that they hadn't used smart glasses, but were interested in trying them. A few participants were concerned that, at events, handheld devices or fixed screens would annoy strangers sitting near or behind them.

Questions D15-D19 related to how descriptions of sounds should be reflected in subtitles and captions. Indicating where the sound comes from, for example ALARM CLOCK RINGS, was by far the preferred option (50.9%, n27) to either onomatopoeic labels, i.e. CRASH! or a description of the sound. Some participants noted that they could hear or surmise sounds, and so did not need a description of them, whilst others said that how they were described would depend on what programme they were watching.

When it came to background music, there was a very clear preference (58.5%, n31) for lyrics to be included in the subtitles and captions, and more importantly for them to be demarcated with a ♪ symbol. The preference for this symbol, rather than the # is of interest because it is the latter that is used on television. It suggests that a possible change in usage of television conventions may be worth considering. Following this choice, the next most popular option (15.1%, n8) was to include the title of the song, which suggests that this would also be a good practice to follow. When creating sound labels for music, participants wanted as much information and detail as possible to be included within it.

The preferences for indicating manner were less defined. The use of symbols and punctuation was the most popular option (37.7%, n20), but some participants also selected an explanation within brackets, (AGGRESSIVELY). A fair number of people (22.6%, n12) also selected 'nothing', so when faced with a choice over what to prioritise, it seems that indicating manner is less of a priority.

In general, the comments relating to sound, music and manner highlighted that while for some participants, including these features is very important, for others, "too much description of things that can be surmised from the screen acting appears condescending and can be irritating" (OL45). In practice, it is important that in every situation, the subtitler reflects first on what is needed, and then the best way to achieve it. Sometimes, combinations of the options I offered as answers may also be appropriate. Some participants commented directly on the inclusion of lyrics, saying that not being able to follow the lyrics made it feel like they were missing out, especially since they often "matched the scene/atmosphere/expressed the character's thoughts and feelings" (OL42) and others said being able to identify the song in question was important as they might want to buy it later. Participants also stressed the importance of these captions being in sync. Although music, manner and sound are especially important in pre-recorded subtitles, the responses suggest that even in live scenarios, they should be included where possible. In section 5.3.3.6 below, participants were asked to rank a range of content to help the respeakers determine what they should prioritise as they respeak.

Questions D20-D22 related to the presentation of subtitles. Participants tended to prefer fewer lines of subtitles being displayed, with 2 lines being the preference on television (67.3%, n35) and at live events (39.6%, n19) and either 2 or 3 lines at the theatre (42.0%, n21). The comments reveal that, once again, central to this choice was the question of how many lines could be included without the action being blocked.

One respondent stated:

On television, I would like minimal space taken up by the subtitles. Likewise because captions are further away from the action taking place, I would like more time to watch the action then turn back to the captions to catch up what's being said (OL15).

Another said that on a handheld unit, 4 subtitle lines or more would be fine, but in an open captioned situation, it would become “distracting and laborious to have masses of text on view” (OL10).

Participants were also asked about their preferred style for sound labels. There was a marginal preference for the use of lower-case letters to upper-case letters, though the most popular option (23.1%, n12) was a tie between (he laughs) and no preference. This is interesting as on both television and at the theatre, the use of upper-case letters is the norm, with square brackets being used by Stagertext. The ‘no preference’ vote suggests that users are comfortable with all the suggested variations.

One final question, from section F of the questionnaire, also related to how subtitles at live events would be displayed, in either word-by-word or block form. Word-by-word was by far the preferred choice (64.7%, n33). The comments revealed that the word-by-word display mimicked the way that people lipread and listen: “my eyes are listening word-by-word just as hearing person listens” (OL2). Others felt that word-by-word subtitles were more likely to be in sync with what was said, and might also lead to less of the screen being obscured. Nevertheless, some participants did prefer blocked subtitles, even if this meant a slight delay as the words collected to form a block. They felt that blocks would be easier to read and would permit people to look around more and move between the screen and the action .

5.3.3.5 Speed and style

Given the importance of speed in the debate about subtitling, discussed in Chapter Three, I also wanted to gauge the opinions of participants about subtitling speed and the related issue of editing. Participants were first asked what they thought of current speeds in each setting.

Table 5.6: Audience opinions on the speed of subtitles/captions

What do you think of the usual speed of...?	Too fast	About right	Too slow
Pre-recorded subtitles on television	0.0% n0	96.2% n51	3.8% n2
Live subtitles on television	5.7% n3	54.7% n29	39.6% n21
Captions at theatre	6.3% n3	91.7% n44	2.1% n1
Subtitles at live arts and cultural events	3.3% n1	90.0% n27	6.7% n2

As table 5.6 illustrates, in most settings participants found that the speed of subtitles and captions was about right. It was only for live subtitles on television that opinions were more divided, and this was the setting that I explored in the most depth, given the context of the study.

I began by explaining the three broad subtitling styles, verbatim, standard and edited and asking which option participants preferred for live subtitles at events. The clips included in the questionnaire can be found in appendices 5.4.11-5.4.1.3 and an extract from each, along with the definition are in figure 5.3⁶⁸.

As well as providing these sample clips of each style, I also explained that wherever possible, respeakers would naturally “try and repeat the original as closely as possible so that the subtitles accurately reflect what has been said... However, there are times when a respeaker must edit what has been said” (appendix 5.4.1:35). I then asked the participants their preferred style and the reasons for their choice, before explaining the challenges that each method might pose to a respeaker (table 5.7).

⁶⁸ The subtitling in the clips was pre-recorded in format due to limitations in the software available to me. I explained that at the live event, the subtitles would appear in a scrolling format. I created the script through respeaking so that any editions would resemble those a respeaker might make. The average speed of the verbatim clip was 172wpm, the standard 155wpm and the edited 138wpm. In comparison, the speeds of the clips in the focus groups were 240wpm, 200wpm and 180wpm respectively. Blank spaces were not included in the display rate.

Verbatim subtitles:
Every word spoken is included – the subtitles are a literal and faithful transcription of what has been said.



Standard subtitles:
This is a middle way between verbatim and edited – the subtitles retain much of the language of the original, but some words have been paraphrased or omitted.



Edited subtitles:
The original words are condensed and simplified – the subtitles facilitate reading and understanding.



Fig. 5.3: Explaining verbatim, standard and edited subtitles, with extracts from O'Neill, 2016

Table 5.7: How verbatim, standard and edited subtitles are created

<p>Verbatim subtitles: These subtitles will take the longest to read, leaving you less time to watch the event. The respeaker will speak quickly, which may cause more errors. Correcting any errors causes delay. Some content may be missed if the respeaker is unable to keep up with the speaker. Particular words may lead to unusual errors, if they are not recognised by the speech recognition software.</p>
<p>Standard subtitles: The time needed to read these subtitles will be between that needed for the verbatim and edited subtitles. The respeaker will automatically make slight edits in the subtitles when the content is dense or words are difficult. Some errors seen in verbatim subtitles will be avoided, but others will not. Correcting any errors causes delay. The respeaker may still miss some content.</p>
<p>Edited subtitles: These subtitles will take the least time to read, leaving you more time to watch the event. The respeaker will often paraphrase the original and will try to choose words that are more likely to be recognised successfully. There will still be some errors. Correcting any errors causes delay. The respeaker will have less words to say, but may need time to think about the editing.</p>

A notable shift in preferences was seen as participants gained a better understanding of the challenges that each style might present for a respeaker, and consequently, for the participants themselves (table 5.8).

Table 5.8: Preferences for verbatim, standard and edited subtitles

	Which of these options do you prefer for live subtitles at events? (Original response)	Knowing the problems associated with each style, has your opinion changed at all? Please indicate your preferred <u>style</u> ? (Revised response)
Verbatim	47.2% n25	26.4% n14
Standard	35.9% n19	49.1% n26
Edited	17.0% n9	24.5% n13

Whereas verbatim had previously been the style of choice (47.2%, n25), standard replaced it (49.1%, n26), and the popularity of edited subtitles also increased, in particular among Deaf BSL, deaf and hard of hearing audience members (table 5.9).

Table 5.9: Preferences for verbatim, standard and edited subtitles by group

	Deaf BSL		deaf		Deafened		Hard of hearing		Other	
	Orig.	Rev.	Orig.	Rev.	Orig.	Rev.	Orig.	Rev.	Orig.	Rev.
Verbatim	100.0%	50.0%	42.1%	21.1%	58.3%	33.3%	50.0%	31.3%	0.0%	0.0%
	n2	n1	n8	n4	n7	n4	n8	n5	n0	n0
Standard	0.0%	0.0%	42.1%	52.6%	33.3%	58.3%	31.3%	43.8%	50.0%	50.0%
	n0	n0	n8	n10	n4	n7	n5	n7	n2	n2
Edited	0.0%	50.0%	15.8%	26.3%	8.3%	8.3%	18.8%	25.0%	50.0%	50.0%
	n0	n1	n3	n5	n1	n1	n3	n4	n2	n2

A variety of reasons were given for preferring verbatim in the original question. Both Deaf BSL users referred to wanting to know “exactly what was said, with nothing missed out” (OL24) and to being on “equal footing to hearing people” (OL30). Similarly one deafened person said that they wanted to know what was actually said, not someone else’s version of it. Some spoke of verbatim being more natural or mentioned that it was easy to follow if the subtitles matched exactly what was being said. As one person put it, “I can still hear the words... Just need help deciphering them, if you like. It gets confusing when they don’t match” (OL14). One person seemed to capture the crux of the verbatim-edited debate in their response, by acknowledging that it was good for things to be made easier, but not changed (OL1).

Those opting for standard tended to note that it was easier to follow without the added ‘ums’ and ‘ahs’. Many appreciated this degree of editing, but would be sceptical of much more:

I wouldn’t be able to trust that edited subtitles are not dumbing down the information/language. It also makes me feel rather patronised. I am also wary it would slow down the subtitling speed and would match a timing of how long someone has spoken for, which would add to the mistrust of being ‘cheated out’ of some information (OL15).

In general, there was a feeling that standard subtitles seemed to strike a balance where the flavour of the exchanges between the two speakers was maintained, whilst allowing comfortable reading with time to look at the picture (OL26).

Participants who did prefer fully edited subtitles noted that full understanding and the essential flavour of the conversation had been maintained. In particular, one person commented that the edited style allowed full understanding with minimal effort for the person reading, whilst the standard subtitles could sometimes be long-winded, despite their lighter content in comparison to verbatim subtitles.

When asked to explain the reason for any change in preference, after the possible challenges for each style were explained, many acknowledged that previous expectations of the verbatim had perhaps been unrealistic. One Deaf BSL user said, “I am expecting too much, need to have a broad outline of what has been said rather than missing half the information with less time to read each word!” (OL30). A deaf participant said that although they prefer the verbatim style, perhaps at live events standard subtitles may be better as a compromise (OL12). Others maintained verbatim as a choice, reiterating the fact they wanted full access to what was said, not what someone else viewed as being important or they stated in their answer that they were capable of reading verbatim subtitles.

Reviewing these answers, it seems that, regardless of whether preferences changed, this was a useful discussion to have with users.

5.3.3.6 Priorities in the different settings

Having had the opportunity to reflect on these different subtitling conventions and styles individually, the final task for participants was to state their priorities in the three settings of live television, the theatre and live events. They were given a list of 14 features and asked to rank them on a 5-point Likert scale, which ranged from ‘Not at all important’ to ‘Extremely important’. The list included features relating to the timing, editing and presentation of subtitles as well as features of subtitle content. The results are captured in tables 5.10-5.12 below. Clear trends were seen across the three settings.

Table 5.10: Priorities for live subtitles on television

Priorities for live subtitles on television	Ranking of 'Extremely important'
Little/no delay in subtitles appearing	71.7%, n38
Error free	35.8%, n19
Easy-to-read font	35.8%, n19
Specialist words appear accurately	35.8%, n19
Verbatim	30.2%, n16
Standard	30.2%, n16
Little/no obstruction of the picture	28.3%, n15
Indication if content has been omitted i.e. (...)	26.4%, n14
Indication if the audio cannot be heard by hearing people i.e. (INAUDIBLE)	20.8%, n11
Indication if the audio is in a different language i.e. SPEAKS FRENCH	17.0%, n9
Fully edited	9.4%, n5
Sound effect labels i.e. APPLAUSE	7.5%, n4
Music labels i.e. NATIONAL ANTHEM	7.5%, n4
Passages to be cued out where possible	3.8%, n2

When asked about live subtitling on television (table 5.10), little/no delay in the subtitles appearing was the highest priority with 71.7% of respondents ranking it extremely important. Then followed a cluster of closely ranked priorities, all selected by more than a quarter of respondents (and highlighted in orange): subtitles being error free (35.8%, n19), in an easy-to-read font (35.8%, n19), having specialist words appear accurately (35.8%, n19), having verbatim or standard subtitles (both 30.2%, n16), little/no obstruction of the picture (28.3%, n15) and having an indication of whether content had been omitted using (...) (26.4%, n14). Indications of the audio being inaudible or in a different language weren't ranked as 'extremely important' as frequently, but were ranked highly as 'very important' features to include (35.8%, n19 and 50.9%, n27 respectively). In contrast, the feature which most respondents (24.5%, n13) ranked as 'not at all important' was 'fully edited subtitles'.

Table 5.11: Priorities for captions at the theatre

Priorities for captions at the theatre	Ranking of 'Extremely important'
Little/no delay in subtitles appearing	62.3%, n33
Error free	47.2%, n25
Specialist words appear accurately	47.2%, n25
Easy-to-read font	45.3%, n24
Verbatim	41.5%, n22
Little/no obstruction of the picture	39.6%, n21
Standard	26.4%, n14
Indication if the audio cannot be heard by hearing people i.e. (INAUDIBLE)	24.5%, n13
Indication if the audio is in a different language i.e. SPEAKS FRENCH	22.6%, n12
Indication if content has been omitted i.e. (...)	20.8%, n11
Music labels i.e. NATIONAL ANTHEM	17.0%, n9
Sound effect labels i.e. APPLAUSE	15.1%, n8
Passages to be cued out where possible	11.3%, n6
Fully edited	5.7%, n3

When asked about theatre captions (table 5.11), a very similar group of priorities gained the highest rankings, from over a quarter of the participants that were ranked 'extremely important'. Indications about the audio were ranked a little lower and were most often ranked as 'very important' or 'moderately important', so, at the theatre, the dialogue is clearly the most important feature and there is, perhaps, an understanding that it will naturally be captured as the captions are pre-prepared. Similarly, there is a clear preference for verbatim over standard captions (41.5%, n22 compared to 26.4%, n14) and a greater expectation for them to be error free, with specialist words appearing accurately. Once again, this reflects the time a theatre captioner has to prepare.

Table 5.12: Priorities for subtitles at live events

Priorities for subtitles at live events	Ranking of 'Extremely important'
Little/no delay in subtitles appearing	49.1%, n26
Easy-to-read font	47.2%, n25
Specialist words appear accurately	45.3%, n24
Verbatim	37.7%, n20
Little/no obstruction of the picture	28.3%, n15
Standard	26.4%, n14
Error free	26.4%, n14
Indication if the audio cannot be heard by hearing people i.e. (INAUDIBLE)	18.9%, n10
Indication if the audio is in a different language i.e. SPEAKS FRENCH	18.9%, n10
Indication if content has been omitted i.e. (...)	18.9%, n10
Music labels i.e. NATIONAL ANTHEM	11.3%, n6
Fully edited	9.4%, n6
Sound effect labels i.e. APPLAUSE	5.7%, n3
Passages to be cued out where possible	3.8%, n2

Last of all, the participants were asked about their priorities for live subtitling at events (table 5.12). The same group of features were prioritised in the scoring. A clear preference for verbatim (37.7%, 20) over standard subtitles (26.4%, n14) was seen, but there was less of an expectation for inaudible, foreign and omitted content to be indicated (all 18.9%, n10). Most participants felt these were 'moderately important' to include. The focus of participants was clearly on the respeaker capturing what was said in an accurate and timely way.

As well as confirming findings from other reception studies, such as the DVT4ALL study (Romero-Fresco, 2015: 157-160), these results also suggest that any work and training respeakers do in new environments will complement their in-house work, since both will serve to meet the audiences' needs in a similar way. The discussions of these expectations feed into the respeaker training programme in Chapter Six. For all groups, having specialist words recognised accurately at live events was a high priority, often higher than in the other settings. Not having delay was important throughout, but perhaps because of the discussion about what the process of respeaking involves, fewer people ranked it as 'extremely important' at live events (49.1%, n26) as compared to on television (71.7%, n38) and at the theatre (62.3%, n33).

5.3.3.7 Access at on television and at events: A thematic analysis of the qualitative responses

In sections B and C of the online questionnaire, participants had the opportunity to comment more freely on access, and reflections on this topic could be found elsewhere in the questionnaire and in the focus groups. I invited participants to recount and comment directly on live arts and cultural events that they had attended, either with or without access (C1). I hoped that their accounts would discuss subtitles but also broader forms of access and that I would learn of a range of experiences, both positive and negative. In addition, I invited them to talk about the changes they would like to see in each of the three scenarios (B13-15); finally, I asked how they would like the events that did have subtitles and captions to be chosen (C2).

The availability of subtitles and captions was a key issue. On the one hand, people wanted options: they wanted more access, at more locations, at a range of times, and for a range of content as not everyone's needs were the same:

I realise not all can be captioned, but those of us who need them are not the same and so we need a range of options to be able to attend... *(48-year-old woman with a hearing aid, deafened at the age of 18, OL43)*

Many referred to the fact that afternoon performances that tend to be captioned were not suitable for people who worked.

For others, it was not a matter of choosing what content should be accessible, but of removing the need to choose. Put simply, access should be there:

I don't think that content is a relevant criteria. We should all be provided with equal access to whatever takes our fancy. *(67-year-old woman, deafened at 20, OL49)*

And:

I have no preference. I am interested in sports, politics, culture, wellbeing, news, social rights, the weather forecast... Everything should always be made accessible. Full stop. *(40-plus-year-old person, no gender, deaf since their teens, who has a hearing aid, OL37)*

Given that I had invited participants to share experiences of events where there was a lack of access, I had expected to find comments about a lack of availability of access in

this sector. What I hadn't expected was for similar comments to be made for television subtitling as well. Often, in academia, we talk of the shift in focus that has taken place from quantity to quality (section 3.3 above), which seemingly implies that coverage on television is sufficient. Certainly, broadcasters come very close to reaching their Ofcom quotas (Ofcom, 2021: 5-6), but for many participants, this is not sufficient. For them, the first step for improvement would be to further increase the coverage of subtitles, across channels and also on the adverts and links that appear between programmes. Technical issues were cited as another reason that subtitle coverage dropped⁶⁹.

Whilst technical issues might present a particular case for television, this issue of access being promised, expected but not actualised was more widespread:

I booked to go to a paid talk. I asked in advance for a loop system – several times. I arrived early to check it was all OK. They said it was working. I switched to the loop setting on my hearing aid and there was nothing at all. I then sat through the whole event not being able to hear. I told them it wasn't working during the event and they did nothing to help at all. Nobody apologised or asked me how I got on at the end of the event. After the event I emailed and asked for a refund. They still insisted it was working which I know it wasn't. They never apologised. I have had experiences like this at many places.... It is a real problem and yet they all advertise themselves as accessible. I find it very frustrating and isolating. *(33-year-old woman who is hard of hearing. She has been deaf in right ear since birth and lost her hearing in her left ear in her early 20s. She now has a hearing aid. OL19)*

And:

I miss everything at uncaptioned events. With nothing to occupy me mentally, my thoughts spiral into self-pity and especially anger if subtitles were promised in the first place. *(68-year-old woman with a CI, who has been deaf since birth, OL2)*

Clarity and transparency about what access is and is not on offer is essential, since the access a person receives depends on what is actually provided, rather than what is promised. It was clear that for many participants a lack of access, or inadequate access

⁶⁹ Although unusual in its scale, the disruption caused to access services, including subtitling provision, following a fault at Red Bee Media in September 2021 highlights the potential impact that technical issues may have on provision and the importance of having backup measures in place and of keeping the public informed and updated if such incidents occur (Davies, 2021; Ofcom, 2021b).

has a decidedly negative impact on a person's health and well-being, as it prevents them from enjoying what is on offer around them and fully participating with family and friends. Mis-advertising only exacerbates this issue.

In contrast, when access is provided, events can change dramatically, as these two accounts from the same participant show:

The live subtitling access made it possible for me to follow completely. It was a fairly dark exhibition hall and the beauty of the handheld units is that I could relax, look at the exhibits and read the text without worrying about trying to be close to the speaker to lipread/hear.

And:

Live subtitles were projected on to the same screen as the slides. A difficult subject but brilliantly and accurately subtitled that lead to a lively discussion after the event for several deaf people at the pub. *(68-year-old man with a hearing aid, deaf from the age of 5, OL10)*

This reinforces the notion of access being a tool to human rights that was outlined in Chapter Two. Without it, barriers within society remain and the possibility of joining in and engaging becomes far harder, if not impossible.

This relates to another theme that I noted within the responses, one I have termed 'avoidance'. Many participants stated that they hadn't attended particular types of events, as they knew the challenges that faced them:

Never been to a live arts or cultural event as I knew it was not worth attending with no subtitles. *(62-year-old woman, Deaf since birth and grew up using BSL, has hearing aid. OL30)*

This avoidance may result from experiencing a lack of access at previous events, or from participants not being aware about the different types of events that are (sometimes) accessible. Whilst mis-advertising causes frustration, inadequate – or an absence of – advertising means potential audience members do not attend.

Many comments referred to broader aspects of access provision and the importance of fully embedding access within the event, in line with the proactive approach (Romero-Fresco, 2013; Greco, 2018, 2019a) outlined above (section 2.2.3.2). Some referred to economics, whilst others referred to awareness of those in charge of organising events:

I think the principle needs to be – how do you make those who organise the events feel that providing access is accessible to them? So that it is affordable and manageable and becomes an essential part of reaching audiences that would not otherwise attend? (38-year-old woman, deaf since birth, who has a hearing aid and uses BSL. OL46)

They felt more awareness was needed about why access is important, and how it could be achieved. Others highlighted what happened when the chain of access (see section 2.2.3.2 above) broke down:

Lack of access everywhere (at the festival). We had an interpreter booked to see an interview, but we were told to queue up like everyone else – but we could not go in the end as it was packed... There is so much that we can't access due to no subtitles and sometimes not enough interpreters. To cover all is not possible. (31-year-old woman, deaf since birth, has a hearing aid, OL29)

Whilst a general desire for access was common to all participants, different points of view were expressed about how that access should be provided and preferences over the type of access that should be offered varied. In the focus group, some expressed a strong preference for STTR over respeaking, feeling that it was more accurate or reliable:

I much prefer stenography. The accuracy is usually higher and the pace much more even. Both these factors – particularly in tandem – can make or break a viewing experience. (60-plus-year-old woman with a hearing aid, who became hard of hearing between the ages of 30-49, FGB6)

And:

I love stenography when it is a highly skilled STTR working in comparison to a poorly skilled STTR and respeaking does not provide an equally comfortable experience just yet. Too many errors means it is disruptive to one's experience. (A female Deaf BSL user, aged between 36-59, FGB5)

Others acknowledged the need for different types of provision, for example BSL and subtitles but also stated that they personally found BSL interpretation distracting if it was in view of the subtitles. In the case of theatre and events, different preferences were expressed over where captions and subtitles should be displayed. In general, participants seemed very aware that many different access needs exist; although happy to take part in the research, many explicitly said that they could not represent everyone

and did not want to become a spokesperson for all DH audience members. Rather, they highlighted instances where other people's needs and expectations might differ from the one they had just expressed. The needs of deafblind and colour-blind people and people with Usher syndrome were mentioned a number of times. By the same token, many participants also suggested other groups of people who might benefit from subtitles, whether at events or in other scenarios, including for example, dyslexic people.

One question that I did not include in the questionnaire was "What makes "good quality" access?" Nevertheless, potential answers to this question seem to be contained within the responses given elsewhere. It seems that availability, reliability, suitability of a particular type of access to a particular setting, use of multiple forms of access, and the embedding of access into the event, along with effective and accurate advertising form part of the answer; more succinctly, any consideration which means that the access provided allows people to engage in an event or situation seem key.

In the case of live subtitling more specifically, low latency and accuracy were the features that participants most wanted to change and improve, and technical reliability was key. Whilst some demanded them all to ensure equal access, others were grateful for what was (now) on offer.

5.3.4 Views from Non-native speakers of English (NNE)

I had originally intended to work as closely with this group as the DH group, but quickly realised this would not be possible. This is a heterogenous group, perhaps even more so than the DH audience group, and in many ways harder to target. Many questions arose as I began to plan these sessions. What level of English does a "language learner", referred to in the literature, have as opposed to a "non-native English speaker", the term I adopted? Do I distinguish between these two groups? What languages would I need to sample to really see how this group might respond to live subtitles? I firmly believe that NNE are a group who may benefit from subtitles, yet exploring this more fully was beyond the scope of the current project. Nevertheless, the data I do share here is an important starting point and a possible springboard for further research. As explained in Chapter

Three, while the use of pre-recorded subtitling among L2 learners has been explored in detail, to date I have not encountered any research into how non-native speakers respond to live subtitles.

One further reason why I did not want to discount this data was that many people I spoke to in this group were part of a marginalised community, waiting for outcomes on their asylum status. Unable to work, and in an often insecure setting, making a contribution to this research was something they were able to do. I do not want to further silence them simply because they were part of a group that was small in number, a choice that was mine not theirs.

5.3.4.1 Development of the focus group sessions

The NNE focus group sessions were closely based on the focus groups for the DH audience. Given that this audience group did not use subtitles extensively, fewer questions were asked, but the resources used to illustrate the different conventions were the same. I included additional demographic questions, which explored the participants' language history (appendix 5.5.1)⁷⁰. Unlike at the DH focus groups, the questionnaires were completed during the session so that they could also act as discussion points.

5.3.4.2 Who were the participants?

Seven participants took part, aged between 29 and 67. Their native languages included Dutch, Korean and Luganda. Employment information was not collected, but their education ranged from college to Masters level. On a scale of 0-10, where 0 was none, 1 was low and 10 high, participants were asked to rank their language level in the four skills of reading, writing, speaking and listening. Almost all the scores fell into the range of 7-9, and the participants seemed more confident in their reading (where 9 was the most common score) than listening (where 8 was most common). All had begun learning

⁷⁰ These additional questions were adapted from those found in the Language Experience and Proficiency (LEAP) questionnaire (Blumenfeld *et al.*, 2007).

English early in life, half around the age of five, and half around the age of 12. Some, but not all, had qualifications in English.

5.3.4.3 Viewing habits

In general, the participants watched an hour or two of television a day, and the amount of time they watched with subtitles varied from less than 10% to 50-75%. As a group, they were more familiar with intralingual than interlingual subtitles and the programmes they used subtitles for included the news, sports, documentaries, films, programmes where people had Scottish and Welsh accents and Jeremy Kyle USA. Their reasons for using subtitles included for better understanding and to access the dialogue, especially when a speaker's pronunciation or dialect, or their fast speech, made it hard to catch what was said.

Some participants had attended subtitled live events, most notably services at the church where I interviewed them, where PowerPoint slides contained the text of songs and most of what was said. In situations where there were not subtitles, the participants said they could understand most of what was said, and that they could usually follow along. When asked whether subtitles would improve their experience of live events, many thought they would and referred to the spellings of individual words, support if they lost concentration and increasing the enjoyment they would have at the event in question. Some said that the subtitles gave them useful information about the music and lyrics. The least useful feature of SDH was the sound labels, but most seemed to be able to ignore them and focus on the content that was useful to them.

5.3.4.4 Priorities in the different settings

As with the DH focus group, I asked these participants what their priorities would be for subtitles and captions in each setting but used an open-ended question rather than ranking system to encourage individual responses.

For pre-recorded subtitles on television, subtitles that were clear in terms of their content and presentation was a definite priority, as was them being in sync. Whilst some

participants would be happy with edited subtitles, as participant N4 said, “no need to subtitle every little detail”, others wanted what was said to match what was written. For live subtitles, summarising was deemed acceptable, but it was important for core words to be contained. Ideally, participants also wanted the pace of the subtitles to match the words spoken.

When it came to theatre captions, references were made to the captions being clear, well-timed, large enough to read and positioned close to the stage. Finally, at live events, as well as the priorities cited for live subtitles on television, it was important that these subtitles did not disturb what was taking place.

Most participants agreed that the subtitles helped them understand and engage with accents, faster speech and lyrics; they also found seeing the spelling of particular words important. When asked to rank their priorities for subtitles, having little or no obstruction of the picture, subtitling key words and well-timed subtitles were key.

5.3.4.5 Comparison of the needs and expectations of the DH and NNE audience groups

The motivation behind the priorities of the two audience groups often differed, with DH participants citing a desire for equal access, whereas NNE participants, who already had some access to the sound, were more concerned with prioritising the visual image. Nevertheless, there was consensus between the two in the expectations they set for live subtitling, with ‘having little or no obstruction of the picture’, ‘subtitling key words and phrases’ and ‘well-timed subtitles’ being priorities for both groups.

Since these expectations applied equally to television and live events, this seemed to confirm that by taking the needs and expectations of the wider audience into account, a better understanding of how to provide live subtitles at events, while promoting the embedding of access into the event as a whole would be gained. Similarly, gaining experience of this new setting would be beneficial for respeakers in practical and professional terms.

5.3.5 Venues

In order to consider how to set up a respeaking service for live events, it was important to talk to a number of venues to understand what was involved in hosting and organising events and what their concerns or priorities might be with respect to using respeaking (appendix 5.7).

To do this, I visited five venues - the V&A, Natural History Museum, Wellcome Collection and Royal Society in London and the Riverhouse Barn in Walton on Thames. At each, I spoke to a member of staff with responsibility for access and inclusion. Sometimes this was the Equality and Access Advisor or Diversity Manager; at the Riverhouse Barn, it was the Director who held this responsibility. At two venues, I also spoke to the Audiovisual and Venue Managers, so that I could ask more technically-oriented questions.

For most venues, respeaking was an unknown access modality, so the interview began with me explaining what it was and how it worked. This led to a number of potential audiences for respeaking being mentioned, including dyslexic people, learning disabled people and people with special educational needs and those accessing venues and museum content for academic purposes.

The remaining content varied across venues. Where the venue was interested in hosting as-live respeaking or a research event, much of the discussion related to what this would involve. At other venues, the conversation focused on more broader aspects of access provision.

5.3.5.1 Key themes

For the venues interested in hosting events with respeaking, the discussion focused on the challenges that using this service might pose, what preparation would be involved, the degree to which speakers or tour guides might be required to change their style when being made accessible and in what scenarios it could be used, especially in relation to the events each venue currently offered.

At the other venues, the discussion oriented towards what accessible provision was currently on offer and how events were chosen and scheduled. It was clear that including access in the event budget was vital in ensuring there would be provision and often the way provision was approached was to make sure that at least one event of each type was accessible, so that audience members had a chance to sample everything.

When it came to implementing new activities, there was a need for discussion about the provision, which often began with a series of trial events so that a venue could explore how to gradually break into a new audience domain. Challenges when setting up new services in the past had included prohibitive initial costs, sharing the correct information for it to be able to run smoothly, finding the right provider, and advertising.

This final point was an important consideration for all venues, for new and existing services. They often promoted events and exhibitions through linked associations, for example Stagetext, and spent time building mailing lists. Many referred to the challenge of having accessible events that no one attends, though it also seemed important that access was provided regardless, so that a venue gained a reputation of being accessible, even if that access was not used on every occasion. There was also a recognition, in line with the expectations of DH audience members, that advertising should be accurate (see section 5.3.3.7 above).

The interviews also included broader discussions of access, which went beyond individual events to include the approach to access that the venue or institution as a whole had, for example in terms of their more permanent collections or galleries. In many cases, it was important to understand how an organisation operated at an institutional level so that access could be embedded, and one approach to this was making access one of the core values. This way, there was an emphasis on there being “a baseline expectation of access” (Venue B) that is present throughout. To do this, it was essential that the venue, and its staff, really understood who their audience were. On the one hand, it is important to offer as much as possible for any access audience, on the other hand, there may be the question or aspiration of becoming a leader in the field for a particular audience (Venue B). It was clear that access was very much a process that

was continually being implemented and updated (Lazard, 2019: 10). Very often, for it to be implemented well, it needed to feature in all areas of the organisation, from initiatives to training, since the responsibility for it being achieved well lay with all members of staff, even if one person had overall responsibility for it.

Providing access involved thinking about the whole visiting experience – from the route, to lighting, to different types of access that would be needed, something which is very much in line with the chain of access (Greco *et al.*, 2012, section 2.2.3.2 above). Very often, having multiple types of access, for example BSL alongside STTR, was the choice a venue went for, and most venues said that access modalities which provided access for more than one kind of audience were particularly beneficial, especially if it was easy to implement. Many venues encouraged their visitors to let them know what access requirements would better help them access the event or exhibition in question.

5.3.5.2 Priorities when using respeaking at live events

When it came to priorities for using a service such as respeaking to provide access at an event, there was general agreement across the venues. It was important that the service was flexible, and ideally one that many could make use of; its cost should not be prohibitive; venues would want to know that they could promote the event and get people to it, and, for that, visitor experience was key. In addition, they would want the security of knowing that everything would run smoothly.

5.3.6 Respeakers

5.3.6.1 Data from respeakers

The data collected from the respeakers came from interviews and an online questionnaire. The respeaking companies I approached imposed certain restrictions on the extent to which their staff members could take part in this research, and this extended in some cases to the questions I was able to ask. For this reason, two versions of the questionnaire were used; most participants completed the full version of the questionnaire (appendix 5.6.2), but a shorter version was sent to one company. The

combined responses to the questionnaire can be found in appendix 5.6.3 and an overview of the content of the respeaker interviews can be found in appendix 5.6.1.

The interviews followed a semi-structured approach and provided an opportunity to discuss the participants' experience of respeaking to date and what they thought about the prospect of respeaking at live events. The questionnaires were developed from what was said and table 5.13 shows the range of topics covered:

Table 5.13: Topics covered in the respeaker questionnaire

Section	Topic	Content
A	Demographics	<ul style="list-style-type: none"> • Age • Gender • Qualifications • Languages spoken
B	Respeaking training	<ul style="list-style-type: none"> • Length of time working as a respeaker • Initial training • Subsequent training
C	Your usual respeaking environment	<ul style="list-style-type: none"> • Working set-up • Conditions that make respeaking most comfortable
D	Audience awareness	<ul style="list-style-type: none"> • What are the audience's needs and expectations? • How aware of them are you as you respeak?
E	Respeaking preferences	<ul style="list-style-type: none"> • Challenges associated with different programmes
F	Respeaking style	<ul style="list-style-type: none"> • Verbatim • Standard • Edited
G	Personal experience of viewing/attending live events	<ul style="list-style-type: none"> • Television in a public space • Live events • Respeaking vs STTR
H	Respeaking at live events	<ul style="list-style-type: none"> • Do you have any prior experience of this? OR • What are your thoughts about new setting?
I	Equipment and Technology	<ul style="list-style-type: none"> • Working with Dragon
J	Working conditions at live events	<ul style="list-style-type: none"> • Duration of respeaking blocks • Preparation materials • Working set-up
K	Co-editing subtitles	<ul style="list-style-type: none"> • Do you have any prior experience of this? OR • What are your thoughts about this?
L	Priorities for successful respeaking	<ul style="list-style-type: none"> • Ranking question
M	Training for respeaking at live events	<ul style="list-style-type: none"> • What would you like to be included?
N	Freelance work	<ul style="list-style-type: none"> • What would a fair hourly rate for live events be?

5.3.6.2 *Who were the participants?*

The respeakers taking part in the interviews came from three companies (A, B and C). Eight men and four women took part and their experience ranged from 6 months to 12 years, with all but one of them having worked as a respeaker for more than two years. Of the fourteen respeakers who completed the questionnaire, one was a freelancer who owned their own company; the others came from companies B and C, but many participants had experience of respeaking at more than one company. Their ages ranged from 18-25 to 36-59. 57.1% (n8) were men and 42.9% (n6) were women. They were based in the UK, Europe and Australia. Many had BA and MA degrees and for one participant their highest qualification was a Higher School Certificate and Certificate IV in Auslan (Australian Sign Language). Their degrees include arts, journalism, languages and psychology. 21.4% (n3) did not speak an additional language and 28.6% (n4) were fluent in two languages. Most participants (50.0%, n7) had more than six years of respeaking experience, and only one person (7.1%) had less than a year of experience.

5.3.6.3 *Respeaking training and knowledge*

All the participants were trained in-house, in line with the prevalence for this kind of training seen in Robert *et al.*'s survey (2019: 111) into who live subtitlers are. The duration of this initial training period varied, from two weeks plus ongoing quality control feedback to six months. The median duration was six weeks. The most beneficial aspects of this training had included getting used to the listening and speaking at the same time, speaking clearly in flowing sentences, learning how to use macros and house styles (see Chapter Three above) and gaining confidence in respeaking through opportunities to practice. More advanced skills were also referred to, including being able to anticipate errors, make (better) decisions, and working and editing at speed. Gaining a better understanding of the software also helped the respeakers to refine their skills. Almost all (83.3%, n10) had received further training after this initial period and over their last year of work, the amount of additional hours of training ranged from 0-20. This

included training on software, new content, reviewing work and NER analysis, captioning for autistic people in the case of one person⁷¹ and voice coaching.

5.3.6.4 Working environment

Most respeakers had worked in different environments. An open plan room was the most common setting, though people also worked in shared booths, or small rooms or office spaces which held three people. Over a quarter of participants had or did work from home. Working in a quiet room with minimal background noise was conducive to being comfortable while respeaking, and many preferred working alone in a room, or at least being the only person who was live.

The set-up and equipment used also contributed to how comfortable the working environment was. Many participants mentioned having a comfortable chair, a desk at the right height and having enough space to work. Having two monitors allowed subtitles to be displayed on one screen and left a second screen free for accessing the audio feed and doing research. Having a good-quality audio feed was essential, and some participants also shared their preferences for headsets (over the ear and noise cancelling). Preferences for microphones varied and included both USB-connected microphones and a standalone set-up. An internet connection was important for research and one person mentioned having water and tea available. The key message, reinforced in the interviews, was that the workstation and workspace should be comfortable and practical so that respeaking could be carried out smoothly, and this is something explored further in Chapters Six and Seven.

Although all the respeakers worked independently, as per the mono-model of Remael *et al.* (2016: 125-126), they were still part of a team; a few people referred to how this team could affect the quality of a person's working environment as much as these physical features. For some, understanding where they fitted into the process was key, and they appreciated working with people who were "prepared or ahead of schedule" (participant

⁷¹ Further information on this type of captioning can be found in footnote 49.

RO6). Others talked about etiquette when sharing a booth, saying they would always check with the other person before training in new words while the other person was live. For others, receiving materials about programmes that had been prepared by the wider respeaking team that meant they could begin working on new content more easily. Information about what topics would be covered and a list of terminology with proper nouns and acronyms were most helpful, as was having enough time to make use of this information and train and practice new vocabulary items.

When it came to respeaking content, some participants preferred to have prior general knowledge of subject matter, whilst others felt that being too involved in the content could be problematic:

I think sometimes if you are a bit emotionally involved in the content that can actually be a bit of a handicap... You might be speaking faster or you might be speaking louder or you might be a bit more animated and that is where your extra words and your failed recognitions can sometimes come up. So even if I was subtitling The Ashes - and cricket is my number one sport - if I was doing The Ashes, I have learnt over time to be in a bubble and not get too emotional about it – who is doing what (RIN4).

Most participants tended to consider that working with unfamiliar content was a natural part of respeaking:

One of the key skills of being a respeaker is you almost have to be able to learn to speak what you hear without necessarily understanding it because we cover so many topics... (RIN5)

Nevertheless, in the questionnaire, every participant said that familiarity or affinity with the subject matter helped their respeaking to some extent, with 16.7% (n2) selecting “moderately”, 50% (n6) “a great deal” and 33.3% (n4) “an extreme amount”.

5.3.6.5 Variations in content and subtitling style

Sections E and F of the questionnaire related to the participants’ experiences of working on different programme content and when different respeaking styles might be more appropriate. As I discuss this area, I also draw in comments made elsewhere in the interviews and questionnaire about the participants’ awareness of audience needs, since this also relates to how the respeaker approaches different kinds of content.

One programme type that was particularly challenging to respeak was sport as a result of its speed, the range of names and nicknames that might be needed, the potential for anecdotes and the background noise that was often present. Although university lectures might be slower-paced, they were often jargon heavy for example in mathematics, chemistry and linguistics lectures and as a result sometimes required unprepared terminology to be spelled out. In many cases, it was a feature of a programme rather than a particular programme type that posed the challenge and banter, quick exchanges, overlapping speech, shortened or made-up words and unfamiliar content were all given as examples.

In general, the participants seemed to have a good sense of what audience needs and expectations were when using respoken subtitles to access a programme. They kept these in mind as they worked either 'a great deal' (58.3%, n7) or 'an extreme amount' (41.7%, n5) of the time. Accuracy, skilful editing, and the need for subtitles to be as in sync as possible were mentioned many times. Some referred to the importance of clear punctuation, colour differentials between speakers and the need for clear subtitle delivery, "without spitting out of subtitles or jerky delivery making reading difficult" (RO10). Many participants suggested that audience expectations might be unrealistic, saying that expectations for live subtitles often matched those for pre-recorded subtitles and another highlighted the gap between what the audience might want and what they might need:

Audience expectations: 100% verbatim record of what is said, no losses due to inaudible or fast speech; audience needs: 99% accuracy, content accurate rather than word for word record, editing paraphrasing acceptable. (RO13)

In addition, the fact that some audience members might see an element of censorship within subtitles was also recognised.

The job of the subtitles is not to censor or to refrain but to represent the original content as closely as possible. (RO10)

Having been given a definition of verbatim, standard and edited respoking styles (table 5.14), the participants were then asked a series of questions to determine what factors,

be it programme type, audience expectation, company policy or personal comfort, influenced their choice of style (Section F).

Table 5.14: Definition of verbatim, edited and respeaking styles for respeakers

Verbatim subtitles: Every word spoken is included – the subtitles are a literal and faithful transcription of what has been said.
Standard subtitles: This is a middle way between verbatim and edited – the subtitles retain much of the language of the original, but some words have been paraphrased or omitted.
Edited subtitles: The original words are condensed and simplified – the subtitles facilitate reading and understanding.

The majority of participants (91.7%, n11) said they used the standard style most, whilst one (8.3%) said verbatim. 75% (n9) of participants said standard was also the style they felt most comfortable using, whereas 16.7% (n2) were most comfortable respeaking verbatim and 8.3% (n1) edited. Some said the predominant use of standard was because of company policy, whilst others said that it was the nature of the content that determined which style was adopted, with certain programmes being simply too fast to be able to respeak verbatim:

I think it's very difficult to create verbatim subtitles in a live setting, especially when expected to get all of the interjections (well, so, therefore...) In my opinion, the overall quality of captions goes down when the respeaker tries to capture everything. (RO14)

On the other hand, there were situations where some respeakers said they aimed to be as verbatim as possible, for example in the news.

Where participants referred to company policies, they seemed to ask respeakers to work as close to verbatim as possible but use editing where required. The aim was for subtitles to stay as close to the original meaning as possible, with some variation expected according to the nature of the content being subtitled. For example, sport would be highly edited, with content such as the ball commentary being omitted and a more edited style would be expected in university lectures where a subtitler worked alone. Similarly, what editing was expected might also depend on what a client had requested. Some

participants, however, said that there was either no company policy, or little oversight about this. Ultimately, the final decision lay with the individual respeaker as they worked on each task.

Although some participants said it was easier to respeak and follow word-for-word, it was clear from their answers about respeaking style that a degree of editing was usually required. When trying to explain what good editing and paraphrasing referred to, the participants referred to clearly capturing the message and conveying the original intended meaning, while including as much information as possible, capturing the tone, humour and personality of the speaker(s), all while maintaining flow and pace in well-structured sentences that were grammatically accurate and with good spelling. Whilst this is a composite answer, it nevertheless captures the complexity involved in the practice of respeaking. They would hope to achieve this by following the original wording where possible, using their own experience of watching television with muted sound to look out for confusing subtitles that they would avoid themselves and doing all they could to avoid delays in the subtitles appearing.

5.3.6.6 Respeakers' opinions of respeaking vs STTR

Having explored audience opinion of the difference between stenography and respeaking, it was interesting to hear from the respeakers themselves, the professionals who are used to critiquing and doing quality control of their own work, as they shared how they felt another form of STT fared. In considering this, it is useful to remember that for many years, respeakers and stenographers worked alongside and so were already likely to be familiar with each other's output. Some felt stenography had been and continued to be better, whilst others thought stenography had surpassed respeaking in the past, but that the situation had now been reversed. Others preferred respeaking outright as a modality.

The comments in favour of STTR referred to stenographers being able to work for longer as voices changed more quickly with fatigue, and also said it was more accurate,

appeared more quickly, was more likely to be verbatim and therefore better able to capture a high degree of information density.

In contrast, those who preferred respeaking suggested that it was more adaptable, offered more scope for editing and was therefore better able to capture the essential content and highlight humour:

I would prefer to watch certain sport with a stenographer. I think panel shows are better done by a respeaker as the content respoken is usually more essential or holds the humour or seriousness better...a personal touch? (RO6)

Having heard so much critique of respeaking on so many occasions, these varied opinions were refreshing to hear.

Others commented on the nature of errors seen across both modalities and pointed out that while those seen in respeaking could often be worked out so that the intended text was understood, nonsense errors might appear in stenography (RO10). Opinions on when respeaking and STTR might best be used varied but many did acknowledge that the quality of the subtitles or STTR would depend on the skill of the person creating them, regardless of the method used, which echoed the feeling expressed by participants in the DH focus group (5.3.3.7 above).

5.3.6.7 The move to live events

Out of those asked, 42.9% (n6) of participants had respoken at a live event and out of these, 83.3% (n5) had worked remotely, for the most part without a visual feed. Some had been able to talk briefly to the presenters or audience during the sound check, but many had no contact. One person mentioned receiving some preparation material as well. During the event and following it, there was little contact between the respeaker and either the presenter or audience members; the only exceptions to this were providing a transcript of the event, or troubleshooting any ongoing audio issues. When asked what, if anything, they would have liked to change about the experience, the responses included improving the audio, more preparatory materials and a more efficient way of contacting the people at the venue if needed. In the last case, the participant referred to

a chat panel within the software that was used for communication, but was not always monitored.

Those participants who had not worked at a live event were asked about the situational challenges they envisaged when respeaking on-site at events. Noise at the event, plus the possible distraction coming from the respeaker if they were heard were both mentioned. Participants were also concerned about conveying the wrong meaning, having stage fright and whether they would have access to preparatory materials, visual feed and equipment that they would require and the challenges involved in transporting and setting everything up correctly. One person added, "I can't think of any advantages to respeaking on site. We never do it". Propositions for getting around these challenges included being based in another room on-site, running AV tests with the organisers in advance of the event, having portable equipment and good preparation.

Many different events were cited as places where respeaking could be used, often with the proviso that measures were put in place to make the task easier (RO13). Some participants noted the need for caption on demand, especially at short notice (RO5), and that many people might appreciate the transcript that followed, for example, and some participants pointed out that having a respeaker may remove the need for someone to take minutes (RO12). As one person wrote:

I can't think of any live event that would be unsuitable for respeaking. It is such an effective means of quickly producing a written version of the spoken word it could be used effectively in all manner of scenarios. (RO11)

Where people were unsure about the use of respeaking at particular events, it related to scenarios where an alternative form of access might be better, for example SLI or pre-recorded captions, and once again to noise, both because the respeakers would inevitably cause a disturbance through lack of a better location, or because the respeakers would not themselves be protected from a noisy environment which could be damaging to them and their ability to work.

Participants were also asked about how they would approach the task of working with Dragon without their in-house subtitling software. House styles and macros were both

listed as tools they would want to set up. One participant realised house styles may not be available:

I think an entirely new Dragon profile might be needed without house styles. Perhaps certain profiles could be set up for certain types of events. I know people who have different dragons for sport or otherwise. A political or musical Dragon might be needed. (RO6)

Others mentioned the need for good training to ensure that recognition was effective, a presentation software that would snake out the captions so that the respeaker would not need to cue them, and also the possibility of a secondary respeaker who could make corrections to the output of the live speaker. This co-editing was something that 75% (n9) of participants had experience of and all participants were favourable towards, as it would allow them to focus on achieving accurate output rather than splitting their attention between output and self-correction. This feature was not tested during the study, but is certainly one to look at in the future.

The duration of a respeaking block varied greatly, from 15 minutes to closer to an hour. Many felt that 30 minutes was the point at which their voice would become uncomfortable, and, once again, it depended on the nature of what was being respoken. 35.7% (n5) said they would either be comfortable being the sole respeaker at an event, or might be, whereas 28.6% (n4) said they would not. The reasons for their uncertainty or discomfort related to the duration of the event, needing back-up in case something went wrong and what was expected of the respeaker in terms of the accuracy of the output.

Similarly, the amount of time that each respeaker would like to prepare for the event also varied. Thirty to sixty minutes seemed to allow adequate time to prepare the voice model, but additional time would be needed for any technical set-up and sound checks. Material-wise, participants would want access to similar information as they had for television work, as well as a run sheet or schedule, and access to this a day or two before the event would also be beneficial. A number of participants said they would like additional links to websites or PDFs so that they could conduct their own research about the content.

The question of where participants would like to be located at an event also led to a mixture of responses, though most wanted to be somewhere that they would not be visually distracting or disturbing to those present, and preferably out of sight. They did, however, want a clear view of what was happening. When asked what they would like presenters and venues to know about respeaking in advance of the event, participants referred to them understanding the nature of respeaking, that the output would not be verbatim and might include some errors, and that it would appear with a slight delay. They also wanted presenters to understand what they could do to help improve accuracy and to know that, as a result of this service, those attending could have access to a transcript of the event (RO12).

5.3.6.8 Priorities for respeaking at events

Having had the chance to reflect on working on television and at live events, I then asked the participants to reflect on what their priorities would be in each setting, on what would enable them to work to their best ability. I wanted to see any areas of overlap across the two settings, in order to be able to design the respeaker training more effectively, and also in order to be better informed about how this training could potentially serve as Continual Professional Development (CPD) or contribute to the professionalisation of respeaking.

The various statements were ranked in a similar way across both settings (table 5.15 below). The two highest priorities, quality of audio feed and familiarity with respeaking software, and the three lowest, knowledge of the audience, personal affinity with subject matter and familiarity with surroundings were the same.

Table 5.15: Priorities for respeaking on television and at live events

Respeaking on television	Score	Respeaking at live events	Score
Quality of audio feed	4.93	Quality of audio feed	4.86
Familiarity with respeaking software (Dragon, subtitling software)	4.57	Familiarity with equipment (microphones, headphones)	4.43
Internet access	4.36	Technical support	4.43
Acceptable payment	4.21	Information about subject matter	4.43
Information about subject matter	4.21	Preparation materials	4.43
Preparation time	4.07	Familiarity with respeaking software (Dragon, subtitling software)	4.43
Technical support	4.00	Preparation time	4.36
Preparation materials	3.93	Acceptable payment	4.21
Back-up respeaker to take over if needed	3.86	(Acceptable) duration of respeaking slot	4.07
Familiarity with equipment (microphones, headphones)	3.71	Internet access	4.00
(Acceptable) duration of respeaking slot	3.71	Understanding of your role as a respeaker for a particular event	3.93
Water	3.71	Back-up respeaker to take over if needed	3.79
Distraction-free environment	3.50	Ability to communicate with fellow respeaker	3.79
Ability to communicate with fellow respeaker	3.50	Water	3.71
Understanding of your role as a respeaker for a particular programme	3.43	Distraction-free environment	3.64
Access to visual feed	3.43	Familiarity with equipment (microphones, headphones)	3.64
Access to a second computer screen	3.29	Contact with the presenter during the day to clarify details	3.50
Ability to adjust table/chair height	3.14	Access to a second computer screen	3.50
Knowledge of audience	2.57	Ability for co-worker to adjust or edit your subtitles	3.43
Personal affinity with subject matter	2.50	Advance contact with the presenter	3.43
Familiarity with surroundings	2.43	Ability to adjust table/chair height	3.36
		Access to visual feed	3.07
		Knowledge of audience	2.93
		Personal affinity with subject matter	2.71
		Familiarity with surroundings	2.29

However, for live events, an increased range of statements were scored in the highest band of 4-5, highlighted in orange. Technical support and having information about the subject matter were higher priorities in the live event setting, and added to this group of higher priorities were preparation materials, (acceptable) duration of the respeaking slot and familiarity with the equipment. In other words, respeakers wanted to know the material and be familiar with the equipment in order to perform well. These were all

factors I considered as I approached the as-live respeaking and sought venues for the research events.

Finally, I asked the participants to rank possible content for the respeaker training programme, in order of how important it would be for them (table 5.16).

Table 5.16: Content for the respeaker training programme

Please rank the following topics where 1 is most important and 10 is least, for a training course in respeaking at live events	Score
Dragon profile management (vocabularies, replacing in house house-styles)	3.50
Voice model time (general practice)	3.93
Presentation software	4.14
Equipment (headphones, mikes)	4.57
Sample exercises on video – Q&A, presentations, open panel discussions	4.79
Respeaking in a recreated live event set up (i.e. lecture hall, walking tour)	5.00
Information on expectations for respeaking at live events	5.57
Co-editing a fellow respeaker's output	6.93
Information on how the audience will be accessing the subtitles (device used, presentation)	8.00
Getting to know fellow respeakers	8.57

In line with their priorities, the four highest-ranked topics were equipment-related, allowing them to develop and refine their profiles in Dragon and become familiar with displaying the subtitles. Next came opportunities for applied practice, working with sample exercises on video and a recreated live event. Less important were informational topics, such as expectations for respeakers in this setting and how the audience would be accessing the subtitles, and team-related factors, including co-editing practice and getting to know fellow respeakers. These would be things that could follow later.

This ranking fed into the development of the training programme that is outlined in the next chapter.

5.3.6.9 The priorities of service providers

Just as some similarities were seen in the expectations of the two audience groups, so too were areas of overlap noted between the respeakers and venues. For both, the

importance of the audience experience and the efficiency of the service were key. Having the opportunity to explore different aspects of the respeakers' working environment which supported their best work was very important, as it was this type of information that would very likely to be helpful for the venue to know, so that their needs could be accommodated during an event. Similarly, it is also clear that some compromises would need to be found. The respeaker training and research events would provide further opportunities for this to be explored, as would the as-live respeaking.

5.4 What is it like to subtitle a live event?

The final strand of the initial data collection involved as-live respeaking. I planned to attend a number of events to trial the process of respeaking, without transmitting any subtitles to the audience. My purpose in doing this was to get a feel for what it was like to work in a live event setting, with people around me, and to test out the effectiveness of the equipment and determine more fully what technical set-up would be required for use at the research events. I also wanted to see what working environment felt best suited to this work, where options were available.

Two opportunities arose for me to do this. The first event was the Annual Diversity Conference at the Royal Society in October 2016. I tested two different locations, the first a small side room, where I had a strong audio feed, but no visuals, and the second the technician's box, where I could see the event taking place, but which was a busier environment, with many technicians present. Out of the two, the quieter location was preferable and being aware of the content and being able to focus better compensated for the lack of visual feed.

By the time of the second as-live event in September 2017, I had compiled the full respeaking kit and I used this opportunity to test it. I respoke a lecture in the Royston Pike series at the Riverhouse Barn, the venue where two research events would be held. Before the event began, I tested the range of the subtitles and audio feed transmission around the building and sought possible locations for a respeaker outside the main event room. During the event, I was based in a space on the first floor of the venue, at the

opposite end of the building to the speaker. Once again, I worked without a visual feed, but the audio was reliably transmitted into my headset, and I was able to adjust the volume as required.

At both events, although there were audience members in the building, I worked undisturbed due to my location outside the main room. The full impact of working with an audience present would need to be explored further during the research events proper. However, the experience of the respeakers at the focus groups also provided some early insights into what it was like.

They were positioned at the front of the room, in very close proximity to the audience. I had thought that they might comment on this, perhaps negatively, in their feedback (appendix 5.8.2); however, it seemed that being so close to the audience actually worked in their favour, as it meant they could get added cues to help them respeak, since the audio feed was much poorer in comparison to what they were used to at work.

The two as-live experiences, alongside this early feedback from the focus group respeakers and my own experience of being respoken by them, highlighted the fact that the respeaking set-up and setting is going to change at every event. In each case, it will be essential to consider where everyone is located and the potential challenges that the environment or content may pose. Seeing that with even minimal preparation respeaking could be used at the focus groups was reassuring, even though it did not come close to providing the quality of access I would wish. Through that experience, however, I realised that even small changes could vastly improve the experience for respeakers and audience alike, and this was very useful to know as I entered more fully into Cycles Two and Three, where the full respeaking process and set-up was tested in action. Most importantly, the as-live respeaking and DH focus group sessions were practical evidence that the theoretical principle of integrating the access into the event was one that worked.

5.5 Review of Cycle One

This initial cycle of action research was an opportunity to explore the needs and expectations of key audience groups who might use a respeaking service at live events

as well as those of the respeakers and venues who would help provide it. What was striking in talking to each group was the commonality that could be seen among the ideas expressed, both between the two audience groups, and also across the audience groups and service providers. This sense of a mutual direction is a positive step both towards the provision of respeaking at live events and to the idea that all involved in event organisation might find it beneficial to work together towards this common goal.

Despite individual experiences, there were common features of subtitles that could be identified that were likely to be preferred by many users and there were certain factors that all agreed should be avoided. For the audience groups, subtitles which captured the main content of the event accurately and in detail, were essential if access was to be provided. Similarly, giving due consideration to where live subtitles would be placed to ensure easy and relaxed viewing would be an important factor for an event organiser to consider. For the service providers, meeting audience expectations and having a service that ran smoothly and reliably was key. Ultimately, both these ideas are complementary, even if the reasons for these expectations differed between parties. It became clear that to meet these expectations, providing the best working set-up for the respeakers, even in the changing environment and dynamic of the live event setting, would be vital to ensure that they could work to the best of their ability. In Chapter Six, the training that professional television respeakers would require and the equipment and technical set-up that would help achieve this is explored.

Chapter 6: Respeaker training and event programme

“Tell me and I forget,
Teach me and I may remember,
Involve me and I learn.”

Xun Kuang, 818AD⁷²

In Chapter Four, the methodological approaches adopted in this study were presented and the precise structure of this research project was explained, composed as it is of interweaving rounds of data collection from focus groups and blocks of action research. Following this, in the last chapter, the data collected in the first cycle of research was presented.

In the current chapter, I present the respeaker training programme and research events, which together form Cycles Two and Three of the research. Their conceptual design is examined, elements of the methodology underpinning them are discussed further and the adaptations which were made to them as the cycles progressed are explained. By the end of the chapter, the final version of the respeaker training programme and recommended style guidelines and technical set-up for use at future live events are presented.

6.1. The role of respeaking training and research events

While respeakers work alone as they create subtitles, there is a wider team around them supporting their work, monitoring output and providing technical support. In technical emergencies, there are alternative respeakers to ensure that subtitles continue to go to air. Preparation materials and programme archives are shared within the team, allowing time to be used more efficiently. In this way, access support companies are able to maintain the high quotas of subtitles expected of them (Ofcom, 2021: 5-6) and within each company, regular checks are in place to ensure respeaking quality is maintained.

⁷² This epigraph is derived from the John Knoblock translation of Book 8, Chapter 11 of the Xunzi (goodreads, no date).

In order to provide a stable and effective respeaking service for live events, similar systems must be put in place to ensure technical requirements will be met and that good quality subtitling will result, despite the variation that will be inherent in working conditions across events. Although other people present will be involved, respeakers at live events will need to take on the awareness, responsibility and support that the wider team previously held, and become familiar with new equipment, while making subtle changes to their usual working practice.

Respeaking is a challenging modality to perform, even in familiar settings, as a result of the real-time, technologically-dependent and un- or semi-scripted nature of the work. Transferring even known skills into a new setting, where content is likely to be less familiar or predictable, and working with a new range of technology in a more interactive setting, will require new knowledge and skills, and the opportunity to practice and refine them. In cognitive terms, such work might be considered as the development of adaptive expertise, where routine work and practices change, either as a result of new job and task requirements, or changes in the complexity or commonness of the situation (Carbonell *et al.*, 2014: 15)⁷³.

By running two rounds of training and events, which formed Cycles Two and Three of the research, the respeakers were first able to adapt their skills to this new working environment and then gain actual practice of respeaking within it, whilst allowing the research questions to be explored. On the one hand, scaffolding this process and providing opportunities for practice, discussion and reflection formed a key part of this training process. On the other, the events were designed to test a range of variables common to live events and expose numerous situations which might potentially pose different challenges for the respeakers.

The research events also provided an opportunity for the respeakers to gain feedback on their work and for their respeaking to be shared with the audience. The initial round of focus group sessions had been an opportunity to determine the key needs and

⁷³ A thank you to Elena Davitti who introduced me to the idea of adaptive expertise through our work on the SMART Project.

expectations of all involved in using respeaking to provide access at live events, from the users of the service (DH and NNE audience members) to those involved in its provision (the venues and respeakers themselves). Each expert group who attended the events was asked to share their feedback on the respeaking and the event as a whole. This feedback was invited through pre- and post- questionnaires and a focus group in the form of a post-event chat. Given the organisation involved in running the event, and the fact that most were held in the evening, running 1-1 interviews was not an option I could pursue.

6.2 Developing the training and events programme

6.2.1 Research Questions

In the first round of the training programme (Cycle Two), I wanted to establish how easily the respeakers would be able to adapt to the new equipment being used and to determine what procedures would need to be put in place. More importantly, I wanted to find out whether they had any questions about the process and how confident they felt about working at live events. In the second round (Cycle Three), I wanted to confirm the effectiveness of the technical set-up and equipment that had been selected and look more closely at the quality of the respeaking, to see how it compared to industry-standards (table 6.1).

Table 6.1: Cycle Two & Three Respeaker training and research events

Cycle Two – First round of respeaker training and research events (2017)				
Questions posed	Actions taken	Format of data collected	End of cycle Reflection	Outcomes
<ul style="list-style-type: none"> How easily can the respeakers adapt to the new equipment being used? What questions do they have about respeaking at live events? What procedures need to be put in place? Do the respeakers feel confident about working at live events? 	<ul style="list-style-type: none"> Training provided for a group of respeakers 	<ul style="list-style-type: none"> Observations Photographs, audio and video footage of training Discussions Questionnaires Participant reflections 	<ul style="list-style-type: none"> What adjustments need to be made to the training programme? Does the technical set-up need to be modified? Are any other changes needed? 	<ul style="list-style-type: none"> Technical set-up revised New equipment purchased if needed Supplementary training modules provided for the respeakers
<ul style="list-style-type: none"> What information does everyone involved need? Is everything in place? What will the quality of respeaking be at the live events? 	<ul style="list-style-type: none"> Four research events held as part of a large reception study Focus group discussion held at the end of each event 	<ul style="list-style-type: none"> Observations Photographs, audio and video footage of each event Transcripts of the respeaking Discussions Questionnaires Participant reflections 		
<ul style="list-style-type: none"> What is the nature of the access currently provided at live events? 	<ul style="list-style-type: none"> Continued attendance and observation at other live events 	<ul style="list-style-type: none"> Observations 		



Cycle Three – Second round of respeaker training and research events (2018)				
Questions posed	Actions taken	Format of data collected	End of cycle Reflection	Outcomes
<ul style="list-style-type: none"> Are the respeakers independent in their work? 	<ul style="list-style-type: none"> Supplementary training provided for a group of respeakers 	<ul style="list-style-type: none"> Observations Photographs, audio and video footage of training Discussions Questionnaires Participant reflections 	<ul style="list-style-type: none"> Are any further adjustments needed to the training programme? Is the technical set-up effective? Has suitable equipment been purchased? Are any other changes needed? 	<ul style="list-style-type: none"> Finalised version of the training programmes produced Guidelines created for RLE Information and training resources created for presenters and others involved in holding live events
<ul style="list-style-type: none"> How effective is the technical set-up and equipment? Will the events run smoothly? Will the quality of the respeaking be of a high enough standard? Is respeaking a feasible method for providing subtitles at live events? 	<ul style="list-style-type: none"> Four research events held as part of a large reception study Focus group discussion held at the end of each event 	<ul style="list-style-type: none"> Observations Photographs, audio, and video footage of each event Transcripts of the respeaking Discussions Questionnaires Participant reflections 		
<ul style="list-style-type: none"> What is the nature of the access currently provided at live events? 	<ul style="list-style-type: none"> Continued attendance and observation at other live events 	<ul style="list-style-type: none"> Observations 		

6.2.2 Defining quality in the context of respeaking at live events

In order to be able to determine the quality of respeaking seen at the events, and the extent to which it compares to industry standards, a precise definition of what exactly quality refers to is required. The concept of quality has long been discussed in TS, AVT and MA, but it remains an elusive one, that is inherently difficult to pin down and define. For this reason, Greco and Jankowska (2019: 8) proposed an agenda for media accessibility quality (MAQ), which was designed to enable “a more systematic and coordinated discussion” of what quality is. Many of the points made in this section reflect the actions they suggest.

With so much (necessary) attention placed on how to calculate the accuracy of live subtitles, there is a risk that in discussions on the quality of respeaking, this feature alone becomes representative of the quality of the subtitles as a whole, an approach described as synecdochal by Greco (Greco and Jankowska, 2019; Greco and Moores, 2021)⁷⁴. However, the fact that quality is multidimensional in its nature, whereby many different features or aspects of a product or service must be considered together to give a true picture of what the quality of something actually is, is increasingly being understood (*ibid.*). Certainly, the qualitative data from the audience groups considered in Chapter Five present quality in this more composite way, highlighting a range of factors that contributed to a judgement about what might be a ‘good’ or ‘poor’ experience both of

⁷⁴ In the Ofcom study into the quality of live subtitling, three dimensions were used to assess quality – the average speed of the subtitles, the average latency and the number and type of errors seen, i.e. the accuracy (Ofcom, 2013: 1). Ten different dimensions of quality had initially been proposed to Ofcom, as Pablo Romero-Fresco explained in a personal conversation, yet these three were selected by them. In the reports that Romero-Fresco and Pedregosa submitted after each round of subtitling, found in Annex 1 of each of Ofcom’s reports (Ofcom, 2014a, 2014b, 2015a, 2015b), they were able to extend the discussion to include the additional dimensions of the degree of editing and technical issues which affected the subtitles. The eventual consideration of Ofcom was a very detailed one and included additional comments on issues such as delays in transmissions to improve quality, late delivered programmes and the use of hybrid subtitles (Ofcom, 2013: 2-3). This process highlights the challenges and constraints that may be involved when considering a range of dimensions of quality.

respeaking, and access more broadly, among the different people questioned. When it comes to trying to determine the quality of access more broadly, no single scale exists⁷⁵. In the two chapters that follow, both approaches are followed. Chapter Seven presents the reception study carried out at the events, where the feedback from those involved is shared. The accuracy of the subtitles features in this discussion, but so do other factors ranging from the speed, latency and presentation of the subtitles to whether the audience felt they added to the access provided at the event. In other words, the *quality of experience* (Greco and Jankowska, 2019: 8) for those involved is at the heart of the discussion; this reflects the desire to provide a good visitor or audience experience expressed by the venues and respeakers who took part in the focus groups. In Chapter Eight, the discussion shifts to the *quality of service (ibid.)* and to the extent to which “a complete and accurate rendition of the original that does not distort the original message and tries to capture any and all extralinguistic information that the speaker might have provided” (Moser-Mercer, 1996: 44) is conveyed in the subtitles, in other words, to the accuracy of the subtitles. However, even in a discussion that focuses on accuracy, additional variables inevitably play a part:

Optimum quality in professional *respeaking* implies that a *respeaker* provides **a complete and accurate rendition of the original that does not distort the original message and tries to capture any and all extralinguistic information that the speaker might have provided** subject to constraints imposed by certain external conditions... Optimum quality is the quality a *respeaker* can provide if external conditions are appropriate. (*ibid.*)

Moser-Mercer’s full definition was one of “optimum quality”, since invariably, in a live setting like conference interpreting from where this was drawn (and hence my use of italics where *respeakers* has replaced *interpreters*), the extent to which this accuracy (quality) can be achieved will always be dependent on external conditions. The conditions Moser-Mercer listed (p.44-45) closely mirrored the factors the respeakers from the focus groups referred to in Chapter Five, and included, for example, the working

⁷⁵ Reports such as the State of Museum and Theatre Access (Cock *et al.*, 2018, 2019) and the Museum and Heritage Access Survey (Cock *et al.*, 2020) rely on user research and provide data on the different types of accessible performances that are available and how the figures for them change from year to year.

environment, content, speaker delivery, and the respeaker's own emotional response to the situation.

6.2.3 Principles behind the respeaker training and events

With this in mind, when developing the respeaker training and events that followed, all the decisions taken in terms of course design and content and choice of resources and equipment related very practically to the work that the respeakers would be expected to do and to what knowledge they would need in order to perform this work to the best of their ability; in other words, it was geared towards them developing expertise that could be adapted into new settings, while achieving optimum quality in their work.

There were times where my role was that of a more traditional trainer, determining the content that would be covered in particular modules and transmitting information and knowledge (Kelly, 2005: 57). At other times, however, my role was more one of a facilitator (Mackewn, 2008: 617), being more concerned with observing, holding the space and learning from the discussions and interaction taking place, only stepping in to engage more actively to explore something that had happened or a point that had been raised.

Whilst on a global level, the respeaker training provided opportunities for *learning* and the events for *doing*, the two fundamentals of action research, the training also included multiple opportunities for doing, through practice, discussion, interaction and reflection and the research events could be considered as examples of situated learning (Lave and Wenger, 1991), where learning could be applied in authentic surroundings (Risku, 2016: 16).

The training and research events were structured in a way which enabled the respeakers taking part to become informed participants within the research project. It was essential that they had the opportunity to engage in the research and understand the principles behind it. The research questions were shared with them, as they were what I wanted to explore at each event. On the day of the event, their role was to respeak; mine was that of researcher, observer and organiser. I was not involved in the respeaking process,

other than to monitor the subtitles, as I had seen Stagertext do (section 5.2.2 above) and in line with the role of access co-ordinator (section 3.7.2). Once the event was over, our discussion resumed. In this way, it was hoped that the training and events would help to create a community of practice (Lave and Wenger, 1991; Berthaud and Mason, 2018), where the practice of respeaking at live events could be developed and improved over time. Having the same respeakers involved in the research from start to end, although not part of the original action research design, strengthened this process.

6.3 Who were the respeakers?

6.3.1 Selection Process

A number of respeakers were interested in taking part in the training and events process, including some who had taken part in Cycle One and some who had been unable to. All were contacted and provided with information about what would be involved, namely approximately twelve hours of training and participation in at least two events, which would each last two hours plus preparation time. These figures were based on eight respeakers taking part in the training. Whilst there would be no remuneration for taking part, travel costs within the UK would be reimbursed and food would either be provided or paid for. The relevant access companies were also contacted to share this opportunity with other potential candidates and to ensure they gave their permission for participation, in accordance with ethics requirements.

The selection criteria for taking part were fourfold: the respeakers needed to be experienced, have enthusiasm for the project (due to the commitment involved), be able to attend training at the University of Roehampton and, where applicable, they needed the permission of their company to take part. I did not qualify what 'experienced' referred to, as I intended to speak to everyone who applied to gain a sense of how long they had worked professionally and whether, during this time, they had gained enough confidence in their ability to pass across a 'confidence threshold' in respeaking, a threshold that was referred to by many respeakers during the focus group interviews. Whilst the minimum period needed for basic training as a respeaker is 2-3 months (Romero-Fresco, 2011:

25) and between 2 weeks and 6 months according to the responses from the respeaker interviews, longer is needed to cross this threshold, which seemed to happen after respeaking for one and a half to two years (appendix 5.6.1).

Although ten respeakers wanted to join the first round of training, five respeakers actually did so. Four had worked at company A, although one was now in different employment, and the fifth was a freelancer. Company B did not give permission for their employees to take part in this initial round of training, as it coincided with a busy output period, but were still considering more participation in later rounds.

As the first round of training and events progressed, it became clear that company B were unlikely to be able to allow their respeakers to take part in the next round. At the same time, I could see the growing commitment of the respeakers I was already working with, as they increasingly stepped into the role of informed participants. For this reason, I invited them to continue in this role through the second round of training and events in Cycle Three. This continued relationship allowed the training programme and events to be refined in a different way than I had originally expected, but one that ultimately allowed the role of the respeaker at live events to be explored more fully.

6.3.2 Respeaker Profiles

Respeaker A was a 28-year-old native English speaker with an undergraduate degree in Russian Studies and a postgraduate degree in Audiovisual Translation and Advanced Russian Translation. At the time of the training, he had worked as a respeaker for company A full-time for four and a half years.

Respeaker B was a 32-year-old native English speaker with an undergraduate degree in English Language and Linguistics and a postgraduate degree in Cross Cultural Communication and Applied Linguistics. She knew beginner's French. She had worked as a respeaker for a total of seven years; five and a half had been full-time at company A and for the previous year and a half she had worked on a freelance basis, respeaking for approximately 10-20% of her working week.

Respeaker C was a 28-year-old native English speaker with an undergraduate degree in French and postgraduate degree in Applied Translation Studies. In addition to advanced French, she had beginner's Spanish and Welsh. She had worked at company A full-time for three years.

Respeaker D was a 30-year-old native English speaker with an undergraduate degree in Spanish and Translation and a postgraduate degree in Audiovisual Translation. In addition to advanced Spanish, he had beginner's level French. He had worked as a respeaker for company A full-time for five years.

Respeaker E was a 38-year-old man from the Netherlands. He was not a native English speaker, but was fluent in English and had studied English Language and Culture as an undergraduate and postgraduate. In addition to his native Dutch, he had intermediate German and beginner's French and Spanish. He had fourteen years of experience as a respeaker and at the time of training was working as a freelance respeaker and translator. He came over to the UK to take part in the training, after hearing about the training opportunity at a conference the previous year.

None of these respeakers had any prior experience of respeaking at live events, but all agreed that the factors that made a programme more comfortable to respeak included having advance knowledge or an affinity with the subject matter and having adequate time to prepare. Good quality audio and visual feeds and a steady pace of speech to respeak, ideally with fewer rather than more speakers to follow, also aided their work. The factors that made respeaking less comfortable for them included lack of preparation time, fast-paced or multiple speakers, swearing, complex or unknown names or terminology and poor-quality audio. These views, shared in the mid-training questionnaires (appendix 6.1.6), were representative of the views expressed by the wider focus groups (section 5.3.6 above).

6.4 Equipment

An important aspect of the training programme was to determine what set-up the respeakers required to facilitate their work in the live environment. The core parts of the

respeaking kit, made up of the laptops, microphones, Dragon licences and presentation software, were purchased for the respeakers to use during the training programme and other items were added as their need was revealed, so that a full kit could be developed for live events. Most items were purchased in duplicate as each respeaker at the event would need their own kit.

6.4.1 Computer and speech recognition software

An educational licence for Dragon Professional Individual version 15 was installed on each laptop, which allowed for multiple users. The laptops purchased were HP Pavilion, with a 7th generation Intel Core i5 processor, 258GB Solid State Drive, 8192MB DDR4 SDRAM, Windows 10, 15.6" display, B&O Play. The use of the two computers was restricted to research purposes to ensure running capacities would be comparable.

6.4.2 Microphones

Classic SmartMics were purchased from TalkTechnologies (<https://talktech.com/>). Originally designed for court-reporting in the US, these muffle the voice of the respeaker to allow them to work in the same room as the audience. A USB audio adapter was used to connect them to each laptop.



Fig. 6.1: SmartMics in use

Each came with an over-the-nose face piece and a mouth-only face piece; a hands-free strap was also available for the respeakers to use. Additional mouthpieces and foam sets were purchased for hygiene reasons so the respeakers could try each style and experience the physical changes needed when respeaking with each type. Talk Technologies have since released a microphone stand, which will be offered to respeakers at future events.

Although only some research events would require the use of this microphone, the respeakers used them throughout training to maximise the time they had to adapt their technique to the face piece for this microphone.

6.4.3 Audio feed

During the training, the respeakers used their own earphones when practising with the software, and listened to the presenter in the live practice exercise in module four without a headset. A more sophisticated set-up was developed in time for the research events, which enabled the respeakers to have the audio feed streamed direct to a headset (see section 6.6.7 below).

6.4.4 Presentation software

Text on Top was the principal presentation software used in training to display the subtitles⁷⁶. The initial version used was 2.7.0 and updates were installed as they became available⁷⁷. Four USB devices were purchased. Two were for the respeakers' laptops, and one was inserted into the main computer to display the subtitles to the audience. The final device was a spare, but was later used at the events on an additional laptop,

⁷⁶ This same software is used by STTRs in the UK and by respeakers internationally at live events.

⁷⁷ Updates to the most recent versions of Text on Top, released after the research events considered here, mean that this software has been transformed from what I defined as presentation software into subtitling software. Dragon is integrated directly into the Text on Top interface and the respeaker now has access to temporary macros and is additionally able to create shortforms, whereby typing one or two letters will lead to a full word or phrase appearing on screen.

so that the presenters could also view the subtitles if they wished. The subtitles were transmitted through radio channels.

A second system, Streamtext, which enabled subtitles to be streamed to multiple tablets at the mobile events, was later introduced to the respeakers⁷⁸. Off-campus training was provided as there were difficulties using this system at the University of Roehampton.

6.4.5 Accessories

On the advice of the respeakers, a number of other accessories were added to the respeaking kit as the training and research events progressed: a freestanding keyboard and mouse to facilitate the use of shortcut keys, an LED light for work in dimmed conditions and a four-socket USB hub to enable everything to be connected. The kit for each respeaker fitted into a single bag.



Fig. 6.2: Respeaking kit

⁷⁸ Since October 2020, Text on Tap has offered its own online streaming service, called Text on Tap. Audience members can visit the text-on-tap-live website where the subtitles are displayed. In addition, an overlay version of this tool also allows subtitles to be positioned over PowerPoints (Text on Tap, no date), for example on the main display screen, if the respeaker is working remotely.

6.5 Respeaker training Round One: Design

6.5.1 Module overview

The respeaker training programme was a bespoke course, designed in response to the needs and expectations revealed by each focus group in Cycle One of the research. In all, eight modules were developed which lasted a total of 14 hours and which were split across three sessions.

Table 6.2: Module overview shared with the respeakers

Module	Contents	Duration (minutes)
First day of training		
1	Project overview <ul style="list-style-type: none"> How can respeaking be used to increase accessibility at live events? Project Design Training and trial events within the research 	30
2	Expectations at live events <ul style="list-style-type: none"> d/Deaf and hard of hearing audience NNE Venues Respeakers (so far...) 	40
3	Equipment <ul style="list-style-type: none"> SmartMics Text on Top Dragon Individual Professional Saving voice models and transcripts 	70
4	Practicalities of live events <ul style="list-style-type: none"> How might the set-up vary? Developing guidelines for good access 	90
5	Respeaking style <ul style="list-style-type: none"> Style guidelines How would you define your own respeaking style? Verbatim or edited subtitles? 	60
6	Practice <ul style="list-style-type: none"> Exploring Dragon as a stand-alone speech recognition tool Respeaking according to the different style guidelines Learning to hand over on Text on Top Saving profiles 	60
7	Review <ul style="list-style-type: none"> Review of today's session Requests for individual training 	30
	Post-training get together <ul style="list-style-type: none"> Informal chance to get to know fellow respeakers 	
Subsequent sessions		
8	Further practice and tailored training <ul style="list-style-type: none"> Voice model development – adding macros and vocabulary Familiarisation with SmartMics and Text on Top Respeaking according to the different style guidelines 	240

Since the respeakers were to be informed participants within the project, information about the overall aims of the project, and the role of the respeakers within it, needed to be shared with them at the start of the training. Module one, the project overview and module two, which outlined the expectations of all involved in live events, were designed with this purpose in mind.

In order to perform effectively, the respeakers needed to familiarise themselves with the practicalities of respeaking at live events, ranging from the equipment used and physical working environment to which style guidelines to follow and modules three to six covered this key content. Finally, there was an opportunity to develop “personal effectiveness” (QA, no date) at working in this new environment, which would come through gaining confidence in the live event setting and feeling supported by the other trainees and future co-workers during the research events. Module seven, the training review, allowed the respeakers to state any training needs they had. This marked the end of the first day of training, and the interval before the individual sessions meant that I was able to develop tailored content, as required, for the final module. The post-training social get-together allowed the respeakers to bond in a less formal setting.

6.5.2 Training logistics

Although designed as a whole group session, it was necessary to run the main training day twice. Three people attended the first session, and two came to the second. The personalised training sessions in module eight were scheduled to accommodate individual availability and the respeakers attended the sessions individually and in pairs. On arrival at the first training session, the respeakers were given a training pack. This included the key information and resources for each module (appendices 6.1.1, 6.1.3 and 6.1.4), questionnaires that the respeakers would be asked to complete during the training (appendices 6.1.5-6.1.7) and consent forms. A copy of the manuals for the SmartMics and Text on Top were also included.

The respeakers were each given a USB where they were asked to save their Dragon profiles, vocabularies and transcripts at the end of each session and after each event for

later analysis; instructions were provided on how to do this. The sessions were recorded and some notes were taken on a flipchart to facilitate the discussions.

6.5.3 Theoretical knowledge and understanding

Modules one, two and five contained the theoretical content of the training programme and will be discussed first.

6.5.3.1 Module one – Project overview

Module one was the most theoretical of the modules. The first half was delivered in lecture style, accompanied by a series of PowerPoint slides, followed by a slot for questions and the completion of the pre-training questionnaire, which asked the respeaker for demographic information and contained questions about their professional history as a respeaker (appendix 6.1.5).

The content enabled those attending to understand how and why the project was conceived; the need for wider access and for respeaking at live events was explained, as were key theoretical and methodological principles including the four-fold focus group approach, the cycles of action research, the collaborative nature of the project, and the concept of Universal Design (Mace, 1988), which was the stance towards access that I had originally adopted.

Particular mention was made of the involvement of Stagertext within the project and the place of the respeaking programme - and therefore the respeakers themselves - within the research project as a whole was highlighted. We also discussed the benefits I hoped the respeakers would themselves gain from taking part and I listed opportunities for professional development, the professionalisation of respeaking and the possibility of respeaking at actual (rather than research) events in the future as the desired outcomes of the research.

6.5.3.2 Module two – Expectations at live events

The second module was designed to facilitate discussion between the participants and to allow them to debate and discover for themselves the key expectations for respeaking and respeakers at live events. The expectations of each group (DH audience members, NNE audience members, venues and the wider respeaker focus group) were discussed in turn. I contributed with any expectations that they did not raise. I also highlighted the areas of overlap seen within the expectations, for example between the user and provider focus groups (section 5.5 above) and between the expectations for respeaking on television and at live events, as this reiterated the fact that the training was likely to be of benefit to them in their daily work.

6.5.3.3 Module five - Respeaking style

By the time they came to module five, a discussion on respeaking styles, the respeakers had created new Dragon profiles, gained some familiarity with the live event setting and completed the mid-training questionnaire where they outlined their own respeaking style. The module began with a discussion of general style guidelines for respeaking at live events. For the most part, these followed standard guidelines for UK television, and a few adaptations were included that were geared specifically towards live events (appendix 6.1.1: 15).

Next came a discussion on verbatim, standard and edited respeaking styles. Together, the participants came up with their own definitions for each style, and we attempted to determine whether the natural style of each respeaker matched that of any other respeaker, and/or one of the three styles introduced here. A working label for 'standard' respeaking was also sought that could better capture audience and respeaker expectations. Further discussion of the three styles follows in section 6.6.3 below.

The slides and training booklet were particularly helpful during this discussion, as they contained some early analysis of the questions on respeaking styles from the DH audience questionnaire (discussed in 5.3.3 above). In particular, the shift in audience

preference from verbatim to standard subtitles, once the potential problems associated with each style during live events had been explained, was highlighted.

When discussing the need for editing, I took care to make clear that this was considered a natural part of respeaking and that no negative judgement was being made on the abilities of the respeakers. Data from the study on the quality of live subtitling, led by Ofcom, was also shared, which revealed that in the 300 programmes reviewed during the two-year study, all contained some degree of editing, with percentages ranging from 13.3% in the news to 22.5% in entertainment shows and 31.6% in chat shows (Romero-Fresco, 2016: 66)⁷⁹.

6.5.4 Practical knowledge and understanding on day one

6.5.4.1 Module three – Equipment

Having the opportunity to explore the new software and equipment that the respeakers would be using was a central part of the training and this was the core purpose of module three. Following an initial demonstration of the SmartMic and Text on Top, and an explanation about the Streamtext software that they would be using at mobile events, the respeakers were able to practice more freely and explore the range of what Text on Top could do, as well as the constraints involved in respeaking into the SmartMic.

6.5.4.2 Module four – Practicalities of live events

Module four was designed to build on this initial familiarity with the equipment and software that the respeakers had gained in module three, and to help them begin to explore the practicalities of working at a live event. The session was split into three parts. In the first, the respeakers were encouraged to share their thoughts about what set-up might be needed at live events and how this could vary. Photographs of a number of venues supporting this project were used to stimulate the discussion.

⁷⁹ These percentages were confirmed in a personal conversation with Pablo Romero-Fresco.

Next, they were given the opportunity to respeak a live lecture and gain first-hand experience of subtitling a live event. There was a single presenter, who used PowerPoint slides and a short video clip within the lecture, which was on the language of music. Complex terminology was used, which included many foreign terms. The subtitles were displayed on the main screen and the respeakers took turns at respeaking and observing the access provided. At various points, the presentation was paused so the respeakers and presenter could give feedback on the process.

Finally, during the review, we worked together to draw up a checklist of steps that the respeakers would need to follow at the event for the service to run smoothly (appendix 6.3). In particular, the question of the pros and cons of working remotely and on-site was raised, since a number of respondents to the respeaker survey had questioned why respeakers would choose to work in person at live events⁸⁰.

6.5.4.3 Module six - Practice

The final module of the first day of training was an opportunity for the respeakers to continue to practice using the software, and to embed what they had learned from the live event simulation and discussion of respeaking style in modules four and five. The respeakers began by working on a video clip of a presentation from the Royal Society, an event which had actually been live subtitled.

6.5.5 Personal effectiveness – Day one and beyond

Whilst there was no specific module for developing personal effectiveness in the live event setting, all the sessions on day one, and indeed the follow-up sessions, were geared towards this. Opportunities for asking questions, clarifying details and learning through discussion and hands-on activities were designed to encourage active learning and confidence building.

⁸⁰ In this instance, 'remote' refers to a situation where the audience attend in person, whilst the respeaker works offsite, rather than being located in either the main event room or another room on site.

Although getting to know fellow respeakers had ranked low in terms of the priorities for training (section 5.3.6.8. above), I still wanted to create opportunities for supportive social connections to develop between the respeakers, and lunch, tea breaks and the post-training get together provided an informal way for this to happen. Most of the respeakers taking part in the training were colleagues, and knew they could rely on each other during the research events.

At the same time, establishing an individualised training programme for the follow-up training would ensure that the development of the respeakers as skilled and informed professionals remained at the heart of the respeaker training.

6.5.6 Follow-up training

In keeping with the action research framework, the shape of these individualised follow-up sessions was determined through exchanges and discussions with each respeaker. At the end of the first day of training, the respeakers were given a final questionnaire (appendix 6.1.7) where they were asked to feed back on the day, state their current confidence with the various items of equipment and software and identify what follow-up training they required.

These additional training sessions allowed the respeakers to work on different types of live event content and to continue to explore the different respeaking styles we had discussed, whilst refining their voice models. Further discussion of the logistics and finer details of the guidelines that would be in place at the events also took place during these sessions.

The respeakers logged their progress during the practice clips and noted how their confidence and familiarity with the equipment evolved.

6.6 Respeaker training Round One: Analysis

6.6.1 Professional approach

The professional interest and approach taken by all the respeakers involved in the project was clear from the outset. Each individual turned up on every occasion with a willingness

and dedication to developing new skills and putting their full effort into all elements of the research, regardless of the challenges involved and the complications posed by work schedules. Despite using the same skillset as at work, there was a sense amongst the respeakers that this was ‘something different’ and they appreciated the opportunity to develop their knowledge in new ways.

The training quickly became a mutual exchange of ideas among professionals, and an opportunity for everyone to learn, myself included. Discussion, followed by trial and error, often led to solutions being found as a team. All were concerned with being as successful possible, which equated to providing the best possible access through respeaking.

6.6.2 Respeakers’ needs

Inherent within the design of this research was a desire to meet the core needs of each of the focus groups in every decision taken. With the respeakers being the main service providers, and respeaking a complex and challenging service to execute well, it was clear from the outset that ensuring an effective basic working set-up would be essential.

6.6.2.1 Core needs

This first round of respeaker training revealed very clearly those conditions which were vital to ensuring the respeakers could work with ease and confidence and which would support them in facing the large challenges involved in working in a live environment. Maslow (1954: 15-19) writes of **physiological** and **safety needs** at the base of his hierarchy of needs, and these terms are equally applicable here.

Even before familiarity with equipment and software, respeaking relies on the effective use of another tool, the voice, accompanied by the breath. Discomfort or dis-ease may affect the basic process of speech and breathing and must be taken into account. A number of respeakers referred specifically to “staying calm” and “not panicking” when describing their respeaking style (Respeaker D, mid-training questionnaire).

The respeakers need to have water to hand and they must be as comfortable as possible with the microphone mouthpiece. During the practice sessions they all tested the mouth

and nose pieces on the SmartMics, with and without hands-free straps, and unanimously decided they preferred the mouthpiece in combination with the hands-free strap, despite the additional burden they then had of remembering to respeak more quietly. Respeaker B, in particular, experienced difficulty in regulating his breath (Session Log, August 6th), resulting in discomfort for him and a long delay in subtitles coming to screen. It was also discovered that the nose pieces were uncomfortable for glass-wearers, and that a hairband facilitated use of the hands-free strap for people with longer hair.

Variations in respeaking style placed additional physical burdens on the respeakers, which will be discussed further in section 6.6.3.2 below.

Maslow's needs of safety and security were particularly visible when considering the environmental dimension of the respeakers' work, a matter which was also highlighted in the online questionnaire (5.3.6.4. above). Individual respeakers had preferences in terms of the computer equipment they used. Even simple changes could add greatly to their ease of work and readiness to cope with greater challenges at hand. For this reason, a computer mouse and keyboard were purchased for each laptop, which the respeakers could use if desired.

The quality of the sound feed was an issue that had come up many times during the focus groups and it was repeatedly mentioned during the training as well. Being able to hear the original audio clearly is essential if the respeakers are to be able to respeak it well. When respeaking for television, the sound feed comes through headphones and the respeakers are able to adjust the sound level of the original audio and of their own respeaking. During training, the respeakers were listening to the audio directly and missed both the physical comfort of the wearing headphones and the control that led to. As one of the respeakers at the DH focus group had noted, when suggesting an audio feed that fed directly into the respeakers' headphones, "the big question is how to get that audio from multiple inputs to the headphones of the respeaker(s)" (appendix 5.8.2). This was a complicated problem to resolve, but silent disco gear eventually proved to be

the solution⁸¹. During training, the respeakers did have a visual feed to the events, and where possible, this was also arranged at the research events.

A bug was noted in Text on Top software during training, which led to joined words and doubled letters appearing. Even though they would be classed as minor errors, with minimal impact on audience comprehension (Romero-Fresco and Martínez, 2015), they were nevertheless frustrating for the respeakers as they detracted from the overall accuracy and therefore quality of their work. As software-specific errors (Moore and Romero-Fresco, 2015), the respeakers were unable to avoid them; they had to continue respeaking and train themselves to overlook them, while looking out for standard or serious errors which they would attempt to correct.

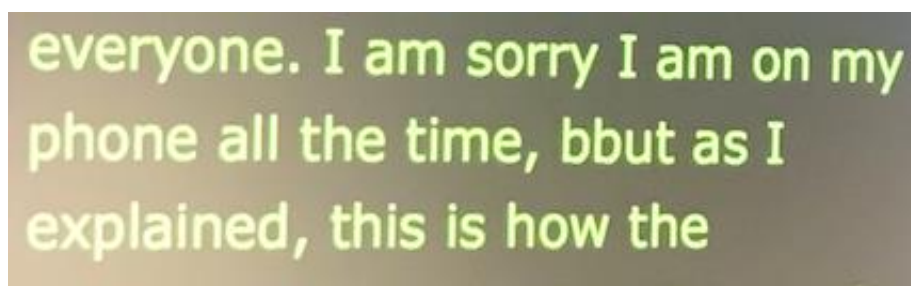


Fig. 6.3: Doubled letters in the respoken output

6.6.2.2 Higher level needs

After Maslow's (1954: 20-22) physical and environmental needs, come **sociocultural** and **emotional** needs. Despite the limited training time, and challenges posed by the equipment, ease and professionalism were shown by all five participants. However, Respeaker E decided during the individual training sessions that he would prefer to step into the role of observer during this first round of research events.

Despite being the most experienced respeaker in the group, his regular respeaking language was Dutch, his native language. Before beginning work in a new setting, in front of an audience, he wanted to gain more experience of respeaking in English. He

⁸¹ A thank you to Ben Wilson who suggested this solution to me.

felt less familiar than his counterparts with the particular challenges of respeaking in English and therefore less able to avoid errors occurring, yet his natural preference was to work in a situation where he was “top of his game” (Session Log, July 14th). His contribution to the research remains a very valuable one, and he returned to the UK to respeak as-live at the second event and now regularly respeaks live events in English. From an analytical point of view, this meant that the respeaking corpus analysed in Chapters Seven and Eight was created by native English speakers.

Respeaker E's decision also reveals the complexity involved in respeaking. The training programme described here is designed for experienced respeakers and respeaking skills per se are not taught. It does not cover the range of the content that is expected and required in an initial training programme (see sections 3.7.1 and 5.3.6.3 above). That is not to say that initial respeaker training cannot be done in a live event setting – it certainly can⁸², but additional modules with scaffolded practice, would be required to achieve this. Trainees would need to develop their ability to listen and speak at the same time, understand how to work with speech recognition software and learn to speak with a software adapted delivery⁸³. This additional training would be necessary for any new respeaker, be they theatre captioners hoping to respeak short sections of ad lib materials, for example in pantomimes, or for people intending to respeak full-time, though the length of this additional training may perhaps vary for each group. What this does, however, mean is that the current programme could be used in multilingual settings. Since it is primarily broader skills that are being taught, it could be easily adapted for respeakers working intralingually in a language other than English, and even interlingually.

The speed and delivery of the presenter and the nature of the content being respokey were other factors that impacted on the performance of the respeakers. For this reason, information about what respeaking was and guidelines on how to speak in a way that

⁸² At least some respeakers in Company C receive their initial training for educational lectures.

⁸³ This term was coined as part of the SMART project. See footnote 33 for further information about the project.

supported being respoken were created for the presenters at the research events. These guidelines and the role of the presenter are discussed further in Chapter Seven.

6.6.3 Respeaking style

In module five, the question of verbatim, standard and edited subtitles (Szarkowska *et al.*, 2011) and the possibility, practicality and appropriateness of using each while respoking necessitated a longer discussion.

6.6.3.1 *Natural respoking Style*

The mid-training questionnaire (appendix 6.1.6) revealed that all five respeakers involved in the training programme shared a broadly common practice with regard to their regular respoking style. All tried to use a style that was 'aiming for verbatim' or 'as close to verbatim as possible'. Their aim in doing this was to create subtitles that were as full, precise and as exact as possible. Where the speaker's speed and language allowed it, they would respeak exactly what they heard; where this was not possible, due to rate of speech, or the presence of multiple speakers, they would begin to edit, in order to achieve "clarity and readability" (Respeaker D) in their output.

When editing, they would aim to drop "non-essential" information, phatic phrases such as 'well', or 'as I said', or content which had been repeated. They would strive to maintain the idiosyncrasies of the original speaker wherever possible and would use macros to do this. Micropauses (Respeaker C) and clear enunciation (Respeakers C and D) became particularly important at speed to achieve accuracy, especially as corrections led to further delay.

The respeakers stressed that such choices were made when working quickly (Respeaker B), but that their judgement over what and how to edit depended on how well they had understood the original, how likely the speech recognition software was to recognise their utterances (Respeaker A) and on the particular speed and content of each programme (Respeaker B).

6.6.3.2 Testing verbatim, standard and fully edited styles

According to the spectrum of verbatim, standard and edited, this natural respeaking style, self-defined as ‘aiming for verbatim’ was a very close match to standard. I had originally intended to make a distinction at the research events between a ‘standard’ style and the ‘respeakers’ natural style’, but following the discussions during training, this contrast no longer seemed appropriate.

Since the fully edited style had not been rated as highly as an expectation of DH audience members (5.3.3.5), I decided to contrast and test the reception of verbatim and ‘aiming for verbatim’ within the first round of research events.

When trialling verbatim was suggested to the respeakers, they took up the challenge with enthusiasm. The aim of ultimately delivering what the audience wanted was important to them, though the question of whether, in reality, the audience would be able to follow the verbatim subtitles alongside the event, was raised.

During the initial discussion, it was clear that the respeakers felt more comfortable about the prospect of adopting a verbatim style than a fully edited style. To achieve success during heavy editing, careful decisions must be taken, yet respeaking, even the potentially slower respeaking that might be possible given the lower word-count expected in edited subtitles, does not grant a respeaker the time for such decisions. When defining what it would actually mean to subtitle verbatim, the idea of “no turning back” came up (Session Log, July 1st). The respeaker would follow what was said and keep going. As in regular respeaking, corrections would be permitted in verbatim respeaking. It would be up to the respeaker to respond to the degree of the error, in accordance with the principles of the NER model (Romero-Fresco and Martínez, 2015), and decide whether a correction was justified.

It was during the individual training sessions that the respeakers began to practice verbatim respeaking in earnest and switch between styles. They were, however, asked to try respeaking verbatim at home after the first training session to develop the style and flow independently of Dragon or Text on Top.

The respeakers were able to adopt this style while training, but found it more laborious. Even through their muffled microphones, I was able to detect the change in pace, volume and effort that was required to maintain this respeaking style. Like the participants in the online questionnaire, the respeakers had previously been unphased by the length of time they might respeak for; however, using verbatim as a style prompted queries over how long they would be expected to maintain it for (Training Log, August 14th). I noticed during the training that all the respeakers opted for durations at the shorter end of the scale as they gained confidence in the new set-up. When respeaking verbatim, these blocks were further shortened during training to under ten minutes.

I had hoped that getting audience feedback on their reception of respoken subtitles created in a verbatim style compared to a natural, standard, respeaking style would be a key area to explore at the research events, as I envisaged it contributing to the discussion of respeaking quality in terms of both experience and service provided. However, it quickly became apparent that adding this dynamic to what was already a new setting and set-up would be challenging for the respeakers and also impact on how the events were organised; in order to be able to get contrastive audience responses on each style, those events where both styles were trialled would require an interval to allow the audience to complete an additional questionnaire. The need for an interval ruled out testing verbatim respeaking at the museum tour in Round One (event 4), though we planned to test it at events 1 to 3. However, technical difficulties at the start of the first event (section 7.4.1.3 below) meant that we were only able to test it at events 2 and 3. During the post-event discussion at event 3, I revealed to the audience that we had changed style across the two halves of the event. The feedback from the group was that they had been unaware of this change in style (section 7.4.3.6). As a result of this feedback, and knowing that a range of variables were already being explored at the events, I decided to stop the verbatim-standard comparison at this point in the research. The qualitative feedback that I had gained by this point was useful and it seemed more appropriate to focus on refining the systems in place in order to provide better access than to explore a style which we were ultimately likely to find fault with and reject.

As a footnote to this discussion of styles, the risks of using any label (Chapter Two above) must be remembered. I have been talking about ‘verbatim’, ‘standard’ and ‘edited’ as distinct styles, however describing them as respeaking tendencies may perhaps be more accurate. In usual practice, it is not so much that a respeaker chooses one style to use in isolation. What is more likely, is that a general approach will be chosen, but a combination of styles will be seen in any stretch of respeaking, where the respeakers shift continuously between verbatim and standard, making slight changes to the original when pace, content or sentence structure demand it, but avoiding doing so when they are able to. This is how I would describe my own experience of respeaking, and this seems to reflect the views of those respeakers who took part in the online questionnaire (section 5.3.6.5 above). Fully edited subtitles are rare and undesirable for the respeaker as well as for many members of the audience; it is likely that they will only be used in circumstances where a respeaker is working under severe constraints.

6.7 Live event logistics

In order to prepare for the live event setting, the respeakers needed to master a number of techniques which differed from their regular practice.

6.7.1 Sending out respoken text

Whilst in-house software automatically pulls the respoken content from the subtitling software and sends it to air, when using Text on Top presentation software, the respeakers must press Enter to cue out the subtitles manually. During training, the respeakers explored the visual effect and impact on latency of cueing out different lengths of text, ranging from one or two words at a time to one or more sentences and in doing so developed a routine that was comfortable for them.

Most settled on an approach where they hit Enter regularly, thus cueing out five or six words at a time, in line with the optimum “speech-to-pause” rhythm suggested for respeaking in Romero-Fresco (2011: 107), a process which meant that words continued to scroll out for the audience. The visual unity of a subtitle block was not a requirement, but Enter was often pressed immediately after a punctuation mark was voiced.

6.7.2 Handing over

Achieving a smooth handover procedure when the active respeaker alternates to resting is essential for good access provision; care must be taken to send out all words that have been uttered before control is given to the next respeaker to avoid any content being lost. The procedure required in Text on Top was the reverse of that seen in-house, so much time was spent practising it in order for it to become a natural routine. Since it was so central to the access provision, this became one of the technical checks that the respeakers would perform on the day of an event to ensure that they did not revert to company ways during the live event.

6.7.3 Cueing scripts

Whilst the focus of the research was on respeaking live, unscripted content, some time was devoted to cueing scripts. The Ofcom study on the quality of live subtitling, highlighted the benefits that hybrid subtitling, where cued blocks of subtitles are combined with live scrolling subtitles, can bring (Ofcom, 2015a: 24)⁸⁴, and some audience members also noted a preference for cued subtitles when the content permitted (for example, FGB1, FGB2 and FGB6). The respeakers tested the transition between respeaking and cueing to understand what options they would have if a detailed script for even a part of the event was made available to them.

6.7.4 Saving transcripts

The ability to produce a full transcript of an event is an added bonus that comes with providing live subtitles and the respeakers were given instructions on how to export the transcript from Text on Top. All agreed that following an event, they would want to do a brief review of the transcript, while it was fresh in their minds and make minor adjustments to the text if needed; they would not, however, want there to be an

⁸⁴ At the same time, it was also noted that instances of rapid subtitles have become common during hybrid subtitling (Ofcom, 2015a: 21-22).

expectation on them to carry out major research to improve or correct the transcripts. The respeakers at events 7 and 8 corrected the transcripts they had created and the process took around an hour.

6.7.5 Dragon vocabulary and macros

The respeakers were all impressed with the accuracy of Dragon Professional Individual. Recognition was generally high during the practice sections and, for the most part, the respeakers responded to errors by typing in corrections rather than setting up complex macro systems within Dragon.

Discussion revealed that the respeakers commonly used this approach when respeaking live for television and it was an approach that the typing facility in Text on Top supported. Since they had had only limited time during the training to develop their voice models, the respeakers were selective about which macros they did create and chose not to create extensive lists. However, at work, they had created such lists and developed their voice models across many years. Respeakers A and D referred to instances at work where a particular macro had come up that they wished to copy in their training profile (Training Log, August 6th).

The focus group interviews suggested some variation existed in the processes that different companies used to avoid and correct errors. It is possible, therefore, that respeakers from companies B and C may have adopted a different approach to the macro lists. However, creating macros is a straightforward process - it is simply a question of allocating time for the respeakers to draw up the lists and then enter, train and practice the new macros to ensure they work reliably. A future modification of the respeaker training would be to allocate time specifically for this purpose.

6.7.6 Liaison with presenters

During the full-day session, the respeakers discussed what they would like presenters to know about respeaking, and the ideas shared reflected many of the views expressed in the online questionnaire (section 5.3.6.7 above). Overall, it was felt that adopting general

'good practice' as a presenter would benefit the respeakers. I prepared guidelines for presenters which incorporated this advice, which the respeakers reviewed during the follow-up training sessions, and which were revised again in preparation for the second round of events. A meeting between the presenter(s) and respeakers was also scheduled before each event began, to allow the respeakers to familiarise themselves with the accent and delivery of the presenter and to help acclimatise the presenters to the respeakers, and to what would be involved when being respoken. A more detailed discussion of these guidelines and the response of the presenters to respeaking follows in Chapter Seven, section 7.7 below.

6.8 Research events: Design and purpose

Whilst the training provided the respeakers with an opportunity to practice the skills needed for respeaking in the new setting of live events and develop confidence and familiarity with new equipment, it was only through respeaking at research events that they would have the opportunity to work in a truly live environment and allow the quality of the access being provided in real-time for the audience to be investigated. These events, and the continuing opportunities for learning they offered, in turn revealed refinements which would be needed for future offerings of the respeaker training programme.

6.8.1 Range of events

A total of eight research events were held. The first four events ran in the summer and autumn of 2017 and the second four in the summer of 2018, allowing a period for analysis, reflection and further training between each round. The events lasted between one and two hours each and were held in various locations around the UK. Each contained a substantial unscripted element, so live respeaking, rather than semi-live cueing was used throughout.

Table 6.3: Event overview

Event type	Venue	Respeakers	Software	Working set-up	Theme	Audiovisual content
Round One						
1	Riverhouse Barn, Walton on Thames	A, B	Text on Top	Respeakers located at back of theatre (raised) Performers on flat stage Subtitles at top of screen	Drama, literature, performance	Multiple speakers Live performance PowerPoint with text and images
2	Riverhouse Barn, Walton on Thames	C, D	Text on Top	Respeakers located at back of theatre (raised) Presenter on raised stage Subtitles at bottom of screen	Theatre and arts	Single speaker PowerPoint with text and images
3	BFI, St Stephen Street, London	A, C	Text on Top	Respeakers in separate room Presenters on flat stage Subtitles at top of screen	WW1, censorship, blindness, filmmaking	Multiple speakers Some use of PowerPoint with text and images Videos with burnt in subtitles
4	Wellcome Collection, London	B, D	Streamtext	Presenters located within museum with a mobile phone audio stream from the presenter Subtitles on individual tablets	Science and history	Spoken word and visual reallia
Round Two						
5	The Watershed, Bristol	A, C	Text on Top	Main discussion group located in open café Respeakers located at the table next to the discussion group Multiple lines of subtitles streaming to two different screens	<i>The Piano</i>	None during the discussion
6	The Depot, Lewes	C, D	Text on Top	Respeakers located at side of screening room Presenters sat on flat stage Three lines of subtitles in the middle of the screen	<i>The Piano</i> , access	None during the discussion
7	Manchester Art Gallery	A, D	Streamtext	Respeakers located within the museum with a mobile phone audio stream from the presenter Subtitles on individual tablets	Varied; art- related	Art works around the museum
8	University of Roehampton	B, C	Text on Top	Respeakers located at the back of the lecture theatre (raised) Presenter on flat stage Three lines of subtitles in the middle of the screen	Philosophy, self-help, wellbeing	No slides; use of gestures and action to support content

Audience members were invited to attend as many events as they were able and the events were publicised through organisations such as Stagertext, the NADP and AOHL and shared on social media, including Twitter and Facebook, for example through d/Deaf organisations and local area groups near to where the events were being held. In addition, some of the host venues also included information about the particular events on their websites or in mailings. People who had completed the DH survey and expressed an interest in taking part in the events were also invited.

The scope of the events was broad and they included presentations, public speakers, film panels, museum tours and post-screening Q&As. This design allowed respeaking to be tested across single/multiple speakers, a seated/moving audience, diverse technical set-ups in and outside the event room, and varied visual and spoken subject matter. The variation in the venue, event type and content were planned in advance, but the variations in the working set-up were determined on site according to the specifics of the location.

In addition to these eight events, a conference I attended in Vigo, Spain, provided a further opportunity to test respeaking at a live event (Table 6.4). Sequentially, this conference fell between the second and third research events; however, I consider it outside the main series of events, since the respeakers and audience differed, a new technical set-up was used, and content was respoken both intra- and interlingually. Nevertheless, the experience of the implementation of respeaking, and the respeaking itself, fed into the research events that followed and informed this study.

Table 6.4: Conference in Vigo

Event type	Venue	Respeakers	Software	Working set-up	Theme	Audiovisual content	
Additional event							
*	Conference	A Fundación, Vigo, Spain	ZM and independent	Swiss Text	Split respeaking team. Two located at back of conference room in an interpreting booth (using an open audio feed and headsets), and two working remotely in Switzerland (with a sound feed carried over the internet). Presenters on raised stage. Multiple lines of subtitles displayed on a side screen positioned on the floor to the left of the stage.	Audiovisual translation, access, accessibility	PowerPoint presentations with text and images displayed on a screen on the right-hand side of the stage.

6.8.2 Simulating 'real' events

In the initial research design, the intention was for the audience to experience the feeling of attending a 'real' live event rather than a research event and I hoped that the research element of the event would be as invisible or unobtrusive as possible.

However, in the first round of events this was particularly difficult to achieve as the explorative and trial nature of the research was so evident. Two events, the live arts presentation and the film screening (events 1 and 3), were created specifically to provide an opportunity for testing. For authenticity, we did try to design them in a similar way to other events held at the venues, but at times the research aspect remained strong. Events 2 and 4, the public speaker and museum tour, were re-runs of events that were held regularly and so felt more authentic. Nevertheless, certain modifications were still required to accommodate the data collection. A more detailed account of each event follows in Chapter Seven.

One factor I had not expected was that many people attended because they were interested in taking part in the research, rather than because they wanted to attend a cultural event per se. For this reason, the research elements were for the most part thought of as acceptable; the challenges and delays that were encountered were viewed less favourably, but since it was precisely these problematic elements that we were aiming to overcome through this research, all we could do was acknowledge and discuss these during the post-show discussions.

Nevertheless, in the second round of events, I decided to focus on adding access to existing events and simplifying the process of data collection. By taking this approach, I was able to concentrate on embedding the access smoothly. With the research element less obtrusive, though undeniably present, I could get a better sense of what the reception of respeaking at future, 'authentic' events might be.

6.8.3 Early attempts at integrating access into the event

A central premise behind the design of this research project is that the earlier the future audience is considered, the better the provision of access will be. Embedded within this proactive approach (Greco, 2018; 2019a, 2019b; Chapter Two above) is the need to consider what form of provision this access will take, and in the case of respeaking, how it will be organised. A number of decisions were taken to embed this integration as the events were organised.

6.8.3.1 Publicity and recruiting audience members

The events were open to all who were interested, but advertising was targeted towards DH and language-related organisations to ensure, as far as was possible, that these groups would be represented in the audience at each event.

The call for volunteers was shared by email and in a signed and subtitled video. It explained the aims and purpose of the research, how those interested could volunteer, what would happen on the day and how audience members would be selected. This information was shared widely, through relevant organisations as outlined above (section 6.8.1).

6.8.3.2 Venue and presenter preparation

Before each event, I conducted a site visit for familiarisation with the venue. Decisions were taken about the location of the respeakers, arrangement of the room and subtitle display position. In venues where some in-house equipment was being used, a technical check was run to ensure compatibility with the SmartMics, headsets and presentation software. In the museums, the internet signal was checked along the route of the tour. Information about respeaking and guidelines for presenters were shared with presenters and key venue staff. The presenters were asked to outline the content of the talk (including key themes and terminology) and to list any technical requirements and visual

aids being used. This information was shared with the respeakers around a week before the event.

On the day of the event, I arrived on-site with the respeakers a few hours in advance of the audience. One hour was scheduled for them to train in new vocabulary and to review live event conventions, such as handovers. Technical checks were also scheduled with the venue staff. As mentioned above (section 6.7.7), there was an opportunity for the respeakers and presenters to meet before the event began. As well as, where possible, allowing the presenters to experience being respoken, this was also an opportunity for the respeakers to raise any queries about the content of the event and any (visual) resources used within it in order to provide a more seamless respeaking service.

6.8.3.3 Availability and entry

A frequent criticism made of accessibility is how limited it is. Where accessible screenings are offered, all too often a subtitled film is scheduled during the day on weekdays, when most people, including DH people are at work (Sharma, 2021). Although it was not always possible due to the complex scheduling that was required between all involved, I tried to be aware of this when organising the research events and avoid falling into the same pattern. An effort was made to hold the events at different times and on different days to accommodate varied working schedules and only one was held during the working day. One event was held on a Sunday afternoon and the other six were evening events.

The events in the first round were based around London/Surrey to enable participants to attend multiple events. However, it gradually became clear that the London-centric focus of the events made them inaccessible for many potential participants, in particular many who had completed the questionnaire. I addressed this when organising the second round of events. Here, in order to achieve wider geographical reach, events were organised in Bristol, Lewes and Manchester; only the final event was based in London, thereby allowing participants from Round One to attend and reflect on how the access provided through respeaking had evolved over the course of the events.

Every venue was wheelchair accessible and all audience members were asked in advance of the event about any other access requirements they had. The standard practice at the events was for respeaking to be the main access service provided, so that every audience member could give feedback about the subtitles, but a BSL interpreter was present to provide sign language interpretation during my introduction and the post-event chat. In addition, if a BSL user wanted to ask a question during the event, this could be done through BSL. This mirrored the approach used at the CaptionCue research events (Romero-Fresco and Fryer, 2016).

This access provision differed slightly at events 5 and 6. Event 5 was part of the Deaf Conversation series at Watershed in Bristol. These events are always BSL interpreted, so here the subtitles were an extra layer of access being tested on this particular day. Event 6 was structured in a similar way, and had been advertised as having BSL interpretation, so the entire event was BSL interpreted as well as respoken.

These initial steps were a good starting point towards the integration of respeaking and accessibility within the events, but many more steps were needed, especially with regard to the poietic role of the audience (Section 2.2.3.2 above). These new steps were gradually incorporated event by event as they were recognised; by the end of the study, the need for a clear chain of access (Greco *et al.*, 2012) was fully understood and a working model was created for achieving this, which will be discussed further in Chapter Nine.

6.8.4 Respeaker allocation

The respeakers each respoke at 3-5 events and were rotated through the events (table 6.5). This allowed the respeakers to work with different colleagues and for the experience gained at each person's previous event to be shared in a practical and relevant way.

Table 6.5: Respeaker allocation

Event	Respeakers
1	A, B
2	C, D
3	A, C
4	B, D
5	A, C
6	C, D
7	A, D
8	B, C

6.8.5 Event organisation

Everyone present at the events were asked to complete pre- and post-event questionnaires and invited to share further reflections during the post-event chat and individually in the week that followed the event. A technical analysis of the respeaking was also carried out after each event.

Video and audio recordings were made at each event to facilitate the data collection and analysis and the transcripts of the respeakers' output were saved. At the seated events, cameras were positioned to record the whole event as the audience viewed it, while at the mobile events, a camera followed one tablet throughout the tour to record the subtitles that appeared. Where possible, additional cameras were set up to film the seated audience as they watched and the respeakers' screens as they worked.

This data collection also influenced the structure of the events. Each began with a brief introduction about the research, where the format of the event and the audience questionnaire were explained.

In the first round of events, a short video clip of a live event broadcast on television with respoken subtitles was shown to the audience as a baseline point of comparison for their reception of the event. After viewing it, the audience completed their first questionnaire.

Then, they completed another questionnaire which asked the same questions, following each respeaking style tested. Where only regular respeaking was used (events 1 and 4), the questionnaire came at the end of the event and preceded the post-event chat. Where both styles were tested (events 2 and 3), close-to verbatim/regular respeaking was used in the first half of the event and verbatim was used in the second half; the questionnaires followed each half (in the interval and at the end of the event, before the post-event chat). My aim in asking the audience to complete multiple versions of the same questionnaire was to enable the reception of respeaking to be compared across style and setting. To facilitate this, in my discussions with the presenters at events 2 and 3, we aimed to ensure that similar content, including opportunities for questions, would be included in both halves, and the respeakers shared the respeaking in each half.

During the reflection period which followed the first round of events, I realised that these multiple questionnaires were unlikely to provide the comparative analysis that I had hoped for; with so many variables across the events, it was unlikely that individual causes for any patterns would be identified with certainty. I could also see that multiple questionnaires risked compromising the user experience and accessibility of the event as it became too research heavy, in a way which I could no longer justify. For this reason, in the second round of events, as well as focusing on regular respeaking, the introductory television clip was discontinued.

6.9 Technology used at the events

Ensuring a good audio feed for the respeakers was a vital requirement of the research project, and silent disco headsets proved a solution for this. These headsets were the last piece of kit to be added to the respeakers' individual kits referred to in 6.4.5 above. In order to transmit the voices of the presenter(s) and audience to them, an audio system was required. A dual wireless handheld microphone system was chosen for the flexibility it provided and a Falcon 3-channel transmitter sent the signal from the microphones to the silent disco headsets.

In Round One, VH2 microphones from QTX were used with either a Citronic 2-channel compact mixer or the venues' own systems. This changing set-up proved complex and problematic and it became clear that the success of the event depended on the quality of the microphones.

For this reason, an independent kit was created, consisting of the Falcon transmitter, the Shure SM56 wireless microphone combination system, and a Xenyx X1222USB mixer and microphone stand. A TX208 Alto speaker allowed amplification of voices in larger venues. The respeakers all agreed this improved the quality of the audio feed.

This was a complex kit, so in order to ensure its smooth and speedy set-up from event to event, I purchased coloured cables (red, green, blue and yellow) and used stickers of matching colours to indicate the ports where each cable needed to be connected. This simple technique was very effective and vastly simplified the technical aspect of the events. Figures 6.4 and 6.5 illustrate how the kit was assembled at stationary and mobile events.

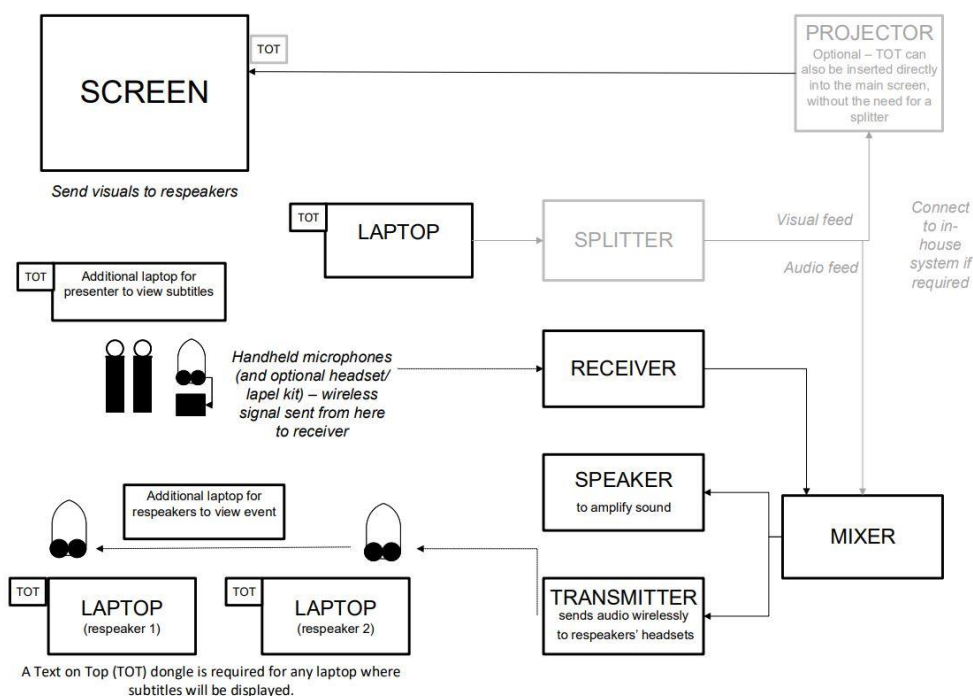


Fig. 6.4: Technical set-up at stationary events⁸⁵

⁸⁵ Elements in grey are optional additions to the core set-up.

The presenter communicates with the respeakers via mobile phone; a headset or ear piece can be used as desired. Only the words of the presenter will be heard, so any comments or questions from the audience must be repeated.

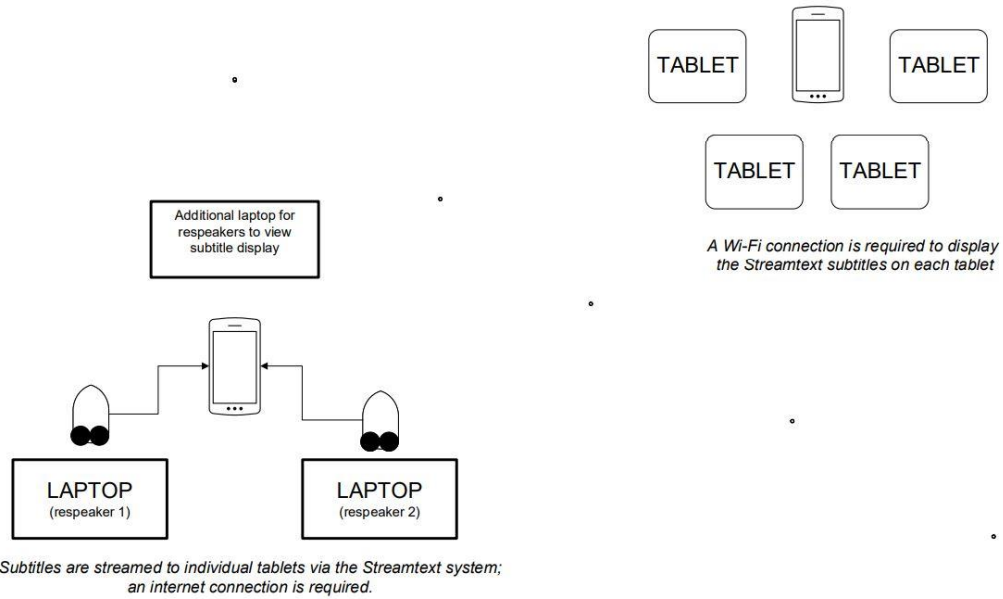


Fig. 6.5: Technical set-up at mobile events

6.10 Style guidelines for live events

The style guidelines for respeaking at live events were based on the preferences expressed in the focus groups (section 5.3.3.4 above) and finalised in discussion with the respeakers during the initial training. Table 6.6 sets out the conventions and compares them to those used in live subtitles at Stagertext's talks and tours, at Ai-Media educational events, in theatre captions produced by Stagertext and in live subtitles on television.

Table 6.6: Suggested conventions for live events, compared to other forms of subtitles/captions⁸⁶

Feature / function	Focus group preferences	Respoken live subtitles at events	Stagetext's talks and tours	AiMedia's educational lectures	Stagetext's theatre captions	Television, using the BBC as an example
Display mode	Scrolling word-by-word	Scrolling word-by-word, or a few words at a time	Scrolling word-by-word		Scrolling lines	Scrolling word-by-word
Font	Arial	Arial			Arial	
Subtitle position	Bottom of screen on TV, both sides of the stage at the theatre	Variable	Variable		Variable, but positioned close to stage where possible	Usually at the bottom of the screen, sometimes at the top
Number of lines	2 lines	2 lines if other content is displayed on screen	2 or 3 lines if other content is displayed on screen	Multiple	3 lines	2 lines
Line spacing for live text	If multiple lines, use a line break between speakers and paragraphs	Otherwise, variable if multiple lines, use a line break between speakers and paragraphs	Otherwise, variable if multiple lines, include a line break between speakers and paragraphs	Single line line break after each speaker name Two-line breaks between paragraphs	Line breaks may be used around sounds	Not applicable
Corrections in live text		Use double dash to add corrections to text already displayed on screen	Use double dash to add corrections to text already displayed on screen		Not applicable	Use double dash to add corrections to text already displayed on screen
Colour contrast for main speaker Speaker ID	No preference	Yellow on black	Yellow on black	White on black	Yellow/orange on black	White/yellow on black
Manner	Name tags	PETE: NEW SPEAKER: FROM THE FLOOR:	Pete: Sometimes capitals may be used	PETE: What he says begins on the next line like this QUESTION FROM FLOOR: COMMENT FROM FLOOR:	PETE: The next word appears two spaces after the colon like this	Change in colour (white, yellow, cyan, green)
	No preference	Indicate when needed, could use symbols ! ? / or (AGGRESSIVELY)			No interpretation of how lines are delivered, but [WHISPER] or [SOFTLY] may be used Capitals (for shouting) or italics (for emphasis or anger) may be used sparingly	Not usually indicated in live subtitles

⁸⁶ This table has been compiled from existing guidelines (Stagetext, 2018; BBC, 2021; Ai-Media, no date a) and images of shows, talks and programmes available online (for example www.stagetext.org and mycleartext.com). Where a feature is greyed out, no information was available.

Feature / function	Focus group preferences	Respoken live subtitles at events	Stagetext's talks and tours	AiMedia's educational lectures	Stagetext's theatre captions	Television, using the BBC as an example
Language	Indicate accents or language spoken; represent accurately	(SPEAKS OWN LANGUAGE) (ALL SPEAK AT ONCE)	Stagetext's policy is to subtitle verbatim, or as close to verbatim as possible	If a speaker has broken grammar, and the respeaker understands it well enough to be able to, they should improve the grammar for readability and to reduce Dragon errors	[SCOTS] [AS PETE]	IN JAPANESE:
Swearing	Represent (all content) accurately	Include	Include		Include	Depends on time of broadcast
Access labels in live subtitles	Wanted this information	(INAUDIBLE) (PLEASE REPEAT THAT) (?) to mark a tentative spelling ... for omitted content	[inaudible] [please repeat that] (?) to mark a tentative spelling	** Audio lost **	Not applicable	Not applicable
Sound labels	Either no preference over presentation / styled as (he laughs) Where applicable, indicate source of sound	(APPLAUSE) (applause)	[applause]	(Applause) (Multiple speakers) (Video plays) (Reads)	VOICES DROWNED MUFFLED VOICES Mid-dialogue, left justified, on its own line: [APPLAUSE] [SCHOOL BELL] If isolated, centred and in the middle of the display, with no brackets: DOORBELL RINGS	i.e. APPLAUSE May be positioned on a separate line from the spoken content Some variation in style seen across channels
Music	Lyrics Include information about the music	Lyrics Include information about the music	Lyrics Information about the music		Lyrics Information about the music	# Lyrics Information about the music

Since the events were being held in the UK, British English spelling was specified in the guidelines, something which would already be standard practice for all the respeakers involved. The issue of swearing had come up a number of times in the DH and respeaker focus groups. DH audience members were concerned they would not get full access to the event if swearing was omitted (OL19, question F2); the discussion amongst respeakers was about how easily swearwords might unintentionally be produced by Dragon without the in-house filters that usually prevent this from happening. The general consensus amongst respeakers familiar with creating company house styles was that was that swearwords were unlikely to appear unintentionally; this meant that no further profile modifications would be needed to prevent them. If a presenter or audience member did swear, the guideline was to include this in the live subtitles.

The subtitles were set to scroll across the screen as new words appeared. Where the subtitles appeared on the screen with other content, two lines were displayed, but where they were the only content on screen, multiple lines were displayed. Exactly where the subtitles were positioned and displayed varied from event to event, according to the physical space.



Fig. 6.6: Subtitle position at event 2

During the museum tours, where the audience had individual tablets, we pre-set the subtitles to appear in a yellow font on a black screen (choice e) in the figure below);

however, audience members could personalise this display. Figure 6.7 illustrates some of their choices.

- a) Some chose subtitles like this,
- b) and others preferred this.
- c) One person chose lime green because it was different.
- d) Some people used this font,
- e) but we set up the tablets like this.

Fig. 6.7: Examples of personalised tablet displays

In the Text on Top software, the respeakers were able to type in corrections before the text was displayed on screen, whereas on Streamtext, the output appeared on screen automatically, but could still be corrected; in this instance, the audience would see the letters being deleted and replaced. A double dash, - - could also be used to indicate a delayed correction.

Very few sound and manner labels were used at the live events. A few macros were created, including (APPLAUSE), (LAUGHTER), (SPEAKS DIFFERENT LANGUAGE) and (ALL SPEAK AT ONCE). As the need for each label arose, the most appropriate terminology was discussed and consideration was given to register, legibility and number of characters each would require. A number of access labels were also needed. '...' would be used to indicate omitted content. In line with Stagertext guidelines, (INAUDIBLE) would be used to highlight any audio problems and (PLEASE REPEAT THAT) was also included. If the respeaker wanted to enter a tentative spelling, they would indicate it with (?).

Whilst the audience had not expressed an overall preference over the use of upper- or lower-case letters, or demarcation with brackets in the case of labels, many television

channels use upper-case letters, as do Stagertext (Stagertext, 2013), so I adopted this practice, and decided to use regular brackets, as there was a slight preference for them in some audience groups. Since this practice does vary across the different settings, lower-case letters could also be used, though they would need to be demarcated with brackets to separate them from the surrounding dialogue. What is most important is to ensure that a standard approach is being used by both respeakers throughout the event. At events, name tags rather than a change of colour must be used to indicate speaker ID, both because this is the quicker option in Text on Top and because colours are lost on the personalisable Streamtext software. Once again, this matches current practice at live events. Whenever possible, the name of the speaker would be used, but when necessary the label NEW SPEAKER: would be used to identify a change in speaker within the panel and FROM THE FLOOR: would be used if an audience member spoke.

6.11 Updates to the training programme

Following the principles of action research, the conceptual design of the training programme evolved as the project proceeded, allowing refinements to be made as more was learned about the training that respeakers would need in this new setting.

6.11.1 Respeaker training Round Two

One additional day of training was organised before the second round of research events began. This was held in Walton on Thames. Two respeakers were able to attend and the other two received updates and a shorter version of the training on the first of the events they attended in Round Two.

The training covered three key areas. Firstly, this was an opportunity to update the respeakers on the feedback received from the first round of events, so that they could continue to act as informed participants. Next, there was a discussion on the wider role of the respeaker at live events. The respeakers were able to share their own reflections on the events and I outlined the role of the designated access co-ordinator at events and the liaison that might be needed between them and the respeakers. Finally, the new

equipment was introduced. With the help of the Technical Manager at the Riverhouse Barn, the host venue for events 1 and 2, we were all trained on the basics of the sound system, outlined in 6.9 above, and how to set it up.

6.11.2 Future offerings of the training

In order for this bespoke training programme to be used independently of this study, a few modifications will be needed to the modules already in place, and some new content in the form of additional modules would also be beneficial. Table 6.7 summarises the original content of the eight modules included in the initial respeaker training programme of Cycle One (the light blue cells) and explains the suggested adaptations to them and the additional recommended modules (light orange).

In addition, two further modules are suggested for other people involved in respeaking at live events (light green). Whilst the latter are not training for respeakers per se, it would benefit the respeakers to be broadly aware of the content, especially as they may themselves be involved in providing some of the training. This training for presenters and audience members is explored further in Chapter Seven.

Table 6.7: Revised respeaker training programme

Module	Module type	Summary of original content	Suggested adaptations
Group training			
1	Project overview (theoretical)	<ul style="list-style-type: none"> Need for the research project Theory behind it 	<ul style="list-style-type: none"> Why introduce respeaking at events? Overview of what access currently exists and where respeaking could fit Refer to the three shifts of AS Supplemented with feedback received at live events Benefits for wider audience groups
2	Expectations at live events (theoretical)	<ul style="list-style-type: none"> Focus groups: key needs and requirements 	<ul style="list-style-type: none"> Extended training on Streamtext
3	Equipment (practical)	<ul style="list-style-type: none"> Introduction to SmartMics and subtitling software 	<ul style="list-style-type: none"> Supplemented with visuals and case studies from the research events
4	Practicalities of live events (practical)	<ul style="list-style-type: none"> Range of events and venues Watching respeaking in action on the new software 	<ul style="list-style-type: none"> Is this verbatim? Looking at transcripts and video clips sourced from live events Introduction to this process What kinds of edits should I make and when? Clips from research events Live editing
5	Respeaking style (theoretical)	<ul style="list-style-type: none"> Defining the regular style 	<ul style="list-style-type: none"> Introduction to the NERLE Awareness of additional interactions and considerations at live events Discussion of subtitle introductions
6	Live editing (practical)		<ul style="list-style-type: none"> Sounds and transmitter equipment Setting up on location Variation across venues Connecting remotely
7	Practice (Opportunities throughout course) (practical); trainee-led	<ul style="list-style-type: none"> Clips sourced online 	<ul style="list-style-type: none"> Text on Top Streamtext Sound equipment
8	Assessing respeaking at live events (theoretical)		<ul style="list-style-type: none"> Exploring the extended role of the respeaker at live events What information do venue teams, presenters and audience members need? Role-playing scenarios What other considerations are needed if an event is to be accessible? Extend to include new topics
9	Setting up on the move (practical)		
10	Technical troubleshooting (practical)		
11	Role of the accessibility co-ordinator (theoretical & practical)		
12	Access at live events (theoretical)		
13	Review (Opportunities throughout course) (practical)	<ul style="list-style-type: none"> Summing up and questions 	
Individual sessions			
14	Personalised learning (practical)	<ul style="list-style-type: none"> Continuation from the practice module above 	<ul style="list-style-type: none"> Added familiarity with the software
Training for others involved in respeaking at live events			
<p><i>The respeakers would not receive this training, but may be involved in providing it. General awareness of this content would come into the module on the Role of accessibility co-ordinator above. This will be explored further in Chapter Seven.</i></p>			
Training for presenters (practical)	Wider training		<ul style="list-style-type: none"> Know your audience Being respooken Communicating with the respeakers How to handle different types of audio and visual content Preparing for the event

6.11.2.1 Refining modules from the initial round of respeaker training

The modules from the initial round of respeaker training provided the respeakers with a solid base of knowledge and understanding to begin working at live events. In future offerings of this course, I would want to draw directly on the experience that the respeakers gained. Resources from the research events, such as photographs, video footage, reflections from the respeakers, case study accounts and event transcripts could all be used to supplement the existing modules. They would demonstrate the practicalities of respeaking at live events more authentically and be a good basis for the practical module five.

Module one, the project overview, would benefit from being updated within an overview of the feedback received from all the events, and a revised account of the three shifts now identified in accessibility studies, rather than Universal Design which I originally referred to. Module two could address the now better understood needs of the wider audience.

It is module five, on respeaking style, that would undergo the most significant change. Where during the initial programme, some of the discussion on respeaking style related to whether we could trial verbatim respeaking at the events, in future applications of the course, the core content of the module would be on the difference between the styles and how, in practice, all are used. The discussion could be more focused on which situations might call for verbatim and edited aspects of subtitling, and the importance of personalising the content as per clients' requests (when appropriate). Examples from the research events would be very fruitful here.

6.11.2.2 New modules

A new module (six) on live editing would be a very useful addition to this training course, so that the non-live respeaker could edit the live subtitles produced by their colleague; this would complement the updated module five well. This new module would cover the practicalities of this process, advice when preparing transcripts and a chance to see the

original respoken text alongside the tidied transcript. Additional practical time would be needed once the respeakers were familiar with the software to allow the respeakers to try live editing in pairs. Similarly, a module on how to assess respeaking at live events (eight) would allow the respeakers to look more closely at the additional interaction required at live events, compared to television work, and would introduce them to the NER for Live Events, or NERLE model, which I propose in Chapter Eight for assessing the quality of respeaking in this new setting.

The top up training that the respeakers received at the start of Cycle Three forms the basis of modules nine, Setting up on the move, and ten, Technical troubleshooting. In the first, respeakers would learn how to put together the audio kit and in the second, respeakers would be presented with a range of challenges which may occur at live events and together discuss how they could be resolved.

In this way, respeaking-related training would help the trainees acquire the additional skills and understanding they would need to work with and step into the role of designated access co-ordinator (module eleven). Whilst it is not likely they would need to take on this role in its entirety, it is very likely that respeakers would need to understand and advise on it. Further discussion of this follows in sections 8.6 and 9.3 below. Module twelve draws this awareness into an even broader discussion and exploration about how the access provided through respeaking can be integrated into the access of the event as a whole.

6.11.2.3 Innovativeness of the design

With these additional modules, and the understanding I gained through the action research design, I believe that the updated training programme reflects the aims and intentions I began with. This programme would provide an opportunity for respeakers to work in a way where they are fully integrated into the event process; something that was a hope at the start of the project, now has a pathway for being realised. In television settings, steps have already been taken towards this; news scripts and outlines are shared, allowing segments of programmes such as the news to be live cued. This was

one of the changes which resulted from the Ofcom Study (no date), though much post-production of access remains, as evidenced when news scripts arrive late, when programme details are not shared and when programmes that should have pre-recorded subtitles arrive so late that subtitling them live is the only option.

Where the initial respeaker programme allowed the respeakers to be informed participants in this study, this revised version should enable respeakers to expand their role from being 'simply' respeakers, already an impressive feat, to being fully involved in the access provision of the event in question. In this way, the community of practice expands from respeaking-related practice, to the practice of access provision more broadly.

6.12 Resources for respeaking at live events

In the course of this chapter, key resources for respeaking at live events have been presented. I have outlined in detail the training that was provided for the respeakers in this study, who joined me as informed participants in this community of practice and contributed so greatly to the research, and illustrated how, in line with action research, it has evolved into the form that I suggest be used for future respeaker training. Stylistic guidelines for live subtitles have been compiled; their component features have been traced back to the focus group research of Chapter Five, and aligned with the guidelines for subtitling in other settings, including captions produced by Stagertext, live subtitles produced by STTR at events and at Ai Educational lectures, and in subtitles seen on television. Finally, I have also given details of the technical set-up and specialised equipment that I would recommend using at live events.

Alongside this, I have presented the conceptual design of the research events that were interspersed between the two rounds of respeaker training, and which also contributed to the development of these resources.

In the next two chapters, I analyse these events and the data collected during them more closely. This analysis will both explain the recommendations I have made in this chapter

and allow the experiences of those present at the events (Chapter Seven) and the respeaking carried out during them (Chapter Eight) to be examined more closely.

Chapter 7: Reception study at the live events

“Never in the world were two opinions identical,
no more than any two hairs or specks of grain were ever exactly alike.
What unites them the most is the difference that exists between them.”

Montaigne II, 37, 611⁸⁷

In Chapter Six, the respeaker training programme and the research events which followed and provided an opportunity to put this training into action, were presented. The consequent changes to the programme and to the procedures and set-up used at the events were also explained. While this provided an initial answer to the question of what training and technical set-up is needed to allow experienced television respeakers to transfer their skills to the live event setting, it is only by exploring the responses of those present at the events and looking more closely at the quality of the respeaking seen there that these can be answered more fully. In the current chapter, the first is done.

Much of the focus will be on the responses of the audience to the events, yet consideration is also given to the experiences of others who were involved, the venues, respeakers and presenters, in this extended interpretation of a reception study. In this chapter, the presenters represent a new focus group. The research events marked the first time these presenters were being respoken and their ability to continue presenting with this new addition to their usual working environment is a vital aspect that must be considered in the integration of respeaking into live events.

By examining the events from the perspective of those who attended and were involved, as well as determining the quality of service that the audience received, there is an opportunity to explore the process of the proactive integration of access at events and the nature of the interaction that would be required for this to be successful.

⁸⁷ This is my own translation of “Et ne fut jamais au monde deux opinions pareilles, non plus que deux poils ou deux grains. Leur plus universelle qualité, c’est la diversité”.

7.1 A recap - event design and procedures

The eight research events lasted between one and two hours each and were held in various locations around the UK. The scope of the events was broad and they included presentations, public speakers, film panels, museum tours and post-screening Q&As. This design allowed respeaking to be tested across single/multiple speakers, a seated/moving audience, diverse technical set-ups in and outside the event room and varied visual and spoken subject matter as well as in a variety of venues.

As people arrived at each event, they were asked to sign a consent form, which included optional photography and video consent. After a brief introduction to the study, I explained how each event would unfold and that the audience would be asked to share their experience of and response to the event through a questionnaire. Following this, the event proper began. At most events, I was able to step back and observe the event, and step in only occasionally in the role of designated access co-ordinator, for example to pass on information about the subtitles to the presenters, or to deal with a technical issue. At other events, however, my involvement was more hands-on.

At the end of the event – and at the start of the start of the interval where there was one – I gave brief instructions about how to complete the questionnaire. In addition, I led the focus group that followed each event.

The focus groups, taking the form of post-event chats, lasted around 30 minutes each and everyone present, including the presenters and venue representatives, were invited to attend. The respeakers also contributed to the discussion, either respeaking their thoughts directly, or being respoken by their colleague.

7.2 Data from those attending the events

7.2.1 Format

Everyone who attended the event was asked to share their experience in taking part. The respeakers, venues and presenters were given short questionnaires to complete before and after the event, designed to gain a snapshot of their experience (appendices 7.1.1-7.3.2); all were able to share longer reflections if they wished.

The data collected from the audience was more detailed. In addition to a demographic questionnaire (appendix 7.4.1), the audience received a questionnaire at the end of the event⁸⁸ (appendix 7.4.2), which asked them to rank different features of the respoken subtitles and the access they provided through a series of Likert scales (section 7.2.2 below). A number of open questions were also included, where they could respond more freely. The audience were also invited to share a reflective response in the week that followed each event (appendix 7.4.2), in writing or in the form of an audio or video recording. Further qualitative data was also collected during the post-event focus groups. The audience data was analysed in SPSS and the approach taken to the data was a descriptive one, designed to present quantitative findings in a manageable form (Trochim, 2021). This use of descriptive statistics allowed me to simplify the large amount of data collected in a sensible way, and summarise the patterns seen in the data more clearly; it is also one that complements the qualitative data collected at the events, and the approach taken when analysing the data in Cycle One.

The very small groups of people across the events when different hearing statuses and languages were considered, the fact that some people attended more than one event and logged multiple responses, and the inherent diversity seen across the events meant that the collated responses to the questionnaires were inappropriate for real-world data tests. Even non-parametric tests, which can often offer a solution to small sample sizes were inappropriate here. Although this did dictate what approach I could adopt, it did not impact negatively on the analysis.

Citing Zimmerman (2010), Romero-Fresco (2021a: 295) writes, “marginal participants are often classified as outliers and their data are routinely discarded in empirical research as problematic, unrepresentative and useless”. Given the desire in this study and in media accessibility more broadly to account for people who are often “excluded from accessing audiovisual media” (Díaz Cintas *et al.*, 2007 in Romero-Fresco, 2021a), in other words a desire to step away from the implicit focus on societal norms (Chapter

⁸⁸ As explained in Chapter Six, the questionnaires that were completed about the clip and attempts at verbatim respoken are not included in this discussion.

Two), taking this opportunity to focus on everyone, including people who may elsewhere be considered “outliers” can only be a positive step. The approach adopted in the current study may therefore be considered as an example of what Romero-Fresco suggests (p.295) when he states that “it should be possible for researchers in this area to engage with these outliers meaningfully in a way that complements the quantitative and statistically significant data obtained with experimental reception studies”.

7.2.2 Working with the scales

A total of 22 scales were included in the questionnaire. They are categorised under areas of interest in table 7.1 below, as determined from the analysis of the audience focus group data.

While there were many areas of interest that I would have liked to have explored further, with more varied statements, this set allowed the features I most needed to explore in order to compile guidelines to be studied in detail, without the questionnaire becoming too long for the audience to complete. In particular, I would have liked to explore audience engagement further, but I realised that given the range of events being tested, this would be difficult to capture in a single Likert statement⁸⁹. Instead, I encouraged the audience to comment on this more freely after the event in the hope I would gain some insights into it.

I did explore whether I was statistically able to cluster any of these statements together during the analysis, using a Cronbach alpha analysis, with a view to producing a single score, for example, for the area of ‘subtitle content’ or ‘environment’, to simplify the breadth of the data. Whilst I was able to do this and obtain a single score for the two statements on ‘Access’, I was unable to do so for any other areas of interest. A full explanation is included in appendix 7.5.

⁸⁹ In the feedback from the event in Vigo, one participant commented that ‘feeling immersed’ and ‘being connected’ with the speaker may not necessarily apply in an academic setting, however interesting the presentation.

Table 7.1: Likert scales from the audience questionnaire

Area of interest	Original Statement
Access	The subtitles gave me extra information (that I would not have heard or understood)
Access	The subtitles improved my access to the event
Understanding	I understood the content of the event
Attention	It was possible to follow both the subtitles and the visual content of the event
Attention	It took effort to follow the subtitles
Attention	I found myself reading the subtitles even though I didn't want to
Attention	I spent too much time looking at the subtitles
Physical effort	I experienced sensations like a headache or tiredness
Physical effort	Following the subtitles and visual content involved too much head/neck movement
Interest and engagement	I found the event interesting and informative
Interest and engagement	I felt immersed in the event
Interest and engagement	I was able to connect with the speakers
Environment	I had a clear view of what was going on
Environment	The brightness of the screen allowed me to follow the subtitles easily
Subtitle Content	The subtitles were accurate (spelling and punctuation)
Subtitle Content	When there were several people at the front, it was easy to know who was speaking from the subtitles
Subtitle Content	The subtitles captured the main content of the event
Subtitle Presentation	The subtitles were readable in terms of format (size, font, colour)
Subtitle Presentation	The subtitles were readable in terms of display mode (lines, scrolling)
Subtitle Presentation	The subtitles took up the right amount/proportion of space on the screen
Timing	The subtitles appeared at the same time as people spoke
Timing	The subtitles disappeared before I could read them

It was striking that even when a particular cluster of features (i.e. an area of interest) scored similarly and had internal consistency in one event, they did not necessarily have it in all the other events. For example, I tried to group the two environmental features together thematically:

Environment	I had a clear view of what was going on
Environment	The brightness of the screen allowed me to follow the subtitles easily

For some events, the scores for the two statements were internally consistent, but for others they varied greatly. This highlights the diversity of environmental conditions at live events. Certain venues may have good lighting, but the position chosen for the subtitles may not be satisfactory. This reflects the comments made by audience members in Chapter Five, where, when asked about subtitle legibility, many cited other factors beyond the size and style of the font/typeface, which might affect how easy-to-read the subtitles were.

The case of the physical effort required to read subtitles may seem more straightforward and a close correlation between 'Following the subtitles and visual content involved too much head/neck movement' and 'I experienced sensations like a headache or tiredness' might be expected. However, there may be situations where an audience member is seated in a position where only minimal head or neck movement is required to view both the subtitles and visual content, yet sensations such a headache or tiredness are still experienced. This might be due to the subtitles themselves, to the event content or even the health or mood of the person attending on a particular day. In other words, while these factors are thematically related in that they refer to physical effort, that does not guarantee that the scores individuals give them will be consistent even within a single event.

Some variation could also be due to the heterogenous nature of the audience, the dynamic and diverse nature of events or other factors entirely, but it highlights the need for careful consideration to be given to how respeaking, and access more broadly, is integrated into any event. While some general principles can certainly be applied, consideration of the particular circumstances of each event will also be needed and the use of respeaking, as with all forms of access provision, will need to be tailored both to the event in question and to those attending.

I have tried to capture this notion in the analyses of the events that follow, and it is also explored further in the questions and considerations when making live events accessible using respeaking (appendix 9.1) and in the model for participatory engagement suggested in Chapter Nine (section 9.3). Both highlight the fact that a one-size-fits-all

approach to access cannot be expected either across events or for every person at a single event. Providing access is not a 'checkbox' activity, but rather it requires thoughtful discussion and reflection to ensure that it has been implemented in the best way possible in a given situation.

Since there was so much variation across events, both in terms of the event type, audiovisual content, procedural changes which resulted from the action research process and with regard to the factors captured in the individual scales, it was not possible to isolate with certainty the impact of any individual subtitle feature within the data. This further supported the case for the use of descriptive statistics to accompany the broader qualitative reports on each event.

I compiled the scores for each statement, so that I had a mean event score for each statement, which ranged from 1-5 out of five, and a total event score, which I will discuss as a percentage score. In addition, I calculated the mean score given at each event, so that I could see how the scores for individual statements were ranked compared to it.

I also calculated these scores for different audience groups, according to the hearing status of those attending (Deaf BSL / deaf / hard of hearing / hearing, but/ hearing) and according to their native language (English / BSL / Other European / Other non-European / Bilingual with BSL / Bilingual without BSL). These were small events, and in many cases the number of people who identified as being in each individual group was very small. While the qualitative report that follows does discuss how responses across these groups vary, had a more standard statistical test been applied, the awareness and impact of this small group size might have been masked. A full explanation of how this data was compiled is included in appendix 7.5.

The analysis of the events that follows begins with an overview of the eight events, highlighting which were scored more and less favourably by the audience; then the data for each event in turn is considered. Here, I draw together feedback from all parties from the questionnaires, reflective comments and any event-specific discussions from the focus groups. Finally, I introduce the wider considerations that were shared in the focus

groups in the style of a reflexive thematic analysis in section 7.6. In this way, I share how the events were perceived by all who were involved.

7.3 Overview of the audience

7.3.1 Who attended the events?

In total, 81 different people attended the research events and formally took part in the research by sharing their feedback; a few people attended events 5, 6 and 7 without completing a questionnaire. The average attendance at each was 14 people, and 110 responses were received altogether, as sixteen people attended more than one event, and one person attended them all.

The audiences had mixed hearing levels and a range of native languages, including BSL and a variety of European and non-European spoken languages. Out of those who replied, 15 people (18.5%) used either a hearing aid or CI to communicate, 5 people (6.2%) used one along with lipreading, and a further 5 people (6.2%) used BSL along with lipreading and either a hearing aid or CI.

Their ages ranged from 13-88; 24 people (29.6%) were male, 56 (69.1%) female and one (1.2%) identified as non-binary. 38 people had completed degrees or some kind of higher education, and 39 had completed postgraduate studies. 24 (29.6%) people watched less than one hour of television a day, whilst 28 (34.6%), the median score, watched 1-2 hours a day. 45 people (55.6%) said they had no difficulty reading or understanding subtitles on television, whilst 30 (37.0%) people said they had difficulty sometimes. Only two people (2.5%) said they did have difficulty.

7.3.2 Overview of the audience feedback

Much variation can be seen in the scores given by the different groups, both within and across the events. Table 7.2 draws together the scores assigned by each hearing and language group; in addition, a combined score from DH audience members, the 'expert users' as I called them in Chapter Five, is also shown, along with the combined audience score of everyone present.

Table 7.2: Overview of the audience scores at each event by group⁹⁰

	1	2	3	4	5	6	7	8
Deaf BSL	48.1%	88.6%		82.9%	75.2%	72.4%	69.5%	
deaf (Deafened, deafblind)	76.6%	85.8%	77.1%	79.1%	82.9%	87.6%	81.0%	81.9%
Hard of hearing	72.4%	81.6%	79.1%	76.5%			88.8%	87.4%
Hearing, but				65.7%	86.7%	90.5%	79.5%	
DH Combined	72.2%	84.2%	78.1%	77.2%	82.7%	84.5%	81.0%	84.0%
Hearing	71.7%	79.5%	75.0%	79.5%	69.7%	78.5%	80.6%	73.4%
English	74.6%	83.9%	78.0%	77.2%	71.9%	80.7%	82.4%	82.5%
BSL	48.1%				75.2%	72.4%		
Other European		79.5%		80.5%	86.7%	84.0%		81.7%
Other non-European			74.1%				70.7%	61.0%
Bilingual (with BSL)		88.6%						
Bilingual (not with BSL)						80.0%		
Total	72.1%	83.8%	77.3%	77.6%	73.9%	80.5%	80.8%	78.2%

These variations, and details within the scoring, will be considered more closely below, alongside the comments made by those who attended the events. Nevertheless, despite this variation, some general patterns can be seen within the scores, which allows us to consider which events the audience responded to most and least favourably.

Table 7.3 provides a simplified view of this data, showing the DH combined score and the Total score for each event in ranked order.

⁹⁰ A greyed-out block indicates that no-one identifying as belonging to that hearing or language status attended the event.

Table 7.3: Events ranked by audience preference

Rank	Audience total	DH Combined	Key	
1 st	Event 2	Event 6		Above 80%
2 nd	Event 7	Event 2		75% - 79.9%
3 rd	Event 6	Event 8		Below 75%
4 th	Event 8	Event 5		
5 th	Event 4	Event 7		
6 th	Event 3	Event 3		
7 th	Event 5	Event 4		
8 th	Event 1	Event 1		

The scores from these two audience combinations matched for 6 events (table 7.4). Events 2, 6 and 7 are the highest-scoring events, each with a score above 80% (and highlighted in green). Events 3 and 4 received mid-range scores, between 75%-79.9% (highlighted in yellow) and event 1 was the lowest scoring event, receiving less than 75% (highlighted in red).

Table 7.4: Events with matching DH combined and total audience scores

Score range	Event / Event Type		
Highest scoring events (80%+)	2 (Public speaker)	6 (Q&A)	7 (Tour)
Mid-range events (75% - 75.9%)	3 (Mixed event)	4 (Tour)	
Lowest scoring events (<75%)	1 (Presentation)		

Considering the ranking in this way is useful for getting a feel for how the audience present responded to the live subtitles and access provided, even whilst remembering that, with all the variations between events, an absolutely direct comparison is difficult to achieve. It is, however, worth noting that the ranking does not appear to be directly related to the event type (with the two tours, for example, receiving medium and high

scores), nor solely related to when the event took place sequentially (with the second event of the series being amongst the highest scoring).

The final two events received conflicting scores from the combined DH group and the audience as a whole. Event 8 received a high score from the former and a mid-range score from the latter, whereas event 5 received a high score from the combined DH audience group and a low score from the audience as a whole.

The range in individual group scores is also revealing. Very high group scores, of above 85%, were given to events 2,5,6,7 and 8, which reflect the DH combined scores just discussed.

Very low individual group scores were seen at events 1 (below 50%), and events 4,5,6 and 8 (below 70%), which do not mirror the trends for the combined scores. These low scores are not reflected as fully in the combined scores, since, for example, event 6 was one of the highest-scoring overall and events 5, 6, and 8 also contained the high individual group scores just mentioned. This is an example of how scores given by a single person can impact on the combined score for a very small group of people, for example, 2 or 3 people.

Whilst discussion of the individual events is needed to fully understand the scoring patterns seen, this overview nevertheless provides a good starting point for understanding how the different events relate to each other.

In sections 7.4 and 7.5, I will outline each event, and some of the feedback received. The discussion of certain events will be more in-depth than that of others because of the particularities of the event and the aspects of learning that occurred. Events 6 and 7, in particular, are only described briefly in this chapter, as I use them as a case study for the NER analysis in Chapter Eight.

7.4 Round One

The events in Round One follow the first round of the respeaker training programme, and use the initial technical set-up that was devised⁹¹.

7.4.1 Event 1

7.4.1.1 Event outline

The very first event was an early evening live arts presentation at the Riverhouse Barn, an arts centre in Walton on Thames and it was the lowest-scoring event. As the very first research event, this may have been expected, but there were a number of factors that seem to have contributed to this.

The event was bespoke, created by Quick Fix Theatre, a local volunteer pop-up theatre group I am connected to, to support this research. Their most recent performances had been mimes, related to the cinema history of the Elmbridge area, and performed at local summer fêtes and festivals. Their performances are dynamic and engaging and their unusual costumes and period music often draw crowds who are intrigued by what is going on. This seemed to provide a good fit to the arts and cultural aspect of the research and the group were familiar with the venue we would be using.



Fig. 7.1: The cast of Quick Fix Theatre in *The Imposter* (Quick Fix Theatre, 2016)

⁹¹ Video samples from these events can be found in appendices 8.3.1-8.3.7.



Fig. 7.2: Quick Fix performing *Alice in Wonderland* (Quick Fix Theatre, 2017)

We designed the event so that each half would comprise a mixture of mime, set to music, and a PowerPoint-based presentation on the philosophy and creativity of Quick Fix Theatre as well as opportunities for questions. This would allow regular and verbatim respoken to be tested. The thematic content would relate both to art, literature and performance, and include particular references to the two plays pictured above. One actor would take the lead in the presentation, with other speakers joining him at appropriate points.

7.4.1.2 The space

The event took place in the Barn, a large Tudor-style hall. Although the staging and seating are flexible, on this occasion we were unable to adjust the location of the seating, presentation area and subtitles, due to an event which immediately preceded our own. Figure 7.3 illustrates the layout of the event; the same key (Fig. 4) is used for all the events.

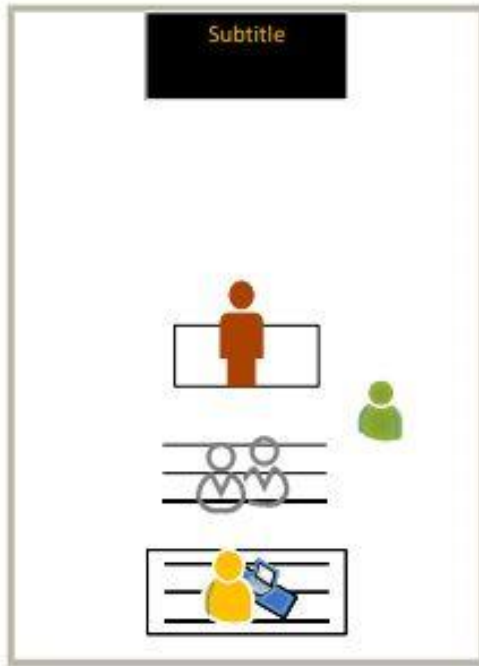


Fig. 7.3: Layout of event 1

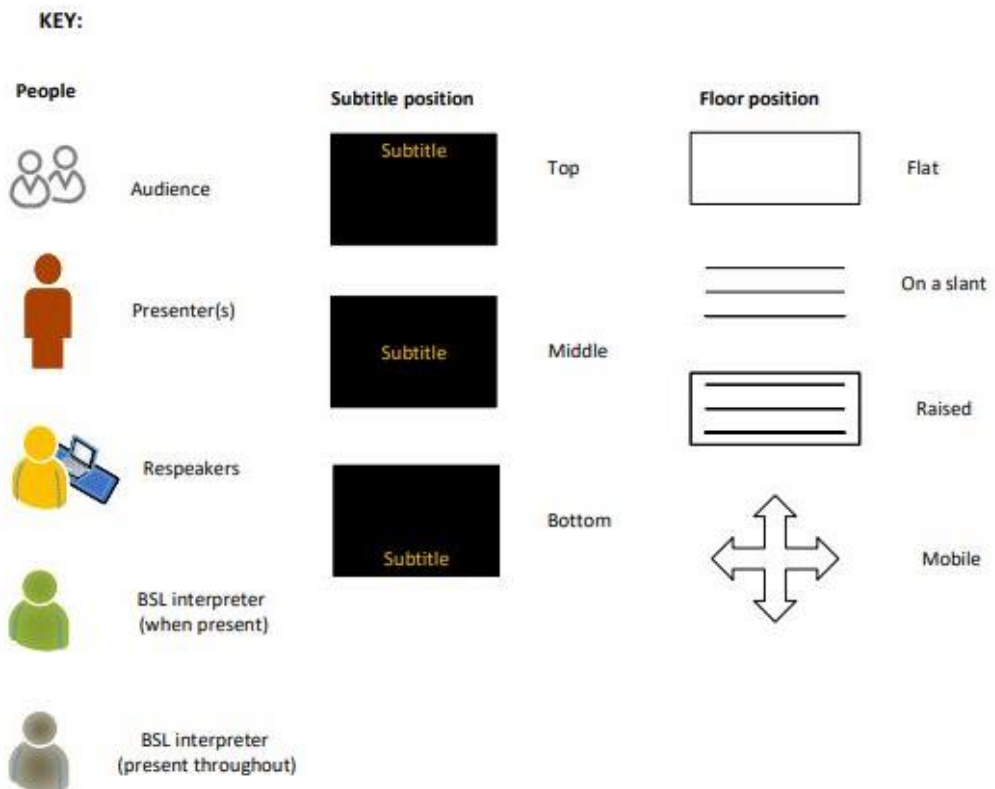


Fig. 7.4: Key for the event layouts

The respeakers were positioned at a raised table at the back of the theatre, and sat directly behind the tiered rows of audience members and in very close proximity to some. The actors and presenters were in the centre of the space and the PowerPoint and subtitles appeared on a projection screen at the back of the hall. Two lines of subtitles were set at the top of the screen.

The screen was a long way from the audience and, in certain seats, the view was obstructed by theatre lights which came down from the ceiling. As the seating was unreserved, audience members were able to choose where to sit. Some audience members noted that they were able to hear the muffled voices of the respeakers, despite their SmartMic masks.

7.4.1.3 Preparation and lead-up to the event

As was the case for each event, the respeakers arrived a few hours early to prepare. They focused on training vocabulary items into Dragon and developing their voice profiles. They also recapped the handover technique that they would be using during the event. The Quick Fix cast arrived a little later, approximately two hours before the event began.

Time was scheduled for both groups to meet and talk before the event began, but the actual meeting was brief. The respeakers needed relocate to the space where they would be working and Quick Fix needed to set up props and work out their staging. As the start of the event drew closer, there was a local power cut, which reduced the time available for technical checks. The equipment was tested and appeared to be working, but no additional time was available for the respeakers to practice handovers using the full sound kit as they had hoped. Unfortunately, as the audience began to arrive, the test subtitles we had displayed dropped off screen. Without a clear reason for this, everything needed to be re-set; on the advice of the Technical & Venue Manager, we minimised the equipment being used.

The performers' music and the key elements of the respeaking kit were unaffected, but the main presenter was no longer able to use a clicker to change slides; instead I did this

on his direction. This delay to the start of the event was unsettling and frustrating for all involved, the audience included, and it meant that we had to adjust the content of the event. The verbatim testing was removed and, as one member of Quick Fix had to leave early, one performance also had to be cut.

The reason I outline this in so much detail is because it reflects the reality of what can happen at an event. At many events, a larger team might be available to mitigate the impact of this chain of events. However, in the live world, situations such as these do arise and may impact on the event as a whole. If we think of broadcast television, it is possible to swiftly replace one programme with another if the original needs to come off air, and a team of respeakers would be available if any individual were unable to stream their subtitles. Live events do not have this facility.

7.4.1.4 Event proper

After the introduction to the event, and audience reflection on a clip of live television subtitles, Quick Fix 'popped up' and performed a scene from Alice in Wonderland. It was clear from the feedback that followed the event, that many audience members were confused by this. The spontaneity and engagement that had so often been seen in fête/festival settings did not transfer to this dynamic. Confusion about the lack of speech during the performance, despite initial sound and music labels, further contributed to this. Many audience members commented that once the presentation began and they understood who Quick Fix were, everything became clearer, but until then they were left feeling lost and wondering why subtitling was being tested for a show without speech. In hindsight, stepping into performance mode later in the event would most likely have worked better. Nevertheless, the audience did engage with the event, with many taking the opportunity to ask questions and staying on for the post-event focus group discussion.

7.4.1.5 Audience responses

In total, 13 people attended this event. One was Deaf, six deaf, three hard of hearing and three hearing. There were no native speakers of any other European or non-European languages.

It was clear from the comments on the feedback sheet that as a BSL user, the Deaf audience member would have preferred the whole event, rather than just the introduction and instructions to be sign language interpreted. They awarded low scores for the accuracy of the subtitles, their legibility, scrolling format and the latency with which they appeared. Similarly, how easy it was to understand the event content, the amount of time spent on the subtitles and the physical movement involved in doing so, plus engagement with the event and speaker also received low scores. However, the statement “it took effort to follow the subtitles” was given a score of four, which indicated little effort being needed, so it seemed it was not the subtitles themselves that were problematic, but perhaps the time and attention they demanded at this particular event. Overall, this audience member gave a score of 3.5 for the access provided by the subtitles. So, whilst they did improve access to the event, it is clear that on this occasion the subtitles did not provide the audience/user experience I was ultimately aiming for.

For all the other audience members, whether DH or hearing, the mean variable scores were higher, sitting in the band 3 to 3.99. Like the Deaf audience member, most agreed that the latency was problematic as the subtitles did appear with some delay, and similar scores for latency were seen across the events for this very reason. Other low scores indicated the effort involved in following the subtitles, either because too much time was spent reading subtitles or because of the physical effort involved. To an extent, this was reflected in some of the feedback as audience members commented that the subtitles were very distant, or partially obstructed, which made reading harder, and noted that the presenter spoke a great deal, and at a fast pace. Higher scores revealed that many felt the main content of the event had been captured in the subtitles, meaning that they did improve people’s access to the event, which was very positive to learn at this stage in the process. In terms of wider engagement, many who attended found the event

interesting and informative and felt immersed in it, with one audience member wanting to invite the main speaker to give a talk to an organisation they were involved in. Others, like the previous audience member, engaged less. More than anything, what was clear was, as the epigraph to this chapter captures so pithily, opinions varied and at this event, and those that followed, whether an event was of interest was a matter of personal choice.

The audience feedback also revealed two technical issues which occurred during the event. The first related to the visual presentation of the subtitles, where the colour and background settings used by the two respeakers did not match, meaning that the letters flickered as each handover took place. This was easily fixed and the solution was incorporated into the preparation material for the next event. The appearance of double letters at the start of some words, referred to in Chapter Six, was also noted. The reason behind this was unclear, though it seemed to relate to how Dragon interacted with the unsupported Text on Top software, and I continued to investigate the issue.

7.4.1.6 Wider reflections on the event

From an organisational point of view, the challenges leading up to the event start highlighted the importance of ensuring sufficient time and flexibility to prepare on site, especially where complex content is involved. In addition, this event highlighted the potential gap that will occasionally arise between what is planned and what can actually be achieved on the day. This is a very important point to understand when considering access. Not all events will be simple, and events should not have to be simple to be made accessible. However, to ensure that access can be achieved successfully, all the steps involved in running an event must be carefully thought through so that different eventualities can be considered and potential solutions found; everyone involved must be aware of their role on the day. In the case of research events, where much of the organisational responsibility lay with myself, it suggested that incorporating subtitles into an already existing event might be a better approach, which would allow me to be free to focus on the research element.

The feedback from the respeakers, venue teams and presenters from this event was also very insightful and an important part of the action research process. Going into the event, both respeakers felt positive and ranked their confidence as 4/5. They knew they had the experience required to create live subtitles. Any nervousness they experienced related to unknowns such as the venue, what audio quality they would have, what the presenters' accents would be like, and the fact that this was the first time they would be using software in a live setting rather than training environment. Both agreed that the focus of their preparation would be on training vocabulary into Dragon and "being (as) aware as possible of what to expect, so as to be able to respeak quickly and accurately" (Respeaker A). With the many references to names and places and the fast pace of the delivery, the content would have been challenging for the respeakers if detailed notes had not been provided.

In their post-event feedback, it seemed that the experience had, for the most part, been a positive one. Although Respeaker B, in particular, would have liked more time, both appreciated having the opportunity to talk to the presenters before the event. Aside from the subtitles not streaming at the start of the event, the equipment generally worked well. The headsets provided good quality audio, although they had to be abandoned midway through the event as a result of continued technical difficulties. One respeaker also experienced problems with Dragon intermittently through the event. Having Wi-Fi at the venue meant they could use their preparation time effectively and they both appreciated the support of the team:

Easy to go up and talk to cast if needed. Definitely great to be able to talk to the Technical Manager, or anyone who is helping set up the sound/lighting etc. Zoe always being present and on hand was great, too.

The respeakers also suggested how their experience could be improved at future events. Whilst it was clear that having enough preparation time was vital, this event pinpointed the fact that preparation time should, wherever possible, be in the actual location that they will be working in, ideally with the full respeaking and sound kit already in place. They also noted that the switch to name tags rather than colour for speaker identification

involved more effort on their part. Not only did it require more time to say each name, but the respeaker also needed to be able to identify who each speaker was if they were to be able to use the list of names provided effectively. At this event, the main speaker introduced the other speakers clearly, but this might not always be the case.

Two presenters shared their experiences after the event. Both appreciated having had an overview of the respeaking process, from the information that was shared before the event, but there was much more they would have liked to have known. Some of this desired knowledge related to a better understanding of the audience and how best to interact with them; both felt more conscious of the words they were using and how they were speaking and were aware of the audience looking past them towards the screen. They also wanted to know how to respond to the subtitles and subtitlers; both were able to hear the respeakers and this took some time to get used to. Whilst they knew they might be asked to repeat content, they were not sure how to respond to any latency in the subtitles and both said they would have liked a screen in front of them where they could follow the subtitles themselves. As the events progressed, it became clear that in addition to the training that the respeakers had received in order to transfer their skills to the live event setting, a similar offering would also be of benefit for presenters, something which is explored in more detail in section 7.7.

Whilst this first event did not provide the smoothest start to this section of the research, it was an extremely valuable learning opportunity, which highlighted the appropriateness of the action research methodology that was adopted. The fact that the next event would be in the same venue, with many audience members returning for it, meant that event 2 provided a great opportunity to put action research into practice and swiftly implement a number of changes.

7.5.2 Event 2

7.4.2.1 Event outline

The second event at the Riverhouse Barn was a presentation by Pete Allen from the RC Sheriff Trust. His talk continued the theme of drama and theatre, but the event was

organisationally extremely different from the first. The presentation was one that Pete gave regularly, so although he did adjust it slightly to incorporate an interval where verbatim respeaking could be tested, the content he included did not differ too greatly from previous talks.

7.4.2.2 *The space*

On this occasion, our event was the only one being held in the Barn that day, so we had more flexibility in how we set up the room.

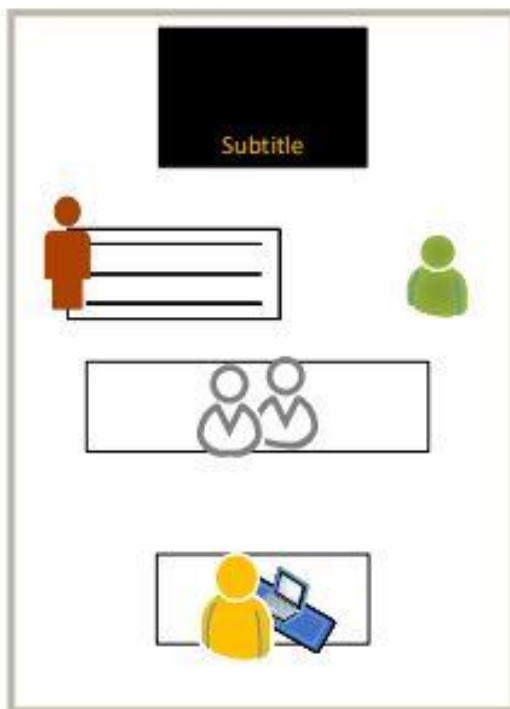


Fig. 7.5: Layout of event 2

Pete was on the left-hand side of the venue on a low stage. A projection screen hung down from the ceiling and displayed the PowerPoint presentation, with two lines of subtitles at the bottom of the screen. The audience were seated immediately in front of the stage. The respeakers were located further back in the hall, and were more removed from the audience than they had been at event 1. They had a clear line of sight of Pete and the subtitles. They prepared for the event in this same location, with the full sound kit already set up. Respeaker E was also present in the venue, respeaking as-live from

the same location that I had used myself (section 5.4). His subtitles were not streamed to the audience, but he did take part in the post-event discussion.

7.4.2.3 Preparation and lead up to the event

Although it was respeakers C and D who subtitled this event, they had been in touch with respeakers A and B to discuss their experience at event 1. In addition, I had incorporated the feedback from the first event into the information I shared with them for the day and in my discussions with Pete. The Technical and Venue Manager had a better understanding of our technical needs and there was time for everything to be set up smoothly.

The respeakers arrived first to prepare for the event and familiarise themselves with the equipment in this new setting. They had received the slides and detailed information about the event, including notes about the content, beforehand. Pete arrived nearer the start of the event, in time for a full meeting between him and the respeakers and the opportunity for him to trial respeaking in action.

The feedback from all involved suggests they felt prepared for the event. Pete in particular stated that, “talking to the respeakers was invaluable in getting a sense of how everything worked and the parameters I had to work within”. In addition to understanding the technical set-up, having a better understanding of who the audience would be was also very helpful for him.

7.4.2.4 Event proper

From the perspective of the audience, this was the highest scoring of all the events, receiving a score of 84.2% from DH audience members as a whole and an only slightly lower score of 83.8% from the full audience. 16 people attended in total, of whom one was a Deaf BSL user, eight were deaf, six hard of hearing and one hearing. When asked about language, the Deaf BSL user identified as being bilingual English/BSL and the person who was hearing was NNE, with a different European language as their native tongue.

The mean variable scores for most groups sat within the 4 to 4.99 range, with the score for the hearing group sitting just below at 3.98. Many scores of five were awarded, from Deaf BSL, deaf and hearing audience members. Clusters of fives were seen across the statements related to interest and engagement and to the physical space in the venue and even where five was not awarded, the statements were generally ranked highly. It seemed audience members had a clear view of what was going on and that the position of the screen and lighting levels were well adjusted so that they did not interfere with the ease of reading the subtitles from a physical perspective. Additional feedback revealed that the lighting could have been further improved by up lighting the eyes of the speaker. Some effort was nevertheless involved in reading the subtitles, with this statement receiving a score in the 3 – 3.99 band across most groups.

There seemed to be a general consensus that the subtitles captured the main content of the event well and that the subtitle presentation, in terms of legibility, display mode and the proportion of space they took up on screen, was also optimum, or close to it. Where errors did appear, they were easy to work out (ID1) and the editing seemed appropriate:

I was impressed with the editing when too many words are spoken for them all to be put on screen – the gist was always there. (ID21)

Where lower scores did appear, they related to the accuracy of the subtitles in terms of spelling and punctuation. Hard of hearing audience members gave this variable a score of 2.83; the Deaf audience member ranked it 4.00 and all other groups ranked it in the band 3 to 3.99.

Latency was still ranked low, with a score in the 2 to 2.99 band, though some audience members who had attended both events commented on the improvement in latency, and the subtitling in general, between the two events. One audience member said:

Very interesting and good timing of good subtitles connected to speaker. Audience Q&A were represented correctly! (ID22)

For this reason, it seems unlikely that given the production method of live subtitles, a score for latency will ever reach 5.00. In the feedback for this event, a number of

audience members also commented that this natural delay was of use to them. One wrote,

I always struggle with talks and plays and although I was able to follow Pete from RC Sheriff it was really useful to be able to check the subtitles if I wasn't sure I had heard correctly... Especially when it came to catching names of individuals, so very useful. (ID19)

Reflecting back on the event, the NNE said that although the subtitles did not increase their access to the event,

They did help at a couple of points of the presentation to visualise why something was funny – Mr Hoare – same pronunciation as whore/minimal pair – and helped me to put dates into context when a lot of information was given. (ID15)

I also knew from pre-event discussions that one audience member was dyslexic. They found the subtitles very clear to view and said, “the size and spacing of letters (were) excellent and meant, for once, I was able to complete the reading before the wording disappeared” (ID23). In addition, the yellow on black contrast worked well, as did the small number of words put up at a time.

All in all, it seemed that not only was the content of the subtitles good, but that they were well-integrated into the event and added to the experience, rather than being a point of distraction. As one deaf audience member said, “I was confident to ask questions based on the subject. Normally I would keep quiet because I usually find it difficult to follow everything” (ID18). This sums up the heart of the research. With good subtitles, people can choose if and how they engage with the event. As another said, “Without subtitles nothing. Sans-subtitles rien!” (ID3).

Despite the clear achievements of the event, a few technical issues remained; in the verbatim section, there was an excessive use of punctuation which did confuse some audience members; they turned back to see whether the speaker was continuing or had in fact finished their sentence. This seems to be more a reflection of the added pressure of verbatim work, than of the implementation of respeaking at live events.

7.4.2.5 Wider reflections on the event

As presenter, Pete's experience of the event was a positive one. He was unable to hear the respeakers, given their distance from his own location, and there weren't any points where he was asked to repeat information. There was nothing he would have liked to be done differently. When asked how this experience differed from other events he had spoken at, he said:

I was conscious of speaking slower and trying not to turn towards the screen. I also felt I held back on going 'off script' even in the anecdotal sections.

From the point of view of organising accessible events, this is certainly something to be aware of. It is clear it is necessary to find the right balance for the presenter so that as well as being conscious of what is said and knowing what information to share with respeakers, they do not begin to omit content for fear of going off script. This is something that will be addressed further in the presenter training below. What was also encouraging, was that taking part in this research, meant he already felt more confident about talking to DH audiences in the future:

I certainly feel that I have had a grounding in how to give a talk to an audience for whom hearing or spoken English is an issue. I hope that it would enable me to give more talks to audiences with these needs.

This is in line with the key hope from this research that with better understanding, any fear related to accessibility begins to dissipate. This reassures me, too, as researcher, that as accessibility and varying needs are better understood, the process of providing for them will become a natural part of what is always done, a phenomenon I refer to as 'stealth access'⁹².

An additional piece of feedback on presenting to DH audiences came from the audience, stating that it would be helpful for there to be longer pauses whenever pictures are displayed to allow audience members to look away from the subtitles, as there was a

⁹² By this I mean that once access is known about, an awareness and application naturally creeps into new settings, just like the stealth aircraft which were able to enter new territories undetected. A thank you to Catherine Pawasarat, who used this idea of stealth in a different setting.

risk of subtitles being missed if the speaker carried on talking. Similarly, more pauses were needed after questions. These were points that had been included in the information for presenters, but that may take time to embed into natural practice. Where sometimes presenters may feel uncomfortable about pausing for the audience to take in ideas or about being silent, in this particular setting it is something that can be extremely beneficial for DH audience members, and, most likely, for all attending.

7.4.3 Event 3

7.4.3.1 *Event outline*

The third event was another bespoke event, this time organised around a film screening and Q&A panel. It was held at the BFI St Stephen St (<https://www.bfi.org.uk/venue-hire/hiring-bfi-stephen-street>) and it was the only event in Round One that took place in the afternoon. The producers and directors of two films, *Blue Pen* (McNamara and Kimyoncu, 2016) and *Notes on Blindness* (Middleton and Spinney, 2016), were invited to present and discuss their work. As a short film, *Blue Pen* was played in full, whilst clips of the feature-length film *Notes on Blindness* accompanied the presentation; subtitled versions of all were used. A film expert and colleague from Roehampton chaired the session and, following some general guidance, the directors chose the format that their presentations would take. Because of the screenings, this was the longest of the events and a slightly longer interval was scheduled for that reason. This was the second and final event where verbatim respeaking was also tested.

7.4.3.2 *The space*

The screening room was intimate, with 36 seats arranged on a slight incline. The films were shown on a large screen at the front of the room, and the live subtitles were displayed here too, as two lines at the top of the screen. The presenters either sat or stood in front or to the side of the screen. The main subtitles were behind them, but were also displayed on a laptop at the side of the room, which the presenters could view if they wished.

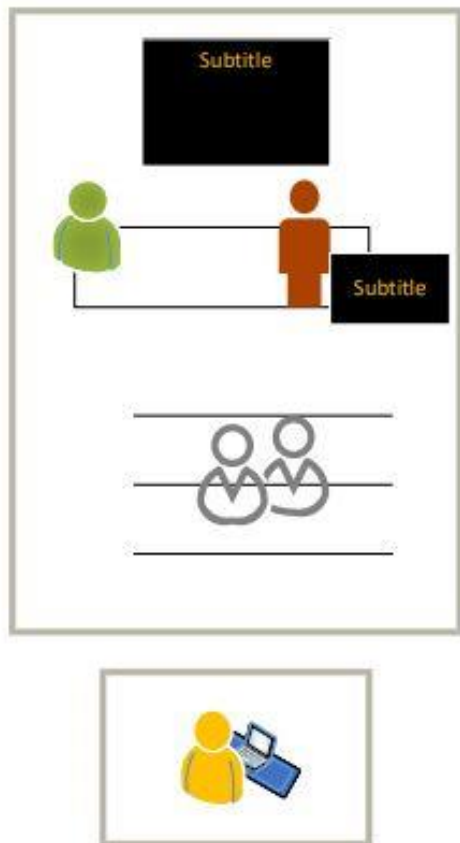


Fig. 7.6: Layout of event 3

This was the first event where the respeakers were not situated in the main event room. Instead, they were located in the projection room behind it, where they were able to connect into the cinema sound system. Being in this different location meant there was no risk of the audience hearing them through the SmartMic mouthpieces, but it was not the most comfortable location. The desk was modified for holding film reels, and so had circular attachments that laptops had to be positioned around, and there was background noise as the room was air-conditioned to cool the equipment. Although there was a view into the screening room and a television monitor showing what was taking place there, the desk was positioned facing the other way. Nevertheless, going into the event, both felt confident since they had now worked in the live environment and were familiar with the process.

7.4.3.3 Preparation and lead up to the event

Connecting into the sound system of a new venue is a potentially complex procedure and one that had required a full system check in advance of the event. I had visited the venue the previous week to do this and resolved a number of complex issues, which including buying a new type of cable. This meant that the procedure was fairly smooth on the day and we had the reassurance of knowing that the systems we were using were compatible. Nevertheless, during the event, we did encounter some challenges with the microphones and sound transfer.

7.4.3.4 The event

Julie MC of VitalXposure was the first to present. After a brief introduction, she asked for *Blue Pen* to be screened, preferring to take questions once the audience were familiar with the film. Pete and JoJo from Archer's Mark talked about *Notes on Blindness*. Pete led with a PowerPoint presentation, which was interspersed with short clips from the film and of related content. I had specifically wanted one event to involve video content as I wanted to see how the respeakers and presenters worked around the switch from live to burnt-in subtitles. I knew from my own experience during the focus groups that this might require some navigating and the switching during the *Notes on Blindness* portion of the event provided ample opportunities to explore this.

Pete was very attentive as a presenter to the subtitles and it was clear that he was following them and, on occasion, modifying what he said to account for any latency. He quickly realised that it would take a moment for the live subtitles feed to pick up after a video and tried to accommodate this as he spoke.

Whilst that was a technicality that was expected and that we wished to explore, we also encountered another technical difficulty, that of the sound transfer through the microphones. Unfortunately, on numerous occasions, the respeakers had to resort to displaying (INAUDIBLE) on screen, as a shorthand request for the presenter to repeat what was said. Although the presenters had been warned that this might happen during the event, it was the first time we had had to use it. I was monitoring the subtitles and

asked the presenters to repeat content when this was displayed. If this happens once or twice during an event, it feels like it is something that the presenters can work around fairly naturally, and the audience can accept. However, it quickly became clear that this was a wider sound issue. During the event I tried to deal with it by switching between the two microphones we were using and by replacing the batteries, but this did not resolve the issue. On reflection after the event, it seemed that the issue was with the quality of the microphones, as opposed to the headset relay. For this reason, before the second round of events began, I refined the sound system we were using (section 6.9 above). Whilst incredibly disruptive during the event, it was, once again, a useful learning step within the action research process.

This interaction with the respeakers and need for repetition also highlighted a different response from the presenters. Since the focus of *Blue Pen* was on censorship, when content from Julie MC's presentation was initially dropped, rather than repeating what was missed, she responded by saying "I've just been censored" and later "Further censorship!", drawing what was happening into her words. This highlighted the dynamism of live events, and also presented interesting new pathways in the NER analysis when it came to analysing the quality of the subtitles from the perspective of what was omitted; this discussion is explored further in Chapter Eight, section 8.6.

7.4.3.5 Audience response

Out of the 17 people who attended the event, six were deaf, seven hard of hearing and four hearing; 15 were native English speakers and three were native speakers of non-European languages⁹³. The majority of scores fell into the 3 to 3.99 and 4 to 4.99 range. As seen at previous events, the lowest scores were assigned for latency and to the fact that a great deal of effort was involved in reading the subtitles, and this was true across all the groups. The hearing group also experienced sensations like a headache or tiredness more than the other groups.

⁹³ This is one example of an event where the totals from the question about hearing status and native language led to a discrepancy in the numbers. The data suggests one participant was bilingual, in English and a non-European language.

The highest scores related to the fact that the audience had a clear view of what was going on, with the subtitles taking up a good proportion of the screen and being well-positioned behind the speakers meaning little head and neck movement was involved when reading them. Similarly, everyone present found the event interesting and informative. The majority of people felt that the subtitles captured the main content of the event, and that the live subtitles were readable in terms of size, font and colour. The scrolling format of the event was slightly less popular, perhaps because of the contrast with the blocked subtitles in the films, and a few people felt the subtitles disappeared before they were able to read them. The scores indicated that speaker identification could also be improved, with hearing audience members scoring this at 2.75 and everyone else between 3 to 3.99.

7.4.3.6. Wider reflections on the event

Although this was another event with complex technical issues that were disruptive for the audience, it was nevertheless a great learning opportunity. The post-event chat led to a very interesting discussion on how language and accent should be displayed within subtitles, that is developed further in 7.6 below. In addition, this event highlighted the effort involved for people using subtitles to follow events.

One aspect of this event that I had not considered was the potential difference between the size of burnt-in and live subtitles, especially if the audience is switching between both. That will certainly be a consideration to take forward as a recommendation when supplying video content at events. This was also a long event. I had focused on the duration of the live subtitling, without paying enough regard to the time that the film content would add. Another recommendation going forward will be to consider the length of the event when making the arrangements surrounding the content.

This does not mean that accessible events should be short or, to use the words of participant OL15 from the DH focus group in a slightly different way, "dumbed down"; rather, it means that the added effort that using subtitles might involve should be kept to the fore as other aspects of the events are planned. For example, this might lead to

longer or more frequent intervals during a full-day event, or even access to the transcripts following an event, so there is a reassurance that content can be re-viewed later.

This was the last event where we attempted to test verbatim respeaking. When, during the post-event chat, I revealed that we had tried to test this, it was evident that the participants had not been aware of the attempt. With so many other variables to consider during the remaining events, plus the added burden on respeakers that I knew verbatim posed, I decided to focus on regular respeaking for the remaining events.

7.4.4 Event 4

7.4.4.1 Event outline

The fourth event was a museum tour in the Wellcome Collection, London, entitled *Medicine Man* (Wellcome Collection, no date), which outlined the history behind many artefacts in the permanent exhibition of the same name. This was the first event where live subtitles were added to an existing event. The Wellcome Collection regularly held events which were subtitled via STTR and this was a tour that the guide had given before. The Visitor Experience Manager (VEM) with responsibility for access and inclusion, who accompanied the tour, was used to monitoring the subtitles as access coordinator for the event. This meant that for the first time I could step fully into the role as researcher and observe the tour taking place.

7.4.4.2 The space

The tour moved through the gallery and each participant had an individual tablet, loaned to us by Stagertext, which allowed them to roam freely. Although the respeakers could have worked fully remotely offsite, on this occasion they were located in the Wellcome Collection building.

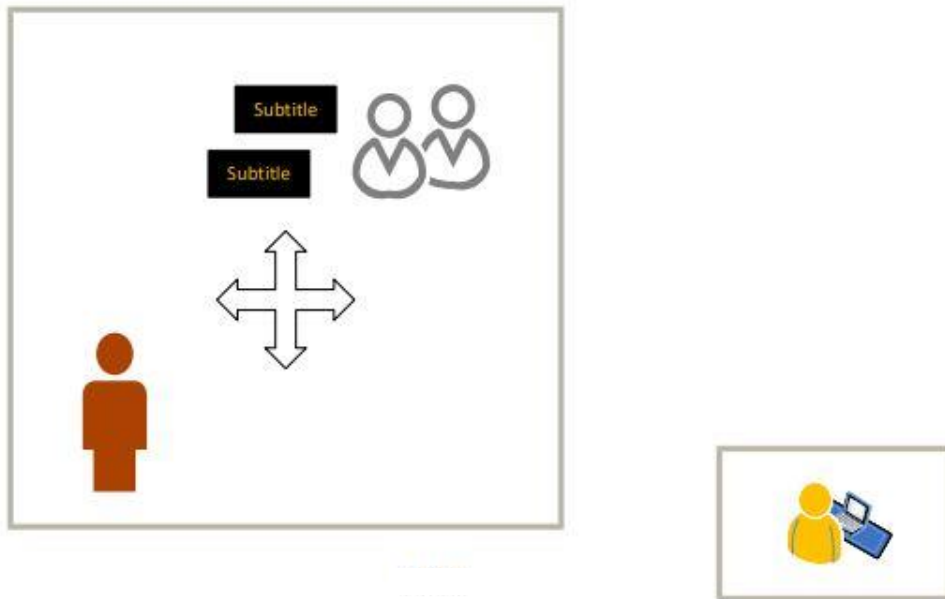


Fig. 7.7: Layout of event 4

On this occasion, rather than using Text on Top as they had done for the previous three events, they switched to Streamtext as this software enabled them to stream to each individual device. As for event 3, the respeakers both felt fairly confident although they were concerned whether the technology and internet would be reliable during the talk. They would have liked visual aids alongside the notes about the content; the respeakers did have the option of visiting the gallery before the event, but preferred to focus on their voice model preparation. The respeakers did note that this remote set-up, and absence of visuals, made the event more like the broadcast work they were used to. The particular challenges they envisioned were the use of Streamtext and also how the audio connection between themselves and the tour guide would work.

7.4.4.3 Preparation and lead up to the event

During the meeting with the guide, the respeakers were able to test the audio connection over the mobile phone and set their headsets correctly. Following that, the respeakers continued with their profile preparation, while the VEM and I moved to our base near the gallery where we proceeded to ensure that the devices participants would use were working correctly.

We checked their battery levels and Wi-Fi connection and then connected to the link for the event that had been generated via Streamtext. We set the subtitles to yellow letters appearing on a black screen and selected an optimum font size. We then alerted the respeakers, via mobile phone, that we were ready to test the subtitle stream and performed a technical check to ensure that they were streaming correctly. We also connected a screen in the room to the Streamtext feed. We used this room for the event introduction, and we returned here after the tour for refreshments and the post-event chat.

During their preparation, time had been set aside for the respeakers to familiarise themselves with Streamtext as it was the first opportunity they had had to work with it. Much time was devoted to perfecting the handover procedure, since this was significantly different from the one they had been using in Text on Top.

7.4.4.4 Event proper

The tour lasted approximately 40 minutes and each participant had their own device. The guide held a mobile phone, set to speaker, and used this to deliver the tour both to the people around her and to the respeakers, who had connected their headsets to their own phone. As she led the audience around the gallery, the guide stopped at numerous artefacts and points of interest to deliver her tour and respond to questions that she was asked. From her previous experience and discussions with the VEM, she was aware that she had to repeat any audience questions so that the respeakers would be able to hear and subtitle them, and she did this confidently.

On two occasions, there was a significant drop in subtitle feed, which seemed to correspond to handover points. We were able to resolve the streaming issue, but it did cause a temporary delay of a few minutes each time.

7.4.4.5 Audience response

15 people in total attended this event. One identified as a Deaf BSL user, five were deaf, four hard of hearing, one 'hearing but I don't hear as well' as I used to and four were hearing. In total, there were three non-native English speakers present.

Looking at the combined scores for the events, the majority of scores across the Likert scales sat between 3 and 4.99. The statement on latency received the lowest scores across the board, which was unsurprising given the two drops in subtitles that had occurred. On the whole, those attending engaged well with the event and found it interesting and all felt they had a clear view of what was going on, most likely because of their personal streaming devices and the ability to move to a new position if they wanted to see an artefact more clearly.

Most people found it was fairly easy to follow the subtitles and visual content of the event, with the majority of scores this statement sitting between 3 to 3.99. The statement that subtitles were accurate in terms of spelling and punctuation receive scores between 2 and 3.99, and it seemed from the comments and post-event discussion that this was due to the reappearance of doubled letters at the start of some words.

On the whole, those present felt the main event content was captured accurately within the subtitles. A few people did feel they spent too long looking at the subtitles and noted some effort involved in this.

For the most part, people set the screens to show multiple lines of subtitles which meant that they were able to look at artefacts and then return to the subtitles to see what had been said. However, this balance between looking and reading is clearly something that needed to be developed across events. In some sections of the gallery, there was a great deal of background noise and at that point many noted that the access provided through the subtitles was even more important for them; without them, they would have been unable to follow.

7.4.4.6 Wider reflections on the event

This event placed a spotlight on the impact that any drop in subtitles has on the experience of the event as a whole for all involved. Even though at this event people could walk around and stay occupied while waiting for the subtitles to return, the fact that content was disrupted clearly stood out in people's minds and it is one factor that must be more reliable. From talking to the tour guide and VEM, they did find working with

respeaking more challenging than working with STTRs, so this was a key factor that we sought to improve as we approached the second round of events.

7.5 Round Two

The events in Round Two followed the second round of respeaker training. All were pre-existing or repeat events, to which access was being added. Three of the events were held outside of London, which meant that new people attended and I went into each event without knowing how many people would be there.

The more distant locations meant that the pre-event site visits were more challenging to schedule, but I did visit each venue in advance of the event. Since the modified sound kit we were using in this second round worked independently of the in-house system, the need for a full technical check during this pre-event visit was removed.

7.5.1 Events 5 and 6

7.5.1.1 *Event outline*

As variations of the same events, I will present events five and six together. Both consisted of a Q&A discussion session, guided by a host, which followed a subtitled screening of *The Piano* (Campion, 1993). The first took place at Watershed in Bristol (www.watershed.co.uk), as part of their regular series of Deaf Conversations; the second was held at Depot in Lewes (lewesdepot.org). The team at Watershed booked the two hosts and the sign language interpreters for event 5, and I arranged them for event 6.

These events differed from the others as they were sign language interpreted throughout. Deaf Conversations at Watershed are always sign language interpreted, and in this instance subtitled access was being tested to see if it would pave the way to making the Conversation more accessible for non-signers. At Depot, we had not originally intended for BSL to be used for the whole discussion, but since it was advertised a signed event, we did not want to knowingly remove any access that people were expecting, and risk falling into the trap of mis-advertising what access would be available (section 5.3.3.7 above).

7.5.1.2 The spaces

Although the event content was broadly similar, the spaces they were held in differed greatly.

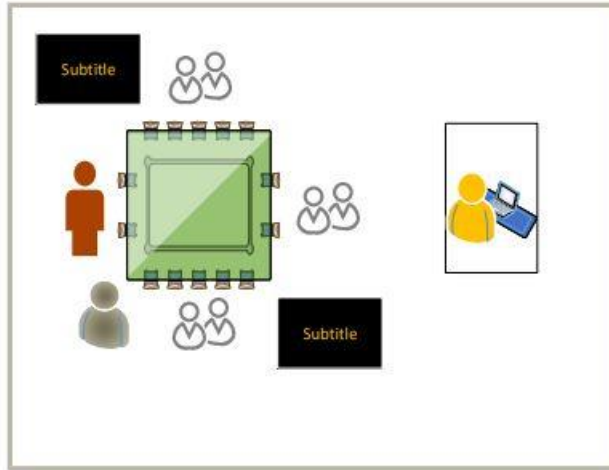


Fig. 7.8: Layout of event 5

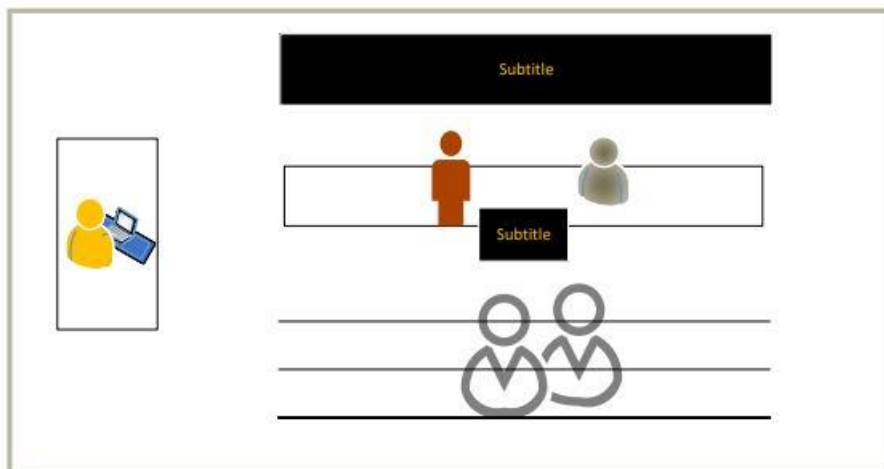


Fig. 7.9: Layout of event 6

At Watershed, the respeakers gave a brief introduction to the event in the screening room, where they were located to the right of the screen, as the audience viewed it. The main conversation, however, was held in Watershed Café. Here, the audience and presenters sat around one table and the respeakers were located at another table in very close proximity to them.

At Depot, the respeakers worked in the main screening room, although they did not watch the film with the audience. Once again, they were positioned at the front of the room, this time to the left of the screen as the audience viewed it.

7.5.1.3 Preparation and lead up to the events

At both venues, since the café and screens were in constant use in the run-up to the event, much of the practice was done in a separate space on site. This required a fairly fast turnaround time in terms of setting up the equipment, but since the respeakers were now trained in doing this, and were on site for much of the day, we were able to work together quickly as a team.

Three respeakers worked across the two events. All had seen the film in advance of the event, which meant that they were able to quickly train in the necessary film-related vocabulary. However, since these were Q&A discussions, which would be guided by those present, it was not possible to prepare as fully for these events. At Watershed, there was only an opportunity for a very brief conversation between the presenters and respeakers; at Depot, this opportunity was more extensive.

7.5.1.4 Event proper

Both events provided a good opportunity to test the new equipment in action and it worked well.

At event 5, two screens were positioned around the main table so that the audience members who wished to could view the subtitles. One screen was a laptop and the other was a display monitor that belonged to Watershed. There was a clear discrepancy in subtitle latency on each device, as the subtitles streamed very slowly to the display monitor, but appeared on screen with an expected latency on the laptop. It is not clear why this happened, but it was an unanticipated finding that such discrepancies could occur. It is something to be aware of, especially if presenters are viewing a different feed from the audience and trying to use that feed to tailor what they are saying to accommodate the latency. For future events, this became something to check as everything was being set-up. At this event, there was an additional latency within the

subtitles as much of the original content was in BSL, which was then interpreted into spoken English and finally respoken. Since the sound quality of the recording was very poor, it was not possible to calculate the latency, or do a detailed NER analysis post event.

In Depot, the subtitles were streamed to the main screen and to a laptop for the presenter to view but no such discrepancy in latency was seen. This event took a very relaxed form - as the audience were very engaged in discussing the process of creating subtitles and providing access, the Q&A organically evolved into the post-event chat.

7.5.1.5 Audience responses

The audience at Watershed was the smallest of all the audiences so far. Out of the ten people who completed the questionnaire, seven were hearing, one identified as 'hearing but I don't hear as well as I used to', one was deaf and one was Deaf, stating BSL as their native language. Out of the other nine people, eight were native English speakers and the final person spoke a different European language.

The scores awarded for this event varied greatly, as might be expected because of the experience across the two screens. When considered as a combined group, deaf and hard of hearing audience members did feel that the subtitles had increased their access to the event and the mean score across all variables was 4.15. Most Likert-scale statements were scored in the 4 to 4.99 band; the remainder scored between 3 to 3.99. Those lower scores related to the effort required to follow the subtitles, the fact they read the subtitles even though they did not want to and the time spent on the subtitles. Similarly, they related to the ease with which the subtitles helped identify who is speaking and the subtitle latency.

Amongst hearing audience members, lower scores were seen. Latency and ease of speaker identification were ranked even lower than among DH audience members as was the effort involved in following the subtitles and the challenge of following the subtitles and the visual content.

In contrast, the subtitles were far better received at Depot. 13 people completed the questionnaire following this event, out of whom one person was Deaf and a BSL user, two were deaf one was 'hearing, but I don't hear as well as I used to' and nine were hearing. In addition to the BSL user and nine native English speakers, there were two native speakers of European languages and one other bilingual speaker.

The statements which received lower scores from all groups were the ease with which different speakers could be identified within the subtitles, subtitles appearing at the same time as people spoke, and the effort required to follow the subtitles, although at least one user ranked all but the statement about effort as a 5.00. Both the Deaf, hearing and hearing but audience members responded that they found themselves reading the subtitles even though they didn't want to and the latter two groups thought they spent too much time looking at the subtitles. In this event, the variation between audience members and across variables was clear.

7.5.1.6 Wider reflections on the events

The event at Watershed was a very intimate one and the appearance of respeakers there, especially in such close proximity, interfered with that. This was less of a problem during the cinema introduction, but very noticeable in the event at the café. Whilst this does not mean that access cannot be provided through live subtitles at an event such as the one in the café, this may have been one occasion where locating the respeakers off site, or at the very least in a further removed position, might have been a better option. It was only as the event began that this became clear. At Depot, a few participants commented on the amount of equipment the respeakers were using, especially with their flashing headsets and unusual masks. Some suggested that given this use of equipment, and the potential of hearing the muffled voices of the respeakers, the audience should be warned about this in advance, simply so they would not spend the beginning of the event trying to work out what it was that they could hear. With reference to this sound, some audience members found it very challenging, whereas others thought that especially given the widespread use of open plan offices, at least in pre-pandemic times,

it should be something that most people could adapt to. Further discussion of event 6 can be found in section 8.5.1 below.

7.5.2 Event 7

7.5.2.1 Event outline and space

The seventh event took place at Manchester Art Gallery (manchesterartgallery.org), and involved an afternoon highlights tour around the museum. As for event 4, the respeakers were located on site at the museum, but did not accompany the tour; the post-tour chat took place in the room that they were located, so the audience had an opportunity to meet the respeakers at that point and see them in action.

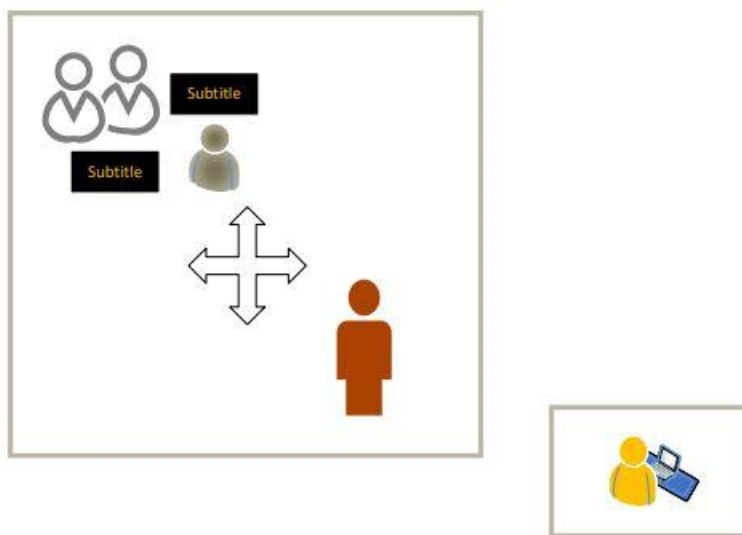


Fig. 7.10: Layout of event 7

Highlight tours regularly take place on a Sunday afternoon in the gallery. They are widely advertised, but there is no need to pre-book, so the exact number on a tour is only determined on the day. A number of guides lead these tours, and they choose what content to include. On this occasion, the artworks were also determined in relation to their proximity to the Wi-Fi hotspots around the building. The tour guide was experienced and had been asked to participate by my contact at the gallery; I met her for the first time on the day of the event, though she had received the information for presenters beforehand, and shared key information about the content of her tour.

7.5.2.2 Preparation and lead up to the event

The respeakers had travelled to Manchester the previous day, so were able to have a relaxed start on the day of the event. On arrival, they set up, trained in new vocabulary and practiced their handovers on Streamtext. Since they knew that there would be natural pauses as the tour moved between the artworks, the respeakers decided to carry out the handover at these points. Although they did have the option of walking around the gallery to see the artworks before the event began, both respeakers decided to rely on the images of the work they had found online instead.

The tour guide arrived in advance of the event and had a chance to talk to the respeakers and briefly test out the sound system. Although interested in taking part in the event, she did appear to be nervous at the technology involved.

Since those attending the tour would each be given a Stagertext device to stream the subtitles, as for event 4, the hour before the event was spent ensuring the tablets were ready for use.

7.5.2.3 Event proper

Twelve people attended the tour which meant that there were tablets available for all who wished to use one. In addition, we had one or two spares. This was very helpful as, on occasion, the subtitle feed to one tablet dropped and we were quickly able to replace it. For the most part the streaming went well, although some of the handovers were slow. As the tour went on, the guide gained in confidence and became more used to the time she needed to leave to allow the audience to view the artwork and follow the subtitles. In addition, it was clear that she had a very good sense of how the respeakers might be coping with what she was saying. When she introduced new terms that had not been shared with them in advance, she also offered spellings and additional information to make their work easier, without being asked.

To save her from holding the mobile phone to her mouth throughout the tour, I had equipped her with a Bluetooth earpiece. Although it worked well during the test, it was patchy in some of the galleries, so she eventually switched to using the phone without it.

In future events, this is something that must be tested more extensively along the length of the tour route.

7.5.2.4 Audience response

The twelve audience members included one Deaf BSL person, one deaf person, one hard of hearing person, seven hearing people and one person who said they were hearing but did not hear as well as they used to. Out of these people, two were native speakers of non-European languages.

The majority of scores across audience members were high, ranging from 3 to 5. Where lower scores did appear, scattered across the groups, they were most consistently attributed to the head and neck movement involved in following the subtitles and visual content, the ease of identification of speakers and the latency of the subtitles. Combined, DH audience members ranked the statements on access being provided, interest and engagement, the clear view within the environment and the subtitle presentation with scores between 4 and 4.99.

7.5.2.5 Wider reflections on the event

Overall, this was the second highest scoring event and the feedback in the focus groups, from the audience and venue contact alike, was very positive. Much related to how this access could be replicated and what could be done to support the work of the presenter; this is discussed further in section 7.7 below.

7.5.3 Event 8

7.5.3.1 Event outline

The final event was held at the University of Roehampton. I had wanted one further event in London, so that audience members from the first round of events could attend and see how the system and techniques had developed. At this event, I was the primary organiser. To simplify what was involved, I booked a public speaker, Paul Connelly, from the Diane Mannering website. This was a talk that Paul regularly gave at different

functions and, based on the book *Why Zebras Don't Get Ulcers* by Robert M. Sapolsky (2004), it was psychological and motivational in its nature.

7.5.3.2 *The space*

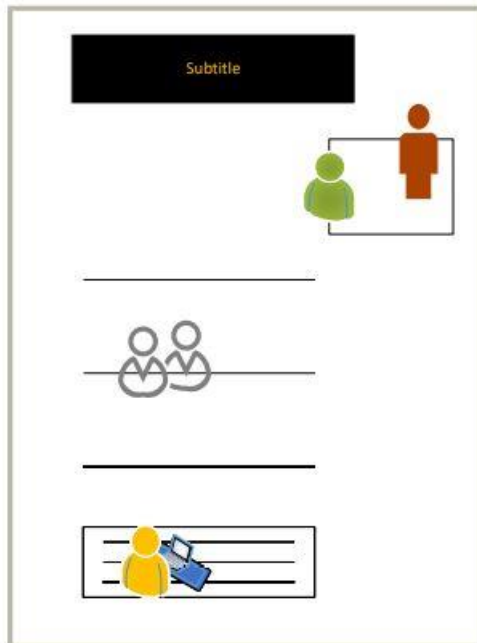


Fig. 7.11: Layout of event 8

The event was held in one of the main lecture theatres on campus. Paul's talk relied on the spoken word; he did not use slides, but did use a number of visual gestures, for example as he illustrated an anecdote about lions running and playing. He stood in front of the main screen, where three lines of subtitles were displayed mid-screen. The seating in the lecture theatre room was tiered and the respeakers sat at the very back, at the highest point in the room. The audience were spread out in the lower rows of seats.

7.5.3.3 *Preparation and lead up to the event*

The respeakers had been supplied with very detailed notes about the presentation which helped them in their preparation. Paul arrived early and was able to chat to the respeakers, experience being respoken and ask questions in good time before the event began.

7.5.3.4 Event proper

This event felt like a social occasion. Some participants came from the Department of Media, Culture and Language at Roehampton and had an interest in translation and respeaking; others were audience members from previous events.

The subtitling went smoothly, though we encountered one slight technical hitch. The main computer was set to sleep if left unused and the appearance of subtitles on screen streamed via Text on Top did not qualify as being in use. The first time this happened, there was a slight delay as we reset the subtitle feed. Following that, I was able to maintain the feed by moving the mouse at regular intervals.

7.5.3.5 Audience response

Eleven people attended in total. Three were deaf, two were hard of hearing and six were hearing; four were native speakers of European languages other than English, and two were native speakers of non-European languages.

Only three statements had mean scores below 3.00. The first was that it took effort to follow the subtitles, which received a score of 2.67 from deaf audience members. Hard of hearing audience members found themselves reading the subtitles even though they did not want to, and scored this statement 2.5 and finally, all groups scored latency between 2 and 2.99.

DH audience members were positive about the lack of physical effort involved in following the subtitles, the clear view and lighting, and generally engaged well with the speaker and the event. They also felt that the subtitles captured the main content of the event and were readable in terms of format and display and took up the right proportion of space on screen, scoring all of these between 4.00 and 5.00. Hearing members of the audience gave slightly lower scores to the main content being captured and to the proportion of the subtitles on screen, which fell in the range of 3.00 to 3.99.

7.5.3.6 Wider reflections on the event

Although this was a talk of a general nature, there were a number of scientific and psychological references within it and some of the content was quite detailed. In addition,

since there were no slides, there was an increased reliance on the spoken word to follow the content. One deaf audience member commented that as a result, the subtitles were even more important. Others said that due to the nature of the terminology, it was useful to see the representation of unfamiliar words. Since the terminology in this event was vital, I also stepped in on a number of occasions as designated access co-ordinator, when I realised that Paul had used a word which had not been given to the respeakers in advance.

A number of non-native English speakers attended this event and they also felt that the subtitles helped complement any vocabulary they might have missed. They did, however, find the accuracy of the subtitles distracting at times and found they got caught up on any errors that did occur. Many of these non-native English speakers were students, and they suggested that subtitles would be helpful in similar technical or scientific lectures where there might also be a lot of terminology to follow.

One person (ID40) wrote at length about the importance of publicising events such as these and setting up a buddy system to encourage hard of hearing people to attend similar events. She spoke of her own experience of such a system, which had given her confidence to socialise more and she also suggested a number of events which would benefit from access provision such as this.

Paul used many gestures as he spoke, which was not something I had anticipated when writing the information for presenters. A number of audience members commented on this, saying that they risked missing them whilst following the subtitles. As a result of this, I later expanded the guidance to incorporate visual gestures as well.

7.6 Post-event focus groups – a thematic analysis

Reflecting back on all the events, it is clear that audience members experience each event in a variety of different ways. Engagement with the content of an event and the presenters is clearly important and in research events such as these, it is quite possible that people will attend with an interest in subtitling, rather than the explicit content of the

event. Nevertheless, most people who attended also found some use in the subtitles, even if they did not rely on them entirely.

In order for the subtitles to be fully embedded into the event, there are more requirements than accurate content alone. The focus groups that followed the events provided an additional opportunity for those attending to share their thoughts about the subtitling process and other aspects of event access. It was an opportunity for them to talk directly with the respeakers and to better understand some of the processes involved.

A number of themes came up in these focus groups. In many cases it was an opportunity to discuss particular features of the event more closely and ways that the subtitling could be improved technically. For example, the flickering of colours as handovers took place, the lack of a screensaver and the need to insert lines between speakers or to indicate paragraphs were all raised in this way. Points such as these could be quickly incorporated into the events that followed. Similarly, challenges related to the specific content of events were also mentioned. The humour of event 2 and the challenge that this presented the respeakers was noted by a number of people.

Many audience members were interested in the process of respeaking and this was discussed at a number of events. Many compared the subtitling of the events to that seen on television, often saying that the event subtitling was better. Since the respeakers were themselves television respeakers, this often led to a discussion about the challenges involved in working in a live television environment and what information respeakers have access to as they prepare for programs such as the news. All in all, these discussions gave the audience members a better understanding of what the role of the respeaker involves. In those events where the respeakers were present in the same room as the event, and where their muffled tones could be heard, the effort involved became even clearer.

Those present were interested in understanding some of the glitches that were seen. The double letters were discussed, as was the importance of having good technical equipment (event 3). Many people also enjoyed challenging the respeakers, suggesting and saying words to see how well they would be recognised. At event 8, where many

people were translators, they wanted to know how the vocabulary was programmed in Dragon, asking whether collocated words would be recognised more easily. They were also interested in the application of Dragon and whether, for example, it could be used directly by lecturers. Elsewhere, the theme of machine translation came up. Phone relay systems and other uses of subtitles were also discussed in event 3. Others were interested in the costs involved in setting up events such as these, as they wished to use such a service in future events of their own.

The focus groups also highlighted what features gave the events quality from the perspective of the audience. Although opinions will always vary, it was clear that continuity of the subtitles stream was vital. As researcher, and coming from a background of television subtitling, my focus was perhaps on the accuracy of the content. However, in practice, audience members do not focus on this in isolation. Their outlook is more holistic since they are taking part in the event in question. Perhaps this is one of the most fundamental shifts when moving from television to live events, or it is a reflection on how, to date, quality has been assessed in the television setting. All too often the subtitles are, in effect, separated from the programme itself. At a live event, engagement becomes more important, as does the ease with which subtitles can be seen.

In a number of focus groups, the discussion moved from live subtitles to those used in pre-recorded content, either at the theatre or on television. Whilst this was not directly relevant to the research, it highlights the importance of having opportunities for audience members to engage in discussions about a tool they use so frequently for access purposes. Many were interested in the use and expansion of creative subtitles and the introduction of caption glasses at theatres. Others reiterated points that had been raised in the focus groups, for example their frustration when subtitle obscured key content.

The topic of accents was also raised a number of times. In event 3, this led to a long discussion about how accent is captured in written form and what different users wish to be included. Since language is part of a person's identity, it was a clearly sensitive issue. Nevertheless, due to the challenges already present in respeaking, indicating accent can be difficult for the respeaker and also risks errors being created, as non-standard words

(from the perspective of Dragon’s vocabulary) may be needed. For this reason, in Chapter Eight, I suggest the addition of a subtitle introduction at live events as a possible solution to this. This would act in the same way as an audio introduction that accompanies an audio description (Fryer, 2016: 155). Whilst not translation per se (Greco and Jankowska, 2020), it supplements the content of the translation of the event and provides more grounding for those present.

A number of other access considerations also came up during these discussions. On the one hand, people present were interested in the experience of people who did not regularly use subtitles and they suggested user groups that might find them helpful, including dyslexic people and people with tinnitus. At the same time, other issues that would need to be addressed for a fully accessible event were also raised. These included the need for designated seats in order to watch subtitles, and the need to educate the hearing audience about why captions or subtitles were being used. More specifically, one person referred to the need for box offices to warn “caption haters” (post-event chat, appendix 7.8) that subtitles or captions would be in use. Although this was not frequently mentioned, it did seem that some audience members had encountered negative attitudes towards the provision of access.

During the focus groups, suggestions were also made about what information to share with the audience before an event made accessible with live subtitles. Delay, the muffled sound of the respeakers and the range of equipment in use plus the reason why access was being provided were all raised. In line with this, at events 3 and 7, the question of whether subtitling experience was better or worse if the respeakers were present in the same room was raised. At event 3, it seemed to be the case that the subtitling improved when they were in the same room.

Other improvements to how access could be embedded in the events they had just attended were also offered. Those present on museum tours would have liked more time to get used to the devices and to understand the range of display settings that were available to them. There was also a great deal of discussion about the role of the presenter and what they could do and what could be done for them to make their role

easier in an event made accessible with live subtitles. I have incorporated the suggestions into the section that follows.

7.7 Presenter training

Having experienced being respoken myself during the DH focus groups in Cycle One of the research, I knew that the presenters would need some information about being respoken in advance of the event so that they could be better prepared for the experience.

With this in mind, I created a presenter information sheet, which explained what respeaking was and how best to prepare for the event (appendix 6.2.1). I gave some tips for good presentation (section 6.7.7 above), and also explained what information would help the respeakers prepare for the event. The pre-event meetings between presenters and respeakers were also designed to support this process.

Whilst this information and the meetings were useful, they were not enough to mitigate the impact of being respoken and, in practice, many presenters still seemed unsettled at the start of the event. It was only some way into it that they seemed to find their feet and acclimatise to what respeaking and being respoken involved. Once they had passed this confidence threshold as it were, they were able to work with the respeakers, begin to sense what content might be challenging, and accommodate delays within the subtitles appearing. Whilst it was very good that this acclimatisation happened, it would have been preferable for all involved if it had happened before the event began and it is for this reason that I have outlined what presenter training would involve. I must also add that suggesting this training is no criticism to any presenter. They all coped admirably in what was a new and incredibly challenging situation for them, all the more so as it was within the context of research, and action research at that.

7.7.1 What would a training programme for presenters look like?

The training programme as I envisage it would consist of five key modules: Know your audience, Being respoken, Communicating with the respeakers, Handling different types of audio and visual content and Preparing for an event.

7.7.1.1 Know your audience

Many presenters said they would have liked to know more about the audience they were presenting for and the first module addresses this.

Deaf awareness training enables presenters to be more aware of potential d/Deaf audience members at an event. Many good courses exist, and the hope would be one could be run alongside this training. Typical courses include communication tactics, understanding the importance of lighting, eye contact, the need for one person to talk at a time, having a clear mouth and line of sight, having non-distracting backgrounds, and being patient (for example, Deaf Umbrella, no date).

Since it is likely that most events will be attended by a range of people, it will also be helpful for the presenter to be aware of the heterogeneity among both DH audience members, and the ways in which the wider audience may use subtitles. Different audience members will have specific access needs, and it would be useful for the presenter to be aware of these. Whilst they may not tailor the event content to meet every one, knowing they exist will most likely make encountering different situations less stressful.

At the same time, it is important that presenters are aware of the bigger picture of access and where their contribution fits within it and the chain of access (Greco *et al.*, 2012) that, hopefully, runs through the organisation of the event as a whole. With this awareness, they have an opportunity to fulfil their role as fully and effectively as possible.

7.7.1.2 Being respoken

When someone is being respoken for the first time and watching the subtitles appear on screen, a very common reaction is to stop talking to allow the subtitles to catch up with what has been said. When I was respoken at the DH focus groups, my gut reaction was to do exactly that, but I knew that if I stopped speaking, there would be no new content generated for the respeakers to continue to respeak and the subtitles would come to a halt. In this situation, instead of pausing entirely, a brief pause would be more effective

in allowing the subtitles to catch up and latency to be reduced. More than anything, it is this skill which the presenters would benefit from practicing before the audience arrive. In order to truly get a feel for how to work with the respeaker this way, the presenter should ideally have the opportunity to be respoken, but also to watch another presenter being respoken so that they can see the impact of different types of presentation styles and pausing for themselves. Ideally, DH audience members would be involved in this training so that the presenter can receive authentic feedback. This was discussed at length in the focus group that followed event 7 and many audience members there expressed an interest in being involved in such training.

This would also allow the presenters to acclimatise to the equipment that is being used, whether it involves flashing headsets or Darth-Vader-like masks, as they were described in the DH focus group sessions.

7.7.1.3 Communicating with the respeakers

All the presenters were aware that there might be a moment in the event where they would be asked to repeat something that the respeakers had missed. Although this happened rarely as the preparation material supplied was very detailed, many presenters were waiting for it to happen. Their feedback suggested that they seemed a little nervous about how they would handle it. In addition, knowing that this might happen also stopped some of them from speaking more freely and veering away from the script when they might have liked to.

Having received this feedback, I realised that what I had intended as a helpful comment, was actually impacting on how the presenters interacted with the audience (via the subtitles).

Having the opportunity to be respoken would also enable the presenters to get used to receiving communication from the respeakers and/or designated access co-ordinators. They might be asked to repeat content or spell a word, or be reminded to repeat a question from a member of the audience. Similarly, this could be an opportunity to expose the presenters to the longer delays that might accompany a handover, so that if

they do occur during the event, the presenter is less phased by this temporary drop in subtitle feed.

7.7.1.4 How to handle different types of audio and visual content

It is also important that the presenter understands how to work with the particular content that they are using. Slides, works of art, artefacts, videos, actions, gestures, questions and various interactions, all require different approaches when an event is being respokey and the way in which different events make use of these elements is likely to vary.

Getting a feel for how much time to leave in these different circumstances is an important skill for the presenter to develop, and it is likely to require, once again, that the presenter is able to both be respokey and watch others being respokey in these different scenarios to understand what does and does not work and why.

Exposure to good and bad demonstrations of this will also highlight the effort that may be involved for the audience. For example, seeing a densely packed slide, with additional information contained in scrolling subtitles should highlight the importance in spreading the content of a wordy slide more skilfully. If a video is being used that is already subtitled, seeing a sudden shift to a very small font will visually highlight the challenge the audience may face or even the fact that the video content, although subtitled, is in fact inaccessible. Similarly, a work of art, such as a large painting in a gallery, is very different from a slide, and more time should be provided to allow the audience to get a feel for it before any explanation begins. As more information is added, or particular details are highlighted, every audience member should have the time to look at those details and think about them, and not simply read the subtitles.

When questions are asked, the respokeyers require access to the audio feed, so the presenter may need to repeat what they have been asked, especially if the respokeyer is connected by phone. At the very least, during a seated event, the presenter will need to ensure the question has been directed into a microphone. At the same time, the presenter should remember to leave a pause after asking a question, so that all the

people in the audience have the chance to respond to it as there will be a delay between access to the question through hearing it and reading the subtitles; this may also be of benefit to people who would prefer longer to reflect on what was asked.

Out of all the modules, it is this one that will be the most tailored to the event type, since the fundamental understanding in the module relates to knowing the impact of speed and pauses on respoken content and audience experience and when to use each one.

7.7.1.5 Preparing for the event

The final section relates to how to prepare for the event. It may well be that this module involves a discussion between respeakers, venue staff and presenters. Questions such as what vocabulary items need to be shared may be raised, since there is often a misconception that long and difficult words cannot be respoken. In fact, they are often recognised more easily than shorter words, as long as they are in Dragon's vocabulary. Rather, it is names, places and proper nouns or obscure technical terms that the presenters should be alert to, so that they can supply these terms to the respeakers. In addition, respeakers may need to know what visuals are being used. If a video will be played, will it have burnt-in subtitles or will it need to be respoken? If an artefact is included in a museum tour, the respeaker should ideally have visual access to it in advance (even as an online image). Access to any audio description of it may also be helpful, as some terms contained within it may also be respoken.

It is also very important that the respeaker be aware of how the event will be organised and of any equipment the presenter will need. If the presenter needs a clicker, will there be enough USB ports alongside the Text on Top dongle? The presenter does not need to make these decisions, but by informing the respeaker of these considerations, the respeaker, in their role as access co-ordinator, will be able to ensure these technical aspects can be handled effectively and avoid complications on the day, which will benefit everyone involved.

Where the subtitles are being displayed alongside PowerPoint slides, there may be an area which must be left blank to house the subtitles. Either way, following good practice

when creating slides, in terms of font and the amount of information contained, will certainly be helpful for any audience member.

Finally, the presenter can help in preparing the audience for any complications that do arise on the day. Hopefully with this added communication and understanding between the presenter and respeaker, these will be minimised, but if the presenter is relaxed and confident throughout, it is more likely the audience will be as well. When the presenter is unnerved, the audience will very likely be too. Although it will be the respeakers that ultimately take care of issues such as these, presenter confidence is key as they create the initial dynamic of the event and set the tone.

7.7.1.6 Access for all = training for all

Live events are dynamic and they are affected and shaped by all who participate in them. Making events accessible does not need to be difficult, but for good access to be provided, planning needs to begin long before the event. For events to be successful, everyone involved needs to understand what is happening and why. The focus here has been on training for presenters, but in fact everyone involved in hosting and attending live events would benefit from understanding how access through respeaking is provided. Hands-on experience is key. Having experience in understanding is beneficial for everyone involved. This comes back to the chain of access (Greco *et al.*, 2012).

Ideally, the people who are (most) involved should have practical experience of seeing respeaking in action and being respoken. Every group involved begins to take on a new role if this approach is followed: the respeakers need to be able to act as access coordinators, presenters need to be able to adapt to being spoken. Venues need to understand the process of respeaking and how to accommodate it at events. Finally, audience members need to understand how and why access is being provided. In fact, they themselves become presenters as soon as they speak and contribute to the event in question.

Since we are discussing in-person events, the training that I have just outlined should ideally be run in-person as well, since its aim is to provide a realistic pre-experience of

the event that will follow; at the very least modules one and four should be run this way, since this is where the interaction with the d/Deaf, deafened and hard of hearing audience members is key. Later sections might be completed online or via correspondence, but the live experience of being respoken is key. Where possible, module four could usefully be run in the same venue as the event will be held.

7.8 From process to product

In this chapter I have reviewed the two rounds of research events where respeaking was tested in action and shared the experiences of those who attended, from audience members to people at the venues, from presenters to the respeakers themselves. Whilst individual experiences and preferences naturally varied, general patterns could nevertheless be detected that suggested which conditions are more and less favourable when the provision of respeaking, and access in a broader sense, at events are considered.

The planning and preparation of each event are key, as are the decisions about where respeakers are placed, what their working environment is like and where and how subtitles are displayed. All of these factors are likely to contribute to a better planned event and a higher quality of live subtitles being produced. In the next chapter, the NER analysis will allow a closer examination of the quality of subtitles seen at each event, and at the end of the chapter, this experiential audience score will be compared to the NER calculations to see the extent to which they align. Latency calculations for each event will also be examined.

The dynamic nature of live events means that the interaction between the different parties present may also impact on the content of, and consequently quality of, the subtitles provided. In this chapter a training programme for presenters to help them acclimatise to access through respeaking was outlined, and a number of recommendations were also made about how to prepare the audience for a respoken event. In the next chapter, the NERLE, a modified version of the NER for use at live

events is presented; this new model incorporates this interactivity of live events directly into the analytical pathway as quality is assessed.

Chapter 8: Analysis of respeaking at live events

“What measures the measurer’s measure?”

Zen Koan, Doug Duncan (Planet Dharma, no date)

8.1 Introduction

Over the last three chapters, different elements of the data collected in this study have been analysed. The needs and expectations of the different groups involved in live events were considered, a (revised) respeaker training programme was presented and the experiences of those present at each event were examined; in doing so, some commentary has already been given on the extent to which different needs and expectations have been met.

For providers, ensuring a quality of service was key, and for the respeakers in particular, a good working environment was complementary to this. For the audience, clear visuals and subtitles which captured the content of the event well, and which were displayed in good time, were vital. At the same time, we learned how beneficial it is when presenters understand the process of respeaking and begin to contribute towards its smooth running. The experiential analysis of the events in the last chapter highlighted those occasions where these different factors seemed to come together for all involved, leading to what might be taken as examples of well- or better-integrated access among the events, and case studies for future learning.

However, in order to fully answer the questions raised at the start of this thesis, the product, or the access provided, must also be explored.

1. What training and technical set-up is needed to allow experienced television respeakers to transfer their skills to the live event setting?
2. How can high quality respeaking and access be ensured?
3. Can the findings from this UK study be applied to the provision of access across borders and in society more broadly?

Analysing the quality of the subtitles seen, in this case the accuracy of their content and the latency with which they appeared, will allow a comparison between the respeaking seen at live events and on television, and with the industry standards established there. This assessment, taken alongside the scores seen in the previous chapter, should highlight those events where high quality respeaking and access came together, as posed in the second research question. Similarly, this should take us forward in our consideration of how the findings of this study can be applied more broadly to the provision of access (question 3).

In this chapter, the NER model used to assess the accuracy of the live subtitles at events is discussed and the corpus of respoken subtitles at live events is analysed. This corpus of respeaking at live events (RLE) is compared to the larger Ofcom corpus (Romero-Fresco, 2016), which stands as the main respoken corpus in the UK, and also the LiRICS corpus (Romero-Fresco *et al.*, 2019), established more recently during the pilot study for respeaking certification in the UK. Events 6 and 7, two of the highest-scoring events in the last chapter, are presented as case studies to illustrate the impact of the errors seen in context, and to explain why I have proposed the NERLE, the NER for Live Events, as a model for assessing respeaking in this new setting.

8.2 The RLE Corpus

The eight research events lasted a total of 440 minutes or 7.3 hours. In addition, focus groups across the three cycles of the study provided an additional 400 minutes of respeaking at live events. In total, 14 hours of footage was collected. Whilst not analysed in the same way, the research discussions around the events (the introductions, instructions and focus groups) also provided useful insights into the practices needed at live events and contributed to the action research.

The NER analysis that follows is based on clips of 10-16 minutes in length that were taken from seven of the events⁹⁴. This sample duration is in line with the clips used by Ofcom in their study on live subtitling, and subsequent studies such as LiRICS (*ibid.*), which is discussed below, and the NER-based quality monitoring in Canada (Nertrial.com, no date). The analysed corpus consists of approximately 9,000 words and 1,350 subtitles were collected across these seven events. Where relevant, examples from the broader corpus are also included in the discussion. The analysis can be found in appendix 8.2 and the clips for each event are in appendices 8.3.1-8.3.7.

When selecting the clips for this detailed analysis, care was taken to ensure that the segments chosen were representative of the event as a whole. A simple checklist was created for each event of the key features that had appeared within it and of learning points from an assessment or action research perspective (see table 8.1 below).

For example, in event 3, there were numerous transitions from the presenter speaking with an image on display, to playing a pre-recorded video. There were also multiple speakers at this event, each with a very different style and approach to the presence of the respeakers. The analysed segment contained all these elements. At event 4, one phrase proved particularly problematic for the respeakers to capture and this was included in the analysed segment⁹⁵. For most events, it was possible to choose a single segment that contained all these elements, but in the two events described above, it was necessary to combine two different segments of the event to capture them.

⁹⁴ It was not possible to apply this analysis at event 5 due to the poor quality of the audio recording.

⁹⁵ The respeakers were initially unable to capture 'Snow Hill' correctly, and it was repeated many times before it eventually appeared correctly on screen. This exchange can be seen in appendix 8.2, Event 4_Assessment, rows 105-116.

Table 8.1: Key features of each event

Event	Key features to include in sample clip
1	<ul style="list-style-type: none"> • Main speaker in flow to illustrate the fast pace • Contributions from additional speakers, including extended dialogues and exchanges with them • Some 'in character' speech from the Queen of Hearts • Use of slides, with text and image • Q&A from audience
2	<ul style="list-style-type: none"> • Slides with images and some text contained within them • Visual realia • Dense content • Humorous anecdotes (where the timing mattered) • Audience interaction (Q&A)
3	<ul style="list-style-type: none"> • Samples from each presentation (<i>Blue Pen & Notes on Blindness</i>) • Audience interaction • Use of (INAUDIBLE) and demonstration of the microphone issues experienced • Transition between video and PowerPoint slides • Interactions between presenters and respeakers
4	<ul style="list-style-type: none"> • Main speaker • Audience interaction (question from the tour guide and opportunity for the audience to ask questions) • Technical/specific terminology • Problematic name (SNOW HILL) and interaction with the designated access co-ordinator • Problematic delays
6	<ul style="list-style-type: none"> • Main speaker • Audience interaction • Use of BSL during the event • Examples of the audience checking how their words had been represented on screen
7	<ul style="list-style-type: none"> • Main speaker • Samples from different points in the event to illustrate how the presenter became more acclimatised to being respoken • Specific terminology • Examples of different types of error correction • Example interventions from the designated access co-ordinator
8	<ul style="list-style-type: none"> • Main speaker • Interaction with audience • Specific and specialised terminology • Use of gesture • Interaction with the designated access co-ordinator

8.3 Models for analysing respeaking quality

When we discuss the quality of the live subtitling product, we are in the realm of the quality of service (Greco and Jankowska, 2019, as discussed in section 6.2.2. above); in this instance, it is the accuracy of the written text, referring both to how much of the original spoken content is included and the precision with which it appears (Moore, 2020b), that is taken as the main indicator of quality and the NER score calculation is

based on this. The latency with which the subtitles appear is calculated, and contextual feedback is contained within the comment that accompanies the NER, but accuracy is the leading factor.

The NER model has been adopted for assessing the accuracy of live subtitling on television in the UK and in many European countries. It is also used in Canada, albeit in a slightly adapted form and its use is currently being examined in the US, with preliminary steps being taken to using it there too (Government of Canada, 2019; Federal Communications Commission, no date).

Before beginning the analysis, I first want to explore more fully what makes the NER the model of preference for analysing live subtitles, and whether this will remain true for the live event setting. To do this, where the NER model sits within the landscape of available models of analysis will also need to be understood.

8.3.1 The NER model as a preferential model for assessing live subtitles

What makes the NER model preferential to methods for assessing accuracy in live subtitling that came before it?

8.3.1.1 Requirements for a model

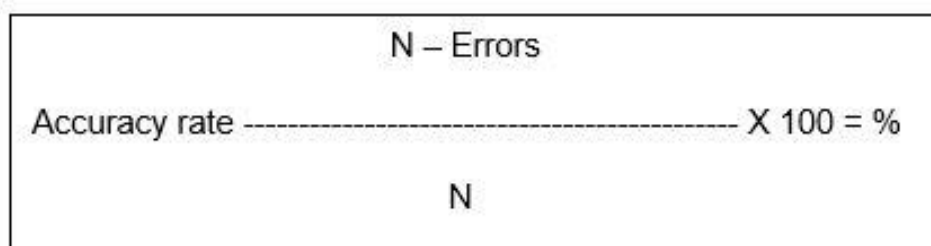
Romero-Fresco and Martínez (2015: 29-30) outline the requirements of a model for assessing live subtitles created through respeaking. The model must be easy to apply and must include an assessment of the respoken subtitles against the original content. Any model that simply calculates the number of errors within a given number of words loses sight of the contextual nature of these errors, and is therefore unable to comment on the actual impact of an individual error on the audience. Very importantly for respeaking in particular, given the use of editing (section 5.3.6.5), the model must also acknowledge and be able to account for the fact that there may be differences between the respoken text and original audio which are not errors and therefore carry no penalty. Such differences might involve a synonym being used, contracted forms, or certain

phatic language being omitted, which, in most cases, will mean that the general sense of the content remains the same even where the exact words used differ.

In addition to incorporating these elements into the score, the model should also consider the impact of any error on the audience, which is highly relevant for the current study⁹⁶. To do this, contextual information is also important and information beyond the accuracy of the subtitles should be provided. This might include delay, position, speed and the use of paralinguistic features (for example, those outlined in section 6.10 above)⁹⁷. The model needs to account for the fact that errors can be caused in different ways, and, as a final point, it should also be informative and provide an element of feedback for the respeaker.

8.3.1.2 Models which preceded the NER

Traditional word error rate (WER) models calculate the percentage of either word correctness or word accuracy by deducting the total number of errors from the total words spoken by a user (N) (Romero-Fresco, 2011: 151-152)⁹⁸, as shown in figure 8.1.



The diagram shows a rectangular box containing the formula for WER calculation. At the top center is the text "N - Errors". Below this, a horizontal dashed line spans the width of the box. Underneath the dashed line, the text "Accuracy rate" is on the left and "X 100 = %" is on the right. At the bottom center of the box is the letter "N".

$$\frac{N - \text{Errors}}{N} \times 100 = \%$$

Fig. 8.1: WER calculation, adapted from Romero-Fresco, 2011: 151

Dumouchel *et al.* (2011) illustrate the range of errors that each measure would include (Fig. 8.2):

⁹⁶ Recent research in Poland and Canada has found positive correlations between the NER scores and users' view of the quality of live subtitling output (Romero-Fresco, 2016, 2020; Szarkowska *et al.*, 2018; CRTC, 2019).

⁹⁷ The range of information that might be included for live events extends even further (section 8.6 below). Mercer-Mosel's (1996) model for interpreting referred to in section 6.2.2 above provides a good guide for this

⁹⁸ This distinction is made by the US National Institute of Standards and Technology, referred to in Romero-Fresco (2011: 151).

Where	is	the	whole	wheat	flour
		D	S	S	I S
Where	is		hole	we	eat flower

Fig. 8.2: Deletion, substitution and insertion errors, adapted from Romero-Fresco, 2011: 151

Here, the errors indicated are created by deletion (D), where a correct word is omitted, substitution (S), where the correct word is replaced by an incorrect one, and insertion (I), where an unspoken word is added. The new total (N – errors) is divided by the original total (N) and multiplied by 100 to give the percentage accuracy score.

If word correctness is the focus, then deletion and substitution errors alone are considered (Fig. 8.3):

$$\text{Accuracy rate} = \frac{N(6) - D(1) - S(3)}{N(6)} \times 100 = \frac{2}{6} = 33.3\%$$

Fig. 8.3: Calculating word correctness from deletion and substitution errors

However, to assess the accuracy, inserted words must also be considered, giving an even lower score (Fig. 8.4):

$$\text{Accuracy rate} = \frac{N(6) - D(1) - S(3) - I(1)}{N(6)} \times 100 = \frac{1}{6} = 16.7\%$$

Fig. 8.4: Calculating word correctness from deletion, substitution and insertion errors

As we will see below, applying this method to a text where respeakers have made many edits may lead to low scores, whilst a model that accounts for correct editions where meaning is not lost, as the NER does, may give scores of closer to 100%.

This potential discrepancy was recognised, and one of the first solutions to it was proposed by the Centre de Recherche Informatique de Montréal (CRIM). Whilst the calculation remained the same, they adapted the model so that a human assessor would review the scoring after the transcripts were aligned, to ensure that any deletions that were penalised had in fact caused a loss of information. Whilst this did allow for some flexibility over editing, the deletion figure remained “ambiguous” (Romero-Fresco and Martínez, 2015: 31) and lacked specificity. It did not reveal the reason behind the deletion, whether as a result of good or bad editing or misrecognition, and it did not account for any corrections by the respeaker which may have followed.

The weighted WER (Apone *et al.*, 2010; Romero-Fresco, 2016) offered a solution which could account for the impact of different kinds of errors, but as an automatic model, with the human judgement of the CRIM removed, a detailed analysis and evaluation of context was less likely. In addition, this model was more focused on STTR output, which meant the potential edition rates of respoken subtitles might still pose a problem.

It is the NER model that seemed to provide a full solution to these different problems.

8.3.2 The NER workflow

The NER model depends on an accurate comparison of the original spoken content of the material being made with the respoken transcript. The assessor will compare both the words spoken and the meaning of the original with what is actually transcribed for the viewer to read and access. Figure 8.5 shows the process of using the NER model to assess the quality of live subtitles as a flowchart.

In order to calculate the accuracy, first, the total number of words in the transcript must be known. However, rather than this being the number of words visible on screen, it is the number of words that have been respoken. This at once reflects the effort being made by the respeaker and their ability to capture the content of the original: saying the

command 'full stop' takes up time where another word could have been spoken. Similarly, commands may themselves be omitted or lead to recognition errors; therefore, for the scoring to be coherent, commands must also be included within the word count. While the respoken transcript does have a lower word count than the spoken original, the total number of words actually spoken by the respeakers when these commands are included is closer to that of the original speaker.

The possible error types are split into two broad categories - edition errors, where the respeaker omits, adds or alters the original text, and recognition errors, where the respeakers' words are correct, but there is an error in them being processed by the software. Within these, there are further subcategories, where the severity of the error is considered. Serious errors carry a penalty of 1, standard errors 0.5 and minor errors 0.25. Since the model is itself user-focused, this severity ranking is considered from a users' perspective and considers the extent to which the meaning of the original is lost or changed as a result of the error.

The scoring of edition errors centres around idea units. An idea unit is a group of words which contain an idea and this often, but not always, corresponds to the main clause in a sentence. The key point is that the idea referred to here is the main or central idea that the utterance is expressing.

The NER compares spoken words to written sentences, so it must be remembered that punctuation alone does not indicate where idea units begin and end. A longer sentence may, for example, be composed of a main idea and additional clauses which express dependent ideas – the who, where, why and how. However, it may also contain two idea units, joined with a conjunction such as 'and', so discretion must be used when determining how to score any omission.

Using the NER scoring system, omitting an independent idea is a standard error, whilst omitting a dependent unit or idea is a minor error. A serious error would be one where the user is misinformed in such a way that they would not realise they have been misinformed without listening to the original spoken content. Consider the example below:

The Drury Lane Theatre was built in 1802 by Jon Miller and Richard Jones, who both lived nearby.

The main idea here is that the Drury Lane Theatre was built in 1802. The dependent ideas include the two builders and the fact they lived nearby, totalling three dependent pieces of information.

If the main idea of the theatre being built was missed out, 0.5 points would be deducted – even if the information about the builders was included, it would make no sense without the situating context of the main idea. If the main idea was included, but one or more dependent ideas were missed out, 0.25 would be deducted for each point accordingly. If the entire sentence was skipped, however, a total of 0.5 would be taken off.

Misleading errors might at first sight seem negligible, but in fact, their impact is large and markers must be on the alert for them when scoring. Take the following example:

She agreed to concede.

The sentence looks accurate and appears to make sense, but if the actual context the sentence appears in is considered, it becomes clear that the sentence is misleading, and should be scored with a penalty of 1.

Trump and May were arguing. He agreed to concede.

In some situations, the misrecognition of 'he' as 'she', might be a minor error, as the viewer could work out who was actually being referred to. However, in this instance, the error is serious as the viewer is misinformed. Simply reading that sentence, the viewer would not know they had been misinformed and that is the reason for the serious classification and score.

In the case of recognition errors, minor errors are ones that the reader can easily determine. 'I have to dogs' is incorrect, but the reader is able to determine what the sentence means.

Standard errors occur when the user recognises what appears on screen as an error, but perhaps struggles to work out the intended word. It may be that the viewer needs to read the sentence a second time to understand it, or that they are unable to, so a gap remains in the information provided.

There was snow on the pizza the mountains.

[There was snow on the peaks of the mountains.]

The serious error penalty is reserved for misrecognitions which mean that the reader is misled. For example, one number might be misrecognised as another; in the sentence "The fee was \$15,000 dollars", whilst \$15,000 was produced by the SR software, \$50,000 was the number respoken. Once again, the viewer would not know any differently.

It is important to note that the seriousness of the error is not a judgement of how important a fact is, but an acknowledgement that the viewer has been misinformed. Assessing in this way avoids making a judgement on what value of a piece of information holds for another person but instead reinforces the notion of respect, outlined in Chapter Two, that is central to the relationship between the respeaker and their audience. A great deal of responsibility lies with the respeaker and in the words they produce. It is this, alongside the constraints of working in real-time, that makes respeaking high pressure work. Being as faithful as possible to what has been said is essential in order to achieve high quality access, but also because that access to information is a right. Knowing how and when to edit respoken text is a central part of training and is also an added element of the work the respeaker is doing while they respeak.

When editions do not remove key content from the subtitles, they are referred to in the NER model as 'correct editions' (CE). They do not alter the final score for a piece of

respeaking, but they are counted alongside other errors, to indicate what has been involved in the respeaking of a passage. Whilst errors are noted with precision, the CE score is a more general one. Omitting 'and'/'so'/'because' or switching 'also' and 'too' may count as a single correct edition, but so would omitting the sentence 'And she said yes,' if the speaker says it three times, and the respeaker only respeaks it twice. Therefore, the CE score should be read as an indicative rather than precise score. Hesitations which occurred at the live events were all transcribed, and so contributed to the CE scores; however, when it came to calculating the speaking rate in the clip and lexical density (see section 8.5 below), hesitations were not included within the word count.

The accuracy score that the NER reveals gives an overall impression of the level of the respeaking. However, it is the comment that comes alongside it that provides the full context of the situation and allows a more detailed reflection on the subtitles as part of the full audiovisual content of the programme. Similarly, if ever there were to be a discrepancy between the accuracy score and final comment, for example if an exceptionally high subtitling speed meant the subtitles were not legible, this final comment would outweigh the calculated score (Romero-Fresco and Pöchhacker, 2018: 159). With this comment, the NER becomes both a qualitative and quantitative judgement of the access being provided, very much in line with the approach I have adopted in this study, mixed methods tilted towards the qualitative side. In the case of live events, in particular within this action research setting, this is also an opportunity to capture the dynamism of the setting, and comment on any missed opportunities for communication or interaction between the respeakers and the accessibility contact.

8.3.3 NER Scoring

The allowance for correct editions, the audience-focused weighting of errors that must be penalised and the contextualisation of the feedback means that the NER has been adopted as a highly specialised model of choice when calculating the accuracy of live

subtitles. The following example highlights this suitability, in comparison to the traditional WER scoring outlined above:

Original text	Subtitle	Comparison
Well, you know, you have to try and put out a good performance, I mean, yeah, it's kind of a stepping stone, isn't it, really?	You have to try to put out a good performance, It's a stepping stone.	Well, you know, you have to try and put out a good performance, I mean, yeah, it's kind of a stepping stone, isn't it, really?
WER SCORE	N - D - S - I	25 - 11 - 1 - 0
	----- x 100 = ----- = 52%	
	N	25
NER SCORE	N - E - R	16 - 0 - 0
	----- x 100 = ----- = 100%	
	N	16

Fig. 8.6: A comparison of NER and WER scoring, adapted from Romero-Fresco, 2016: 58

The reliability of the weighted scoring system is underlined by the very low disagreement that is seen when the same script is marked by more than one trained NER evaluator. In the Ofcom study, this interrater disagreement was reported at 0.09%, based on the average scores given across scripts by first and second markers, indicating a minimal difference between the accuracy score a script would receive (Romero-Fresco, 2016: 63).

When calculating accuracy with the NER, a benchmark of 98.0% has been set as the minimum acceptable level of accuracy. Anything falling below that level would be considered as poor. Subtitles within the range of 98.0-98.5% are considered fair, between 98.5-99.0% good, 99.0-99.5% very good and above 99.5% excellent. More recently, these ranges have been converted to a more standard 10-point scale (table 8.2) (Romero-Fresco and Pöchhacker, 2018: 160)⁹⁹.

⁹⁹ The 10-point equivalencies for the data that follows can be found in the overview of the NER analysis chart in Appendix 8.2.

Table 8.2: The 10-point conversion scale for NER accuracy (Romero-Fresco and Pöchhacker, 2018: 160) alongside the traditional accuracy ratings

Accuracy (%)	10-point scale	Accuracy rating
<96	0/10	Poor
96.4	1/10	Poor
96.8	2/10	Poor
97.2	3/10	Poor
97.6	4/10	Poor
98.0	5/10	Fair
98.4	6/10	Fair
98.8	7/10	Good
99.2	8/10	Very good
99.6	9/10	Excellent
100	10/10	Excellent

8.4 NER Analysis: Overview

Before commenting on the spread of errors seen at the events, one further note is required. Two different calculations were made for the accuracy at each event due to the large number of software-specific errors that were seen (Moore and Romero-Fresco, 2015). The majority of these errors were due to software glitches, whereby the spaces between words were duplicated (by far the most common), removed, or the initial letters of certain words were doubled (section 6.5.2.1 above).

At the mobile events, software-specific errors also occurred when the respoken content was not transmitted in full to the tablets being used by the audience. All these errors were classified as recognition errors. The first accuracy score includes these errors in the count (table 8.3). A revised accuracy score (table 8.4) was also calculated to determine what the accuracy would have been without these software-specific errors, as the glitches were not considered permanent. I use an asterisk (* or RLE*) to indicate the revised scores. Table 8.5 compares these two data sets.

8.4.1 Accuracy at the research events

Table 8.3: RLE Accuracy data (including software-specific errors)

Event	Round One				Round Two		
	1	2	3	4	6	7	8
Edition Errors:							
Serious	1	2	3	2	2	3	1
Standard	11	6	17	20	20	11	7
Minor	24	17	19	22	22	12	23
Total	36	25	39	44	44	26	31
Recognition Errors:							
Serious	0	0	0	2	1	1	0
Standard	0	1	0	3	0	6	1
Minor	71	107	111	43	113	16	112
Total	71	108	111	48	114	23	113
Accuracy	98.2%	97.3%	98.1%	97.8%	97.3%	98.7%	98.5%
Rating	Acceptable	Substandard	Acceptable	Substandard	Substandard	Good	Good

Table 8.4: Revised accuracy data (RLE*) (excluding software-specific errors)

Event	Round One				Round Two		
	1	2	3	4	6	7	8
Edition Errors:							
Serious	1	2	3	2	2	3	1
Standard	11	6	17	20	20	11	7
Minor	24	17	19	22	22	13	23
Total	36	25	39	44	44	27	31
Recognition Errors:							
Serious	0	0	0	2	1	0	1
Standard	0	1	0	0	0	0	1
Minor (ex.)	9	27	17	23	41	15	18
Total (ex.)	9	28	17	25	42	15	20
Revised accuracy*	99.1%	98.8%	98.9%	98.3%	98.4%	99.0%	99.1%
Revised Rating*	Very good	Good	Good	Acceptable	Acceptable	Very Good	Very good

Table 8.5: RLE and RLE* accuracy data compared

	Round One				Round Two		
Event	1	2	3	4	6	7	8
Accuracy	98.2%	97.3%	98.1%	97.8%	97.3%	98.7%	98.5%
Rating	Acceptable	Substandard	Acceptable	Substandard	Substandard	Good	Good
Revised accuracy*	99.1%	98.8%	98.9%	98.3%	98.4%	99.0%	99.1%
Revised Rating*	Very good	Good	Good	Acceptable	Acceptable	Very Good	Very good

The accuracy rates seen are as might be expected for professional, experienced respeakers transferring to a new setting. They range from substandard to good, when software errors are included, and good to very good, when excluded. The highest accuracy scores, both including and excluding errors, were seen in the final two events, where the respeakers had grown more accustomed to this new setting. However, the revised score at the first event was comparably high and variation across event type is also visible, as will be discussed in section 8.5 below. This illustrates that to fully understand the NER score given, it must be interpreted alongside the context of the event.

8.4.2 Comparison between respeaking corpora

Many similarities can be seen when the data collected at live events is compared to the Ofcom and LiRICS corpora. As seen in table 8.6, the average accuracy rate at live events including errors is 98.0% (acceptable), slightly lower than that seen in the Ofcom and LiRICS corpora; when software errors are excluded, the average rating rises to 98.8%, an average that is higher than that seen in Ofcom and LiRICS and which narrowly misses a 'very good' rating.

Table 8.6: Summary of results for the RLE, the RLE*, Ofcom and LiRICS pilot

	RLE	RLE* (excluding software errors)	Ofcom	LiRICS
Average accuracy rate	98.0% (Acceptable)	98.9% (Good)	98.4% (Acceptable)	98.5% (Good)
Excellent subtitles	0%	0%	5%	7%
Very good subtitles	0%	43%	22%	30%
Good subtitles	29%	29%	29%	33%
Acceptable subtitles	29%	29%	21%	7%
Substandard subtitles	43%	0%	23%	23%

In table 8.7, the errors that occur are split according to whether they are edition or recognition errors. Amongst professional respeakers, the common trend is for there to be a higher proportion of edition errors than recognition errors, as the data from Ofcom shows. One key reason for this is the challenge that keeping up with the original audio poses. At live events, respeakers will be working in unfamiliar settings, so a similar trend could be expected.

Table 8.7: Total edition and recognition errors

	RLE	RLE* (excluding software)	Ofcom	LiRICS
Total edition errors	29%	61%	69%	65%
Total recognition errors	71%	39%	31%	35%

In the original RLE data, this is reversed and the errors are 29% edition, 71% recognition. However, once the software-specific errors are excluded from the RLE data, the same trend as in the Ofcom and LiRICS data is seen. It seems that the respeakers' attention is, primarily, going towards editing. The number of recognition errors in the RLE corpus

is higher than that in Ofcom and LiRICS. This can be attributed to a number of capitalisation errors which resulted from the settings used when macros, or voice commands were created. Initially, these were considered to be software-specific errors and beyond the control of the respeakers. It was later realised that these errors were preventable and it is likely that when they are corrected for at future events, the percentage spread of edition and recognition will be even more aligned across the three corpora.

For reference, tables 8.8 and 8.9 show how the total number, as well as percentage of edition and recognition errors, varies across the RLE and RLE* corpuses.

Table 8.8: RLE: Total number and percentage of edition and recognition errors

Including software-specific errors (RLE)								
	Round One				Round Two			Error type (%)
Event	1	2	3	4	6	7	8	All events
Edition Errors:								
Serious	1	2	3	2	2	3	1	
Standard	11	6	17	20	20	11	7	
Minor	24	17	19	22	22	12	23	
Total	36	25	39	44	44	26	31	29%
Recognition Errors:								
Serious	0	0	0	2	1	1	0	
Standard	0	1	0	3	0	6	1	
Minor	71	107	111	43	113	16	112	
Total	71	108	111	48	114	23	113	71%

Table 8.9: RLE*: Total number and percentage of edition and recognition errors

Not including software-specific errors (RLE *)								
	Round One				Round Two			Error type (%)
Event	1	2	3	4	6	7	8	All events
Edition Errors:								
Serious	1	2	3	2	2	3	1	
Standard	11	6	17	20	20	11	7	
Minor	24	17	19	22	22	13	23	
Total	36	25	39	44	44	27	31	61%
Recognition Errors:								
Serious	0	0	0	2	1	0	1	
Standard	0	1	0	0	0	0	1	
Minor (ex.)	9	27	17	23	41	15	18	
Total (ex.)	9	28	17	25	42	15	20	39%

Where the RLE and RLE* corpora do differ from the data collected in Ofcom and LiRICS is in the seriousness of the errors seen. Table 8.10 shows the total of serious, standard and minor errors in each corpus.

Table 8.10: Total serious, standard and minor errors

	RLE	RLE* (excluding software errors)	Ofcom	LiRICS
Serious	2%	4%	5%	8%
Standard	12%	24%	39%	36%
Minor	85%	72%	56%	56%

In all these corpora, as the severity of the error is reduced, the number of errors seen increases. However, in the live event data, the proportion of minor errors is far higher than that in either the Ofcom or LiRICS data, whilst fewer serious and standard errors are seen. Whilst this trend at live events is an advantageous one, since minor errors

have less impact on the audience, further investigation is required to understand the reason behind it.

In Table 8.11, the errors are classified by severity and error type (edition or recognition).

Table 8.11: Total serious, standard and minor errors categorised as edition or recognition errors

	Edition				Recognition			
	RLE	RLE*	Ofcom	LiRICS	RLE	RLE*	Ofcom	LiRICS
Serious	5%	5%	2%	9%	1%	3%	7%	6%
Standard	38%	38%	42%	38%	2%	1%	31%	34%
Minor	57%	57%	56%	53%	97%	96%	62%	60%

The general spread of edition errors was comparable across all corpora, although there was a higher proportion of serious errors at live events than in the Ofcom study. There is no clear reason for the raised number of serious errors, so each must be evaluated within the context they occurred in.

The spread of recognition errors was very different in the live event corpus, where almost all errors were minor. One explanation is that collaboration between the respeakers and presenters meant that the respeakers were well-informed about the content of each event and the preparation time allocated enabled them to prepare their voice models well, thus avoiding many of the potential serious or standard recognition errors which might otherwise have occurred. Similarly, the speech rates experienced at these events were lower than those in many programmes seen on television, which might have prevented more serious edition and recognition errors.

Another explanation lies in how the respeakers corrected errors during the events. When respeaking on television, respoken text is pulled and broadcast automatically, so most corrections, indicated with (--), must follow any error made; 17 such corrections were noted across the seven events. The software used at the live events allowed the respeakers more flexibility when making corrections. The Streamtext software, used at the two mobile events (4 and 7) allowed respeakers to edit content that had already

appeared on the audience's screen. The respeakers did not need to use the (--) on these occasions, as the corrected version replaced the original. Video footage of the events provide evidence of these corrections. In the Text on Top interface, used at the seated events, the respeakers are able to view and edit the subtitles before sending them to the audience's screen and the respeakers reported making corrections in this interim interface. These were not recorded on screen or in the transcript. Since any correction adds to the latency with which a subtitle appears, experienced respeakers would tend to limit corrections to serious or standard errors, which would mean a higher proportion of minor errors appearing on screen, as is the case here.

8.4.3 Latency

Latency also impacts on quality and the experience of the audience at an event. The report from the fourth sampling exercise of the Ofcom study (2015a: 15) discusses latency range reported across the whole study. Whilst the overall latency seen in each round ranged from 5.1–5.6 seconds, reported at 5.4 seconds in round one, 5.6 in rounds 2 and 4, and 5.1 in round three, it was also noted that this was based on a mixture of 'pure respeaking' and semi-live respeaking where pre-prepared blocks were cued out. When pure respeaking alone was used, comparable to the respeaking in the current study, the average delay was 7-8 seconds, with peaks of 10-21 seconds. No data on latency was collected in the LiRICS pilot.

I followed the sampling method used in the Ofcom study (*ibid.*) and identified 2-3 words per minute and measured the delay between the word being spoken and appearing in the subtitle; the averages for each event, and across the events appear in table 8.12, and the full set of measurements can be found in appendix 8.2.

Table 8.12: Latency at each event

	Round One				Round Two			Total
Event	1	2	3	4	6	7	8	
Latency:								
Average	6.7	7.5	5.9	4.5	5.8	4.3	5.7	5.8
Low	2.3	4.4	2.3	3.1	2.4	2.8	2.3	2.8
High	17.4	19.7	13.4	9.3	11.7	9.6	13.9	13.6

At the live events, the average latency in the samples ranged from 4.3 to 7.5 seconds, with 5.8 seconds being the average latency across all the events. This is lower than that seen in the Ofcom study, but still above the rarely attained 3 seconds that is recommended¹⁰⁰.

Some of the peaks at the live events were also high, though they remained in line with those seen in the Ofcom study. The highest recorded peak was 19.7 seconds at event 2. What was noticeable at the live events is that the peaks were kept to a minimum. Sometimes the event itself facilitated this, by allowing the latency to be reset to zero, for example when video clips were played (event 3) or the tour moved on to a new object (events 4 and 7). At some events, the presenters began to monitor the subtitles of their own accord, occasionally pausing to allow the subtitles to ‘catch up’.

It is noticeable that the lowest average latencies were seen at events where the Streamtext software was used (4 and 7); here, subtitles are displayed directly onto the audience’s screens, without the interim interface seen in Text on Top. The NER score for these events differed, with event 4 receiving a score of 97.8% including software errors and 98.3% without, and event 7 receiving 98.7% with software errors and 99.0% without, scores which were reflected by the audience who ranked event 7 highly (81.0%) and event 4 as a medium-scoring event (77.2%); however, as will be illustrated in the next section, many other features of the event were very similar, including the speech,

¹⁰⁰ The report of the fourth sampling exercise notes that on six occasions, this was achieved when a mixture of pure respelling and blocks were used (Ofcom, 2015a: 15).

respeaking and subtitle rates, which were noticeably lower than at the other events where Text on Top was used.

These higher latencies seen with Text on Top may have been due to error correction or may have been the result of the respeakers having to remember to regularly cue out the text they had respoken. Whatever the cause, navigating this fine and important balance between accuracy and latency has a significant impact on subtitle quality.

The sampled clips, and consequently data on latency, do not include those moments when technical issues caused the subtitles to freeze completely and which notably happened at the two events where Streamtext were used. These freezes seemed to be linked to handover points between the respeakers. At event 4, these freezes were more evident and interrupted the flow of the guide (7.4.4 above), whilst at event 7, the respeakers minimised their potential impact by handing over as the group moved between artefacts. Incorporating this technical aspect of tours even more smoothly is something considered further in section 8.5.2 below and which is addressed in the guidance and training for both presenters and respeakers (section 7.7 above).

8.5 NER Analysis: Case Studies

In order to better understand the transference of this service into the live event sector, a closer consideration of the data collected at individual events is needed. As in Chapter Seven, case studies allow the respeaking to come alive, as errors, and the scores they led to, can be seen in context.

I have decided to use events 6, a post-film discussion, and 7, a museum tour, for this purpose; both were Round Two events, where the refined respeaking kit was used. Together, they allow the potential impact of varied event contexts on the accuracy of the respeaking to be considered more closely, and they provide the opportunity to examine possible reasons behind the variation in the scoring that these events received from the audience (7.3.2 above) and in the NER analysis.

For the audience, these events were consistently among the highest scoring. They received 84.5% and 81.0% respectively from the combined DH audience, and 80.5%

and 80.8% from the audience as a whole. In contrast, the NER scores they received were more varied. The revised RLE* data recorded a score of 98.4% (acceptable) for event 6 and 99.0% (very good) for event 7; when the software-specific errors were included, the ranking for event 6 dropped to substandard (97.3%) and to good (98.7%) for event 7.

In addition to data on the spread of errors (tables 8.7, 8.8 and 8.9 above), additional data will also be analysed, including speech, respeaking and subtitling rates, the reduction rate from the original content to the respoken content and the resulting change in lexical density¹⁰¹. This data can be found in table 8.13.

Table 8.13: Speed, reduction rate and lexical density at each event

Event	Round One				4 (Tablet view)	6	Round Two		
	1	2	3	4			7	7 (Tablet view)	8
Speech rate in clip WPM	143	135	136	123	123	145	112	112	133
Respeaking rate WPM	119	134	114	99	99	111	104	104	130
Subtitle speed WPM	104	119	102	86	84	99	92	85	116
Reduction rate in subtitles	27%	12%	25%	30%	32%	32%	18%	24%	13%
Lexical density of spoken transcript	27%	37%	38%	34%	34%	25%	38%	38%	32.6%
Lexical density of respoken transcript	32%	39%	42%	41%	41%	29%	41%	42%	33%

8.5.1 Event 6: Post-screening discussion

Although the event was highly rated by the audience, the accuracy of the post-screening discussion of *The Piano* at Depot in Lewes was the lowest seen. Given that this was one of the later events, a higher score than 98.4% (acceptable) when software errors were excluded might have been expected, yet it was at this event that the highest number of both edition errors and recognition errors were seen. What exactly does this score reveal about the access at the event and what was the impact on the audience's access to it?

¹⁰¹ Lexical density was calculated using the text analyser on the usingenglish.com website (<https://www.usingenglish.com/resources/text-statistics.php>).

As with the other events, as the severity of the error decreased, the number of errors increased. There were 2 serious, 20 standard and 22 minor edition errors and 1 serious, 0 standard and 113 minor recognition errors including software-specific errors; excluding software-specific errors, the number of minor recognition errors fell to 41.

The high number of recognition errors is in part due to the macro settings used at this event; there were numerous capitalisation errors that could have been avoided (for example, the one seen in table 8.14, segment 4 below). The high number of edition errors is reflected in the reduction rate in the event, calculated as the percentage of words from the spoken transcript that are omitted in the respoken subtitles. The rate at this event was one of the highest at 32% reduction, a rate more than double that seen at the events with public speakers (2 and 8), for example. Unlike those two events, event six was based on an audience discussion; whilst there was a chair who opened the event with a series of questions, the audience were equally responsible for determining where the conversation went. Consequently, the respeakers only had a very general notion of what content they would meet. As it happened, in the portion of the event analysed, the discussion moved from the storyline of *The Piano* to how music in films is made accessible. This trait of a high reduction rate is not unexpected at a chat/discussion-based event like this, and mirrors the rates seen in the Ofcom study for chat shows (31.6%) compared to the more heavily scripted news (22.5%) (Romero-Fresco, 2016: 66). Whilst overlapping speech was avoided at this event through the use of microphones, the speech rate at this event was one of the fastest and the conversational tone and informal exchanges provided many opportunities for editing, if the respeaker required it.

Table 8.14 contains segments of the NER analysis for event 6, which illustrate a range of serious, standard and minor errors. In segment 1, the sentence 'I know what's going to happen in a piano' is respoken as 'I know I'm going to hear any piano'. On the first reading, there is a difference in meaning between the two sentences and consequently, a penalty for a serious edition error was applied. However, in the recording, it is notable that the person who has spoken these words pauses and reads what has been subtitled,

Table 8.14: Examples of errors at event 6

Segment	Spoken	Respoken	Comment
1	I can hear a piano better than hear a human voice. I know what's going to happen in a piano. I get it right. But I don't know what the next word is going to be from someone else so I've always got that difficulty.	I can hear a piano better than I can hear a human voice. I note I'm going to hear any piano. -- I know.	1 Correction: --know [I note]; no penalty applied. 1 Serious edition error (misinformation): 'I know I'm going to hear any piano is not necessarily' what 'I know what's going to happen in a piano' means. 2 Standard edition errors (content omitted): But I don't know what the next word is going to be from someone else + so I've always got that difficulty.
2	I know she wasn't deaf... Um... I know she wasn't deaf, but she had a disability.	I know she was not deaf. I know she was not deaf but she had a disability.	2 Correct editions: Um..., repetition of 'I know she wasn't deaf'
3	She couldn't speak. Can't really work out the reason.	I can work out the reason.	1 Standard edition error: She couldn't speak. 1 Serious recognition error: Can/can't. This is likely to be a recognition error, since the two words are similar. A new, plausible sentence is created.
4	It's, it's a kind of very small growing movement but it's all to do with the inclusion of, er, so media access professionals, so subtitlers and audio describers in the filmmaking process itself.	JOSH: it is a small, growing movement, But it is to do with the inclusion so subtitlers an audio describers In the music process themselves.	5 Correct editions: It's (doubled), a kind of, very, er, so 1 Serious edition error (misinformation): 'music process' 3 Minor edition errors: punctuation (,) (0.25); media access professionals (0.25); try to create a dialogue (0.25) 1 Minor recognition error: it (Logged as a respeaker error, rather than software-specific error, as this was caused by a macro-setting) 5 Software-specific errors: Highlighted in bold (all listed as minor recognition errors)
5	So it's not just, "Well, we've got subtitles there we go, um, we're, we're ticking boxes," but it's actually thinking about subtitling and media access in general as a new, as another cinematic tool in many ways.	It is about thinking about subtitling and media access in general As another cinematic tool.	4 Correct editions: But, actually, as a new, in many ways 1 Standard edition error: omitting 'So it's not just, "Well, we've got subtitles there we go, um, we're, we're ticking boxes,"' 2 Software-specific errors: Highlighted in bold and listed as minor recognition errors

before continuing, which seems to suggest that the subtitled content has been accepted as accurate since no attempt is made to correct it. Whilst the scoring was not adjusted, this illustrates the complexity involved in assessing respoking at live events, since their interactive and participatory nature means that the speaker has the potential to affect error correction (and sometimes creation) in a way that is not possible in the recorded content broadcast on television.

The second serious error is a recognition error (segment 3), where 'can't' is respoken as 'can'. Since the resulting sentence, 'I can work out the reason' looks correct, the audience is misled.

The final serious error occurred when 'film making process' was replaced with 'music process', most likely because the link between both industries were being discussed (segment 4). Whilst this error was not corrected, the audience had the opportunity to gain a fuller understanding of the filmmaking process as the discussion continued, so although misleading, the impact of the error was not as serious as it might otherwise have been.

On the other end of the scale were the minor errors (segments 4 and 5), the most frequent kind in this event. The omission of 'media access professionals' whilst retaining 'subtitlers' and 'audio describers' is an example of a minor edition error; in this case, the audience received the basic information to understand what was being said, but an additional detail was lost. The misrecognition of 'and' as 'an' in segment 4 is an example of a minor recognition error; the audience are usually able to spot and understand the intended meaning.

Falling somewhere between these on the error scale, come the 20 standard errors that occurred in this event, all of which were edition errors, where the oral equivalent of a sentence is omitted. These errors often occur when the respeaker prioritises the central idea; for example, in segment 5, 'thinking about media access in general' is captured, but a comment about 'ticking boxes' by having subtitles is omitted. Sometimes a standard omission error might follow a correction, which seems to be the case in segment 1. In segment 3, a standard edition error, where 'She couldn't speak' precedes the serious edition error (can/can't) discussed above. Although the serious error carries the greater penalty, in this instance, it seems that the standard error has a greater impact on the meaning, since it provides the connection between segments 2 and 3. For this reason, when possible, a respeaker should aim to correct both serious and standard errors.

Although a significant percentage of the original was omitted, the impact on the overall content of the event does not appear to be too severe; the correct editions seen in these segments support this, as does a lexical analysis of the text. The density of the spoken transcript was found to be 25.1%, whilst that of the respoken transcript was higher, calculated at 29.5%. Although fewer words were respoken, and sentences were shorter,

the lexical density actually increased. Editions were made, but, for the most part, the respeakers were able to preserve the content and complexity of the discussion and the response from the audience was positive: those audience members who were d/Deaf, deafened or hard of hearing stated that the subtitles improved their access, one saying they gave “100% understanding of the content”.

Having the respeakers in close proximity to the audience also raised awareness about the importance of access and what respeaking involves. Many commented that although they did not need the subtitles themselves for access, they were happy to have them there for the benefit of those who did. Another noted that the respeakers must make on the spot decisions about what to include and stated, “the Q&A in the discussion were of high quality and increased my perception of respeaking”.

Having positive audience feedback about the respeaking alongside the NER analysis is essential. Unlike television, where most people are able to watch (sections of) programmes more than once, a live event cannot be rewound if something is missed. The audience want to engage fully and the respeaker must be able to navigate what is at hand to provide that access effectively, as they have done here.

This event also suggests that 98% is a good benchmark for acceptable accuracy. The two NER scores fell on either side of this benchmark, and the audience feedback was very positive. At the same time, the range of feedback collected shows that whilst accuracy and latency are important, and poor scores in either are likely to lead to a poor overall rating, they are not the only aspects of the event which affect the event quality for the audience.

8.5.2 Event 7: Museum tour

The seventh research event was a Highlights tour at Manchester Art Gallery. The accuracy at this event was high, with a score of 98.7% (good) including software errors and 99.0% (very good) excluding them. As a Round Two event, where the respeakers had good time to prepare and the opportunity to meet with the presenter, a high score was certainly desired. On the other hand, this was an event where the respeakers had

limited access to the visual content, so in some ways the respeaking scenario was one of the more challenging ones.

Unlike the post-screening discussion at event 6, the core content of the museum tour was planned by the guide in advance and shared with the respeakers. They were able to enter and train specific items of vocabulary relating to the artwork that would be viewed during the tour during their pre-event preparation and these were recognised accurately. The tour was not scripted, and the audience were invited to ask questions, so there was also spontaneous speech to be respoken, but given the focus on particular works of art, the respeakers could prepare for this with a degree of prescience and very few vocabulary items posed problems during the tour.

A striking feature of the data for this event is the speech rate of 112wpm, and subsequent respeaking rate of 104wpm and subtitle rate of 85wpm. This is very low in comparison to most of the other events seen in this study and also to television content; the reason for it is the mobile nature of the tour. Whilst the tour itself lasted 50 minutes, time was needed to move around the gallery to the next item of interest. As with any audiovisual content, the audience must have enough time to process both the image and spoken word as 'the whole' comes from the combined content. When the spoken word is displayed visually in subtitles, more time is needed for this and when the focus of what is being said is the visual content, it is vital to ensure that time is left for the audience to see and explore that content after reading the subtitles, before new spoken content is introduced. While the words spoken may be uttered at regular speech rates, the pauses between blocks of speech reduce the average rate for the event.

It cannot necessarily be said, however, that this made the event 'easier' for the respeakers to provide access for. Given the mobile nature of the tour, the respeakers were not present alongside the guide and therefore had limited visual access to the content they were respeaking. Whilst they had been able to find online images of some pieces in advance of the tour, there were others that they had only read about. Nevertheless, they were able to respeak the full tour relying on audio feed alone.

The fact that the audience had personal screens on which to view the subtitles also added to the complexity of the event. On the one hand, the number of tablets available limits the number of people for whom access can be provided; on the other hand, the tablets allow individuals to personalise the presentation of the subtitles, adjusting the size, colour, font and background to a combination of their choice. Time is needed before the event begins to inform the audience of this possibility and to allow them to explore the range of options available and become familiar with the settings. Logistically, this may mean extending the length of the event. Further, in order to allow time for the audience to read the subtitles and have time to look at artwork, it may be necessary to schedule longer at each piece of work than in a non-accessible tour.

In this tour, the subtitles were streamed to the audience's tablets and this required a reliable internet connection at all locations in the tour and while moving between them. When reviewing the footage, it was noted that drops in internet connection had occurred during the tour which meant that not all the subtitles the respeakers were producing were being received by the audience as they toured the gallery; this impacted on the quality of access the audience received and it led to some variations in the experience of access for different audience members, since it seemed that these drops only affected certain tablets. Although the number of errors at this event remained low, the range of errors seen differed from that at other events, since serious and standard software-specific errors were recorded. Elsewhere, the software-specific errors had been minor, resulting from unusual spacing and the occasional doubled letter; at this event, the software errors seen resembled edition errors, as new sentences were created as content was omitted (table 8.15).

Table 8.15: Example of software-specific errors seen at mobile events

Segment	Spoken	Respoken content (excluding software-specific errors)	Audience tablet feed (including software-specific errors)
1	I showed you the pictures of Joshua Reynolds, just let me show you the sort of paintings that the pre-Raphaelites admired from the past.	I showed you the pictures of Joshua Reynolds, let me show you the sort of paintings the pre-Raphaelites admired from the past.	I showed you the pictures of Joshua Reynolds, admired from the past.

The omission of 'let me show you the sort of paintings the pre-Raphaelites' creates two recognition errors. Firstly, there is a missing space between 'Reynolds,' and 'admired'; more significantly, a sentence with new meaning is created, since now the audience are led to understand they will be looking at pictures Joshua Reynolds admired. In the actual respoken content, there were 27 edition errors, out of which 3 were serious, 11 were standard and 13 were minor, and only 15 recognition errors, all of which were minor. In the content streamed to the tablet recorded during the tour, 26 edition errors were found, of which 3 were serious, 11 were standard and 12 were minor, and 23 recognition errors, including 1 serious error, 6 standard errors and 16 minor errors.

In a mobile tour, it is therefore essential to check the internet connection throughout the building in advance in order to ensure that the audience will be able to receive the full content of the respoken subtitles. In certain buildings, this may either restrict the areas where the tour can move through, or mean that a mobile boost is required to ensure the connection remains stable.

The fact that the respeakers are working remotely also have implications for how the guide (presenter) must conduct the tour. Their words will need to be shared with the respeakers via a mobile phone, either handheld, or through a headset; allowing the presenter to get used to this before the tour begins is of great benefit. At any event, it is helpful to have a nominated person in charge of monitoring the subtitles and communicating any issues with the presenters and the presenters need to understand how they can facilitate the work of the respeaker; in a museum tour, this is all the more important as the guide is unlikely to have their own tablet. Communicating with the audience via tablets is a different experience for any presenter and, in a tour, audience and presenter are in particularly close proximity. Remembering to repeat questions from the audience and to leave time for responses and viewing time also mean the flow of the tour is a different one.

In practice, this guidance is important for presenters at all events made accessible through respeaking, and the discussion of this in the focus group that followed this event

contributed greatly to the development of the presenter training outlined in section 7.7 above.

8.6 NERLE: The NER for Live Events

Throughout my analysis of the respeaking seen at the live events, it was the NER model that I used to calculate their accuracy. However, as I worked, I increasingly found that before I could assign an error, especially in the case of omission, I had to carefully consider why something had been said, and what the communicative function behind the words were. On many occasions, it was not immediately clear in the context of the event, whether omitted words could be counted as correct editions, or whether a penalty needed to be applied. Whilst all the scores that I assigned fitted within the serious-standard-minor framework of the NER, the analytical pathways I took as I made them often varied from the traditional NER pathways illustrated in section 8.3.2 above (Moore and Romero-Fresco, 2019). It was in the course of carrying out this analysis, and thinking through these decisions, that the assessment pathway for the NERLE developed.

8.6.1 The need for the NERLE

One fundamental difference exists between respeaking on television and at live events, a difference which ultimately meant that an adapted version of the NER is required in this new setting. It lies in the interaction that is possible in each setting (Fig. 8.7).

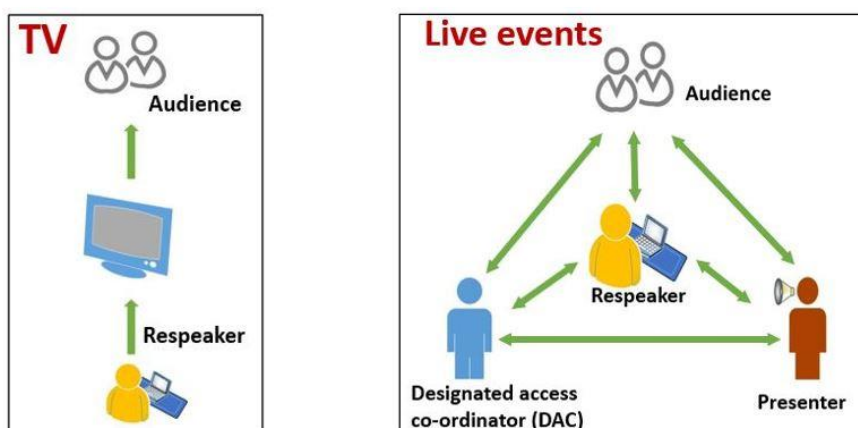


Fig. 8.7: Interaction on TV and live events

When respeaking on **television**, the respeaker is able to research the content beforehand and may even receive scripts which assist this preparation. Once the programme begins, the respeaker watches, listens and respeaks, essentially following and responding to the content that unfolds before them; however, they cannot communicate with anyone in the programme. The only adjustment that might be made to this broadcast content and subtitles is that an antenna delay may be used to improve their synchronicity/limit their latency

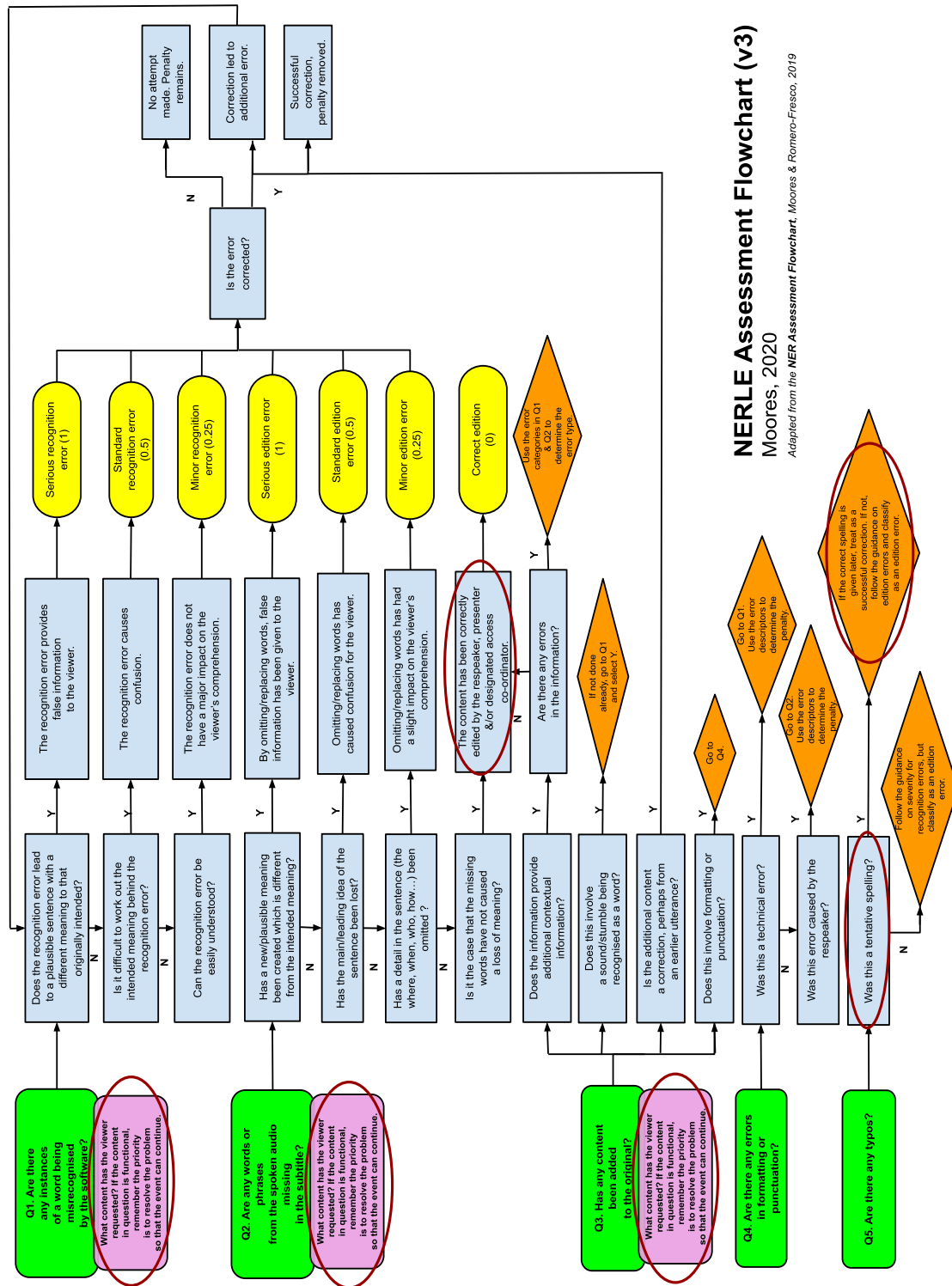
At **live events**, however, the scenario is very different. Everyone attending can interact with each other and with the respeaker, and these interactions may, at times, affect the decisions the respeaker takes and the content of the subtitles. There is often a designated access co-ordinator (DAC) who monitors the subtitles and intervenes if an error appears or some content is missed, and the audience and presenters may also respond to the content they see.

The scoring system for the NER and NERLE remains the same, as the LE stands for Live Events and figure 8.8 below illustrates the analysis pathway of the NERLE. The rings indicate the occasions on which the two pathways differ.

In the sections that follow, I provide an overview of the decision-making process involved in the NERLE. A more detailed explanation of the scenarios that might arise at live events and the associated scoring considerations can be found in appendix 8.4.

8.6.2 Determining what content should be prioritised

I identified three general communication types at live events. The first type is the **formal event content**, the content you would expect at the event. Also included here is any communication related to new or complex vocabulary being used that the respeaker might struggle with, any errors in the subtitles, and occasions where the respeaker is addressed directly, in what I call a 'reversed aside'.



NERLE Assessment Flowchart (v3)
Moore, 2020

Adapted from the NER Assessment Flowchart, Moore & Romero-Fresco, 2019

Fig. 8.8: NERLE assessment flowchart (v3)

Next, is **functional or problem-solving communication**, that is used to resolve any issues that arise. There may be missing content, speaking which takes place off-mike, technical problems with the audio/subtitles or slowed or interrupted delivery, many of which have already been discussed in the analysis of the events above. The final type is **additional information**. Some of this additional information is common to most types of subtitles, for example, speaker ID and sound labels, but there are also additions that are more specific to live events, for example, the inclusion of information about the event, positional/directional information about the action taking place, and a subtitle introduction. I explore this new content, and any implications when assessing its accuracy in more detail, as this also represents a significant change from the work of professional respeakers on television, and a potential new area to explore at live events. When using the NERLE to assess quality, as the person scoring reviews each segment of the subtitles, there is a reminder for them in questions 1-3 to consider what needs to be respoken. The first part of this reminder reads, **“What content has the viewer requested?”** At larger events, such as the events in this study, where subtitles are displayed to the whole audience, the respeaker will most likely be expected to respeak all the formal event content, in a similar way as they are asked to do on television. However, there are scenarios where a single person may be viewing the subtitles, and they may have a more specific request and only require some of the content. In this case, omitting content that is not required would be a correct edition, rather than a standard omission, and it would be vital that as well as the respeaker, the person marking is also aware of this request so that they can assess the accuracy correctly themselves. Examples of these scenarios might include a 1-1 appointment, or small business meeting (Eichmeyer, 2017; Eichmeyer and Kurch, 2017), and a possible request might be that only the words of the other(s) attending the meeting, and not of the person for whom the access is booked, be respoken¹⁰².

¹⁰² On television, such a distinction surrounding what content might be edited out is rare. Sports is an example where respeakers may be directed to edit in-play action more heavily to allow commentary information to be prioritised (BBC, 2021); where multiple speakers talk, it is

Even for larger audiences, at events there may be other content that can be omitted. The reminder to the person marking continues: **“If the content in question is functional, remember the priority is to resolve the problem so that the event can continue.”**

Here, the person scoring must consider why what was said was said. As we will see, in some situations, the words spoken may seem to belong in one communication type, but are actually part of another.

This reminder is designed to encourage such reflection. And this explicit focus on the communicative function of what is said brings the NERLE a step closer to models such as the IRA (Eugeni, 2017) and WIRA (Eichmeyer-Hell, 2020, 2021). These models approach quality from the perspective of how ideas are rendered and the extent to which communication is established.

8.6.3 Formal event content

From a communicative point of view, the formal event content is most similar to the content that is usually assessed by the NER, and it is the planned content of the event. However, the dynamism of live events means that this content often evolves on the day in response to who is present and what is said and the respeaker may often become part of this process.

8.6.3.1 *Reversed asides*

Exchanges between the people at live events are often spontaneous, and everyone who takes part brings a new dynamic to the event, potentially taking what is said in a new direction. Just as those participating will respond to the other people present, they may also respond to the live subtitles. During the research events, the subtitles were a prominent feature of the event and on many occasions, once they began to feel comfortable being respoken, many presenters began to consider how the respeaker

understood that some prioritisation will be needed. However, the general guidance is to respeak close-to verbatim, capturing most of the content in the subtitles.

might be coping with following their words, either commenting on whether what they were saying would be difficult to respeak¹⁰³, or trying to test the skill of the respeaker and capability of the speech recognition software.

There were also many instances of communication between the respeaker, presenters, designated access co-ordinator (DAC) and audience members. Very often this communication related to ensuring the content could be captured in the subtitles or technical issues, and could be considered a kind of aside, akin to something said in parenthesis, rather than part of the formal content of the event; such parentheses involved functional and problem-solving communication, considered in 8.6.4 below.

However, just as the subtitles became part of the event, many of these communicative exchanges evolved from straightforward asides into an active element in the unfolding of the event. For example, in event 3 at the BFI, when asked to repeat a detail the respeaker had not been able to capture, rather than complying, the presenter responded, “I have just been censored!”, incorporating what had just happened as a real-time example of the censorship her film referred to. The words spoken in moments such as these represent formal event content, albeit spontaneous, and any errors or omissions which occur should be marked accordingly.

Although there was a particular emphasis on the presence of subtitles at the research events, this tendency to involve the access provider is more widespread. On the Stagertext website, there used to be just such an example, from the Festival of the Spoken Nerd (2014), a live, semi-improvised comedy show with live subtitles produced through STTR. As the video text reveals, the crew “wanted to make STAGETEXT part of the show, not just a token thing added on” (1.58). The stenographer is visible to the audience as she works and at one point in the show, her hands and keyboard are even displayed on screen (2.09). The first laughs happen when one of the comedians points out the subtitles to the audience. He stops talking while watching the words appear on screen and then waits for the subtitles to continue, forgetting he needs to feed the words

¹⁰³ An example of this happened at event 7 and can be found in appendix 8.2, Event 7_Assessment, row 73.

(0.17). Here, everyone's attention is focused on the subtitles, so the break in content is enough to alert the presenter to what is happening. No functional cue is needed. He comments, to laughter, "I think this is going to occur several times throughout the show." Later, the comedians decide to test the skill of the stenographer, musing to themselves, "What would happen if we all started talking at the same time?" (2.17), which they proceed to do. The stenographer responds with [gibberish] to much laughter and applause.

8.6.3.2 Vocabulary

Whilst some presenters may offer the respeaker a spelling for a complex word during the event, the respeaker may also need to seek the presenter's help. They might offer a tentative spelling and signal it with a question mark in brackets (?), hoping that the presenter will offer the correct spelling. Such exchanges over spellings and terminology also fall under the formal event content. Since this is not something which would occur when respeaking on television, it marks another of the differences between the NER and NERLE, and an additional flowchart step, 'Was this a tentative spelling?' has been added to question 5 in the NERLE.

8.6.4 Functional and problem-solving communication

It is functional and problem-solving communication that can be the most complex to score. Take the scenario where, unrelated to any technical issue, the respeaker omits something that was said. When working on television, the moment will have passed and the respeaker will (hopefully) move on and focus on what is said next. At a live event, different responses might follow.

The respeaker may cue (PLEASE REPEAT THAT), the DAC may ask the presenter to repeat the missing content or the presenter might notice on their own and repeat the content. However, the presenter may also decide not to and continue presenting new information.

All these possibilities affect the content of the subtitles for the audience, as they will only see these exchanges if the respeaker is able to respeak them. The priority for the respeaker should always be to continue respeaking the formal event content, or the content that has been requested. When scoring, the first consideration should be whether the main content has been captured. Only then, would the scorer consider whether these functional exchanges needed to be subtitled.

Very often, omitting these functional exchanges would be considered correct editions, but occasionally, the missing content itself becomes a joke or talking point, and implicitly becomes part of the formal event content, as just described in 8.6.3, and if this happens, omitting it would lead to an edition error being applied. It might be a standard error, but this would depend on the context of that event. Since scoring happens after the event, capturing these subtleties can be difficult.

8.6.5 Additional information

Additional information refers to content that the respeakers capture in their subtitles that go beyond the words that were spoken. Sound labels and indications of manner, that are seen in many forms of accessible subtitles, are examples of such additional content. However, some additions are specific to the live event setting.

8.6.5.1 Speaker identification

Whilst on television, a change in colour is most frequently used to identify a change in speaker, at live events, speaker identification is marked with name tags or speaker labels. There are a number of possible errors that might occur in relation to speaker identification, and at live events such errors are likely to have a greater impact, since the people being labelled are present in person. A person's name may be spelled incorrectly, a generic label such as NEW SPEAKER may be used for one of the key presenters, or the wrong ID label may even be assigned.

In order to limit the possibility of such errors occurring, the respeakers should connect with the presenters and organisers in advance to ensure all speaker identification labels

are listed correctly. If (typing) errors do appear in them, this would certainly be something to feedback to the respeaker in the post-event evaluation. The severity of the error assigned through NERLE marking would depend on the extent to which the label impacted upon the audiences' comprehension of who was speaking. Questions to consider would be how long was someone speaking with the incorrect label, and whether being in the room meant that the audience realised the error. It is possible that at live events, errors can be spotted more easily than when they appear on television, which may limit the impact of such an error.

8.6.5.2 How people speak

During the initial focus groups and post-event chats, the question of how to indicate the accent of the person speaking, and what they sound like, arose on many occasions. What was clear from the initial focus groups was that participants wanted access to the same information that hearing viewers would have. During the post-event chat at the BFI (event 3), the question of accent came up again and dominated much of the discussion. Different opinions were expressed, but for some participants and at least one of the presenters, the way their voices were represented was a matter of personal importance (section 7.6 above).

However, in contrast to this, the general guidance on capturing accents in subtitles is to give it a lighter touch, only indicating accents in pre-recorded subtitles when relevant for the viewer's understanding. The BBC guidelines (BBC, 2021b: 12.1) state:

Do not indicate accent as a matter of course, but only where it is relevant for the viewer's understanding. This is rarely the case in serious/straight news reports, but may well be relevant in lighter factual items. For example, you would only indicate the nationality of a foreign scientist being interviewed on Horizon or the Ten O'Clock News if it were relevant to the subject matter and the viewer could not pick the information up from any other source, e.g. from their actual words or any accompanying graphics. However, in a drama or comedy where a character's accent is crucial to the plot or enjoyment, the subtitles must establish the accent when we first see the character and continue to reflect it from then on.

Where accent or dialect does need to be indicated, the suggestion is that it is done “sparingly”, providing “a flavour” of it for the viewer. The reason for this is to avoid slowing the reading process and so that the speaker is not “ridiculed” by having their words transcribed phonetically (12.2). The guidelines suggest spelling a few words phonetically, including any unusual vocabulary items: “For a Cockney speaker, for instance, it would be appropriate to include quite a few “caffs”, “missus” and “ain’ts”, but not to replace every single dropped “h” and “g” with an apostrophe” (*ibid.*).

In live subtitling, the real-time nature of the production and use of speech recognition software make indicating accent even more complicated. In particular, using non-standard spellings may be more complicated via speech recognition software. Information about a presenter’s accent would certainly be useful for the respeaker to have beforehand, but even then, it may be challenging to subtitle using non-standard spelling throughout. How can this position be aligned with the desire for “equal access” expressed in the focus groups?

My suggestion is to include a *subtitle introduction* before the event proper begins as a possible solution. This idea is based on the audio introduction which has become an established feature of audio-described performances for live events such as theatre, opera and ballet (Fryer, 2016: 155). This audio introduction provides a framework for understanding the play and is an opportunity to “describe the visual aspects of a production in more detail than can be fitted within the soundtrack” (p.156). It might include the AD of a complex set and more detailed descriptions of characters, locations and even action sequences. In a similar fashion, a subtitle introduction, could be shared with the audience before the event begins, through either respoken or cued titles, so that those attending have additional auditory information about the speakers.

When moving from spoken to written language, many auditory cues are lost. Even straightforward changes in respoken text, such as moving from contracted spoken forms (isn’t, hasn’t) to uncontracted written forms (is not, has not), adds an edge or formality within what is written that was not initially voiced. There is a subtle shift in implicature. Only some of these can be captured by sound labels and formatting. Including some

reference to how the presenter speaks and comes across in a subtitle introduction might be a way of reintroducing some of these cues and could potentially allow presenters to be represented more fully. It could be a very short description, or, if everyone is willing, it could become more ambitious.

My initial thoughts are that the description could capture how a person's personality is captured in their voice, perhaps by describing their delivery, and refer to their accent. Attempting to describe my own voice, a subtitle introduction might read along the following lines:

Zoe is friendly and cheerful and has a pleasant voice, which is quite low-pitched. She says she will try to speak slowly and clearly to help the respeakers, but when she feels nervous or gets animated by what she is saying, she tends to speak quickly! Zoe comes from London and has a standard British accent.

When required, such an introduction is also an opportunity to introduce a person's pronouns and, if this access was also required, it could be combined with an audio description of the presenters' appearances. It is well-established that a person's voice is a key feature of their identity (Meizel, 2011; Filmer, 2019) and scripting a subtitle-introduction could be a sensitive task. Accents play a critical role in how cultural identity is formed and categorised and is often taken as "a meaningful indicator of an ethnic category" (Dehghani *et al.*, 2015: 231), even more meaningful than a person's visual appearance (Rakić, Steffens and Mummendey, 2011). A person's voice can also reveal how comfortable they feel in a situation and the emotions they are feeling (Laukka *et al.*, 2008). A voice may also reveal deafness or disability. By including any of this information in a subtitle introduction, potentially in both spoken and written form, it becomes more marked, so the choice of phrasing will be key. Audio description and audio-introductions respectively provide some useful guidance here. On the one hand, is this kind of description really needed? In response to this, it must be remembered that the information being captured is information that some audience members are absorbing implicitly (Fryer, 2016: 152). When this information is not provided, the automatic response may be to align a person's visual or auditory appearance to one's own. For

example, when reading someone else's words without any accent indicators, a person might automatically assign their own accent to the words before them, thus erasing a feature of the speaker's identity. In the case of audio description, it seems that people do want to have more detailed visual information. Blind and partially sighted participants in a number of studies (Gerber, 2007: 31-32; Fryer, 2016: 150) stated that details such as race and disability were important elements of cultural literacy. One particular playwright, Maria Oshodi, said that if these features were not mentioned in an audio description, they spent the performance trying to guess them, which detracted from the actual performance (Fryer, 2016: 151). This seems to mirror some of the comments expressed in my own focus groups. Yet, writing such visual descriptions raises similar issues of sensitivity:

The actors they describe may listen to the AI, or at least may be told by an AD user how they have been described in the notes. It is possible for an actor to take offence at a description. That said, remember that it is the character you are describing rather than the actor (p.158).

As we saw with speaker identification, at live events it *is* the person and not a character whose voice is being described. This is the reason why if such a practice were to be adopted, the subtitle introduction would be written by or together with each of the people being described, in advance of the event, so that appropriate language could be found that everyone was comfortable with. This would mirror the approach seen in accessible filmmaking (Fryer, 2016: 152; Romero-Fresco, 2019). An important first step would be to test this in a reception study and get audience feedback on whether they would like to see this feature introduced and the kind of content, and language they would want to be included.

8.6.5.3 Event announcements

Finally, at the start of most events, announcements are made with information about what to do in case of emergencies, for example indicating where the fire exits are. As well as being essential information that must be made accessible for everyone attending,

and therefore included in the subtitles, this also means that an opportunity already exists to provide information that is supplementary to the formal content of the event, where the subtitle introduction could be included. Additional information which could also be shared at this point includes reminding the audience why there are subtitles (I say reminding, because the audience should have been alerted to this at the point of booking), perhaps alerting them to the fact that they may notice the respeakers if they are in the same room (as suggested in Chapter Seven) and giving the audience a little guidance on how to interact with the microphones, or why they themselves might find the subtitles useful.

At respoken events, it is likely that the microphones will be transmitting the speaker's words directly to the respeakers, so it is very important that the speaker waits for the microphone to arrive and positions it close to their mouths. In times of Covid, there may be additional health and safety considerations implicit in this practice that need to be explained. This does not simply have to be instructional; it could be done in a fun way, and whoever is making the announcement could point out that if the speaker stops talking... the subtitles do, too, as happened in the comedy show referred to above. This will also serve to alert the audience about the latency in the subtitles. The subtitle introduction would be well-placed to follow this information, before the formal content of the event begins. By making these direct references to the subtitles, it means that by the time the 'main event' begins, every member of the audience understands how the subtitles work and have already become at least a little accustomed to using them. The position of subtitles at live events means that audience members are more likely to consciously choose to look at them, than is the case when watching subtitles on television where eyes may be automatically drawn to them. This makes it all the more important to accustom audience members who are less familiar with using subtitles to them, since, as I have shown, they can be of use to and used by everyone present. And the more on board everyone is in their use, the more likely it is that (the technical side of) the access will run smoothly.

8.6.5.4 *Visual information*

There is one final place where the respeakers may add information to the formal content of the event. Where an event has important visual information, that could be missed if an audience member's eyes are locked on the subtitles (see section 3.6.3, table 3.2 above), the respeaker might choose to add positional or directional information to alert the audience member to look up (LOOK UP!). To an extent, the usefulness of such a technique will depend on the latency seen at the event. This is likely to vary according to the event type and also according to the individual presenter. Some presenters may be skilled at managing – and minimising – the latency themselves in their style of delivery. Others may find this far more challenging.

The guidelines to presenters, and future training, advise presenters about this and encourage them to think about this especially when using densely packed visuals such as slides, or in a visual tour at a museum, for example, where there is much to see. However, some presenters communicate visually, using gestures to animate their words and it is in these instances that such deictic labels could be useful.

Just as television presenters are asked to move subtitles so that they do not cover key content, here respeakers would be attempting to redirect the eyes of the audience so key content can be noticed. As with the subtitle introduction, this idea developed in response to the final research event and should be fully explored in a reception study.

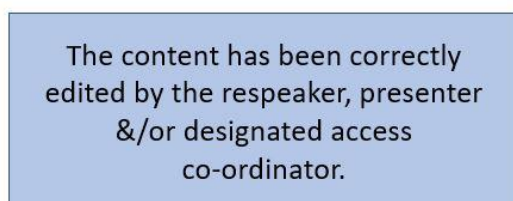
8.6.6 Shared responsibility for errors

The interaction between everyone present at the event means that the respeaker, and the access they provide, are integrated into the event as a whole. It also marks a final difference between the NER and NERLE Assessment Flowcharts, since at live events, the responsibility for errors shifts.

On television, errors might be caused through software misrecognitions, or by edition errors made by the respeaker. In both cases, it is the respeaker alone who can correct them. At live events, along with the respeaker, the presenter, DAC and audience

members may all contribute to errors and help to correct them. They may repeat information that was missed, correct errors in the subtitles, or supply a correct spelling, with or without being asked. However, they may also cause content to be missed. Presenters may forget to repeat a question or comment from the audience, or choose not to repeat missing information even after being asked. Although not caused by the respeaker, penalties would still be applied for these errors if they impact on the viewer's comprehension.

This final error is captured in the explanation of what a correct edition is, as illustrated in figure 8.9:



The content has been correctly edited by the respeaker, presenter &/or designated access co-ordinator.

Fig. 8.9: Explanation of correct editions in the NERLE

8.7 How do the NER analysis and audience feedback compare?

Having considered the feedback from those who attended the events (Chapter Seven) and the results of the NER analysis in the current chapter, I want to briefly consider any general patterns that can be found within them. In doing this, it must be remembered that these are two very different sets of data; the audience feedback is more personal, based as it is on the questionnaires that most people attending the events completed, and it responds to the event as a whole; in contrast, the technical data is based on a sample, albeit representative, of the event, and should be less subjective. Nevertheless, some interesting insights may be revealed through this comparison. Table 8.16 contains the combined DH scores and the total audience scores from table 7.2 in Chapter Seven and various strands of data from the NER analysis; I have applied a generalised coding system to them, to illustrate low, medium and high scores (table 8.17)¹⁰⁴.

¹⁰⁴ Although lower than the average latency of 7-8 seconds seen for pure respeaking in the Ofcom study (2015c: 15), I decided to take 5.8 as the boundary between a medium and high score, since 5.8 was the average latency seen across the events in this study (see section 8.4.3 above).

Table 8.16: Total audience and combined DH scores alongside data from the NER analysis

	1	2	3	4	5	6	7	8
Audience total	Low	High	Mid	Mid	Low	High	High	Mid
DH combined total	Low	High	Mid	Mid	High	High	High	High
RLE accuracy	Acceptable	Substandard	Acceptable	Substandard		Substandard	Good	Good
RLE* Accuracy	Very good	Good	Good	Acceptable		Acceptable	Very good	Very good
Average latency	6.7	7.5	5.9	4.5		5.8	4.3	5.7
Speech rate in clip WPM	143	135	136	123		145	112	133
Subtitle speed WPM	104	119	102	84		99	85	116
Reduction rate in subtitles	27%	12%	25%	32%		32%	24%	13%

Table 8.17: The value of low, medium and high audience, NER and latency scores

	Audience scores	NER scores	Latency
Low	Below 75%	Substandard	Above 5.9
Medium	75% - 79.9%	Acceptable	5 - 5.8
High	Above 80%	Good/very good	Below 5

I have included the subtitle speed or reduction rate for reference, but since such rates relate primarily to the event content and may vary according to genre as well as the respeaker's decisions, I have not attempted to code them.

Event 7 received high scores throughout and it seems to represent the model of a good event. It is interesting to note that this is true despite a fairly high reduction rate in the subtitles; similarly, although there were some technical issues (8.5.2 above), they occurred in the transitions in the tour, so do not appear to have impacted too negatively on the tour as a whole.

Events 6 and 2 received similarly high scores from the audience, yet the original NER score, when software-specific errors were included, was substandard, rising to

acceptable and good when they were excluded. In addition, event 2 had fairly high latency in places. These events highlight the fact that other aspects may guide the audience's perception of an event. Perhaps the low reduction rate was obvious to the audience, and this justified the slightly higher latency to an extent; and perhaps the general atmosphere and interest in the event, along with the subtitles being placed in a suitable position in relation to the set-up of the event as a whole, meant that the audience could focus on the main content of the subtitles and ignore any software-specific errors that did appear. Certainly, the contrast between events 1 and 2 suggest that these other factors were important. At event 1, the overall accuracy of the subtitles was higher and their latency was lower, yet the audience scores were low. Recalling the feedback in section 7.4.1 above, many audience members talked of the fast pace of the event, yet the subtitle speed was fairly low, and the latency was lower than at event 2. It is possible that the audience were aware that not all the content was being captured in the subtitles. This certainly suggests that audience perception and engagement also plays an important role in reception, and that these deserve further exploration.

Finally, the scores in event 3 seem to align in a different way; the audience scores and NER scores were mid-range overall, with the accuracy rising to good when software-specific errors were excluded. This seems to reflect the effort experienced in participating in the event, with the technical issues that were experienced.

At event 8, what is noticeable is how the high DH audience score corresponds to the accuracy of the live subtitles and a very low reduction rate, whereas the total audience score corresponds to the average latency. A number of NNE audience members attended this event and commented on the latency, and it is possible this influenced their scoring.

Certainly this very general comparison of the two sets of data raises a number of interesting questions for further consideration, and also highlights the importance of the qualitative comments that accompany any NER and NERLE score, since it is these that will address these wider considerations that the audience at a live event will be just as aware of as the subtitles.

8.8 Reflections

In this chapter, the accuracy of the live subtitles created at the research events has been shown to meet, and often exceed, the industry-standard benchmark of 98%. Given the limited opportunity the respeakers had to develop their voice models and familiarise themselves with this new environment, it was an impressive achievement. Whilst some of the event content was slower than that found on television, with subtitle speeds ranging from 85-119wpm, other aspects of the live event setting, such as the interaction with presenters, lack of a wider support team and variable physical environment, means that this environment remains a challenging one.

The latency seen at the events follows a similar range to that seen on television and whilst antenna delay is not an option at live events, some latency may be mitigated by skilful interaction with the presenters and by the careful timing of handovers, especially during museum tours, such as events 4 and 7.

Alongside this analysis, the NERLE model was proposed, with its adapted pathways which incorporate the interactivity of live events into the marking system of the NER. As well as being a tool for monitoring accuracy to ensure audience expectations are met, I hope that this model could also allow respeakers to improve their own practice, both by reflecting on the accuracy of their work, and by gaining a deeper sense of the communicative function that lies behind the content they are asked to respeak. This attention whilst editing would be of benefit to their television work, and also to any opportunities they have for more personalised respeaking jobs, as this service grows. The live event setting will be new for most respeakers, so the NERLE could also be used to help identify what jobs or genres individuals perform best at and where the limits of their capabilities lie when considering the duration of respeaking blocks (Szczygielska *et al.*, 2020: 12), something which, as in interpreting, can be difficult to determine on one's own (Moser-Mercer *et al.*, 1998). Knowing this may serve respeakers well in becoming more confident in their work and strengthen the mutual respect between respeaker and audience member proposed in Chapter Two as an ethos behind their approach.

Chapter 9: Conclusions, discussions and further work

"Don't accept the world as it is,
dream of what the world could be -
and then help make it happen."

Peter Tatchell (no date)

9.1 Where are we now?

In this final chapter, I reflect back on what has been learned and achieved during this study and consider the impact that this research has had, both as it evolved and with reference to the ways in which the findings have already been and continue to be used. Some of the discussion is specific to respeaking, however, in the course of the chapter, I also illustrate how these findings can be applied to access more broadly and highlight the fact that any broad discussion of access must eventually include the specifics of the provision, and vice versa, if a continuous chain of access (Greco *et al.*, 2012; Chapter Two above) is to be upheld - something which, I maintain, must be achieved if access provision is to be effective for those who use it.

Given how variable access needs are, because people are themselves inherently diverse, this provision is fluid, with what is required changing according to who is using it, where and how. The live event setting highlights this particularly well as there are so many different factors to consider as an event is organised and held, but it is true for access in any context. There is a risk that considered like this, access may feel like an added burden that must be dealt with, or, when only realising its need late in the planning stages, become an afterthought. With this in mind, I propose a Model of Participatory Engagement, which shows how access can be embedded into the planning in a natural and holistic way, and illustrates that access itself is a process and a way of thinking. I end the thesis with some broader reflections on this, and on the personal journey I have followed in the course of this research.

9.2 The impact of this study

Three key questions were posed at the start of this study that guided the exploration into the potential of respeaking as a solution for providing access at live events:

1. What training and technical set-up is needed to allow experienced television respeakers to transfer their skills to the live event setting?
2. How can high quality respeaking and access be ensured?
3. Can the findings from this UK study be applied to the provision of access across borders and in society more broadly?

Over the previous chapters, I have sought to address all three questions. Whilst the most detailed answers have, perhaps, been provided for questions one and two, through the guidelines and resources that were created, and the practical accounts and analyses of the training and events, these answers also contribute necessarily to the third research question, and how these findings can be applied more broadly, in the UK and further afield. In this chapter, both in the current section and sections that follow, I outline how these findings can be applied more fully.

During the course of this research, I talked to different focus groups, including audience groups and service providers, and established their expectations for a respeaking service at live events. I purposely took the widest approach I was able to when considering who a future audience at live events might be. I worked closely with DH audience members, and also invited NNE to participate, wanting to understand different perspectives on what access might be and to ensure that these could be incorporated into the outcomes of the research that followed.

I used this knowledge to develop, test and refine a respeaker training programme to prepare professional television respeakers to provide this service at live events in a way which would best meet these expectations. Throughout, my attention has been both on discovering how to set up an optimal respeaking service and exploring how this could be embedded within an accessible approach to live events as a whole, so that the chain of access runs through every step of the event, for a smooth(er) audience experience.

Eight research events were used to test these findings. When the NER industry-standard quality analysis was applied, the accuracy and latency seen at the events were shown to be comparable to those seen on television in the Ofcom study. Whilst this does not mean that the respeaking could not be improved even further, it is also impressive that this level was achieved after the respeakers had worked on only two to four events each, and that in a very different setting and under different conditions than they are used to. The findings therefore suggest that respeaking is a serious contender for providing access in this new setting. Most importantly, the audience who attended the events seemed to share this view. Individual opinions did vary, and some criticisms of respeaking remain, but for the most part people were impressed at the access that was provided through respeaking. What was interesting to note was the extent to which the audience scoring aligned with the industry-based NER analysis. The NER provided a good baseline for determining quality, but other factors mattered too: the continuity of the subtitles, the ease with which they could be followed, how integrated they were into the event. All these considerations are just as valid when it comes to considering the quality of live subtitles on television and would be worth sharing with the main providers. The NERLE model for the analysis of live subtitles at events tried to capture the new dimensions that the live event setting brings, with its interactivity, communication and mutual responsibility as live subtitles are created. Appendix 9.1 presents a comprehensive series of questions and considerations for those involved in making live events accessible through respeaking, which draw on the knowledge gained in the course of this study and presented in Chapters Four to Eight.

In running these focus groups, training programme and research events, many opportunities for discussion and collaboration have been created. One of my personal aims during this research was to raise audience awareness of what respeaking was, as I knew it was a widely misunderstood profession and service. This was certainly achieved. Many audience members had the opportunity to see respeaking in action in a very personal way and also to talk to the respeakers, learn more about their work on television and at each event, and understand for themselves the challenges involved in

this profession. I think that this level of collaboration and contact would be good to take forward into future studies.

Behind all of this sits the action research methodology that was so fundamental to the theories underpinning the research and the findings that were achieved. It was this framework that provided a space for the processes we used and the resources that were created to be developed, refined and fine-tuned. I see this methodology as an underused resource within translation studies and one that could be used more widely, and I believe it is well-suited to the new area of accessibility studies. When it is accessibility that is at the heart of the discussion, I believe that collaboration is necessarily part of the solution. With this in mind, there is one further finding and proposal from this study that I would like to share. It addresses the third question of the broader applicability of this research more directly.

9.3 Model for participatory engagement¹⁰⁵

In section 8.6.1 above, I highlighted how the possibility for interaction with a respeaker is different at a live event compared to the scenario when making television content accessible; it is this interactivity that drives the need for the NERLE, and which lies behind the model for participatory engagement. This model is designed at once as a visualisation of how access occurs and as a framework or scaffold for discussions about access that need to take place. It may be particularly beneficial for venues who are exploring access for the first time, or when planning begins on a new (type of) event, to spark different ideas. Venues with established access provision may also find it helpful, if it allows them to approach their current provision in a fresh way, reviewing what is in place and identifying how it might be refined or expanded; doing this ensures that access remains an active and adaptable process.

¹⁰⁵ The creation and naming of this model evolved at the same time as Di Giovanni's article on participatory accessibility (2018) and the label 'participatory' also appeared in earlier discussions of access and disability, for example, Kleege, 2016; 2017; Kleege & Wallin, 2015.

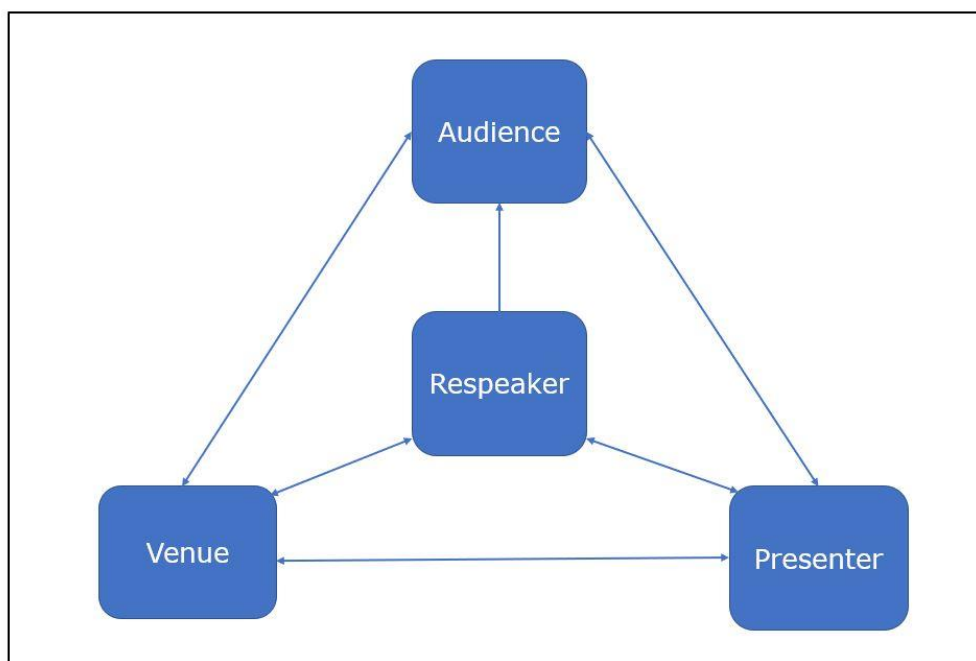


Fig. 9.1: Model for participatory engagement

The model, with its bidirectional arrows, aims to highlight the contact and communication that is required by all involved (Fig. 9.1); each group has a responsibility and agency, be it large or small. Whilst each might seem to function independently, understanding the role of the other parties, and communicating with them is essential and continuous. It is needed as the different stages of planning, holding, and evaluating an event take place. This means that traditional roles grow and expand as access is incorporated.

As we saw in the observations with Stagertext in Chapter Five, and in my own capacity during the research events, there is often a need for a person to step in to monitor the access. The ACT project referred to an access manager (Remael *et al.*, 2019); I refer to the role of designated access co-ordinator (DAC). My reason for doing this is not only to highlight the need for this role to be clear in advance, but also because who this person is may change from event to event, and because, as I conceive it, this role may be held by more than one person. For example, even when there is a main DAC at an event made accessible through respeaking, when it comes to adding or correcting the content of the subtitles, it may be the presenter or audience members who supply that information and of course, the respeaker must respeak it. On a more practical level, multiple members of staff may offer technical support for the audience's tablets or

provide stools for anyone who requires a seat. This reinforces the idea that while one or two people may provide the main access, everyone is involved. In the case of the audience, this may range from stating individual access needs, to respecting the need for this access on the day of the event. These roles are outlined more fully in appendix 9.2.

The respeakers are positioned centrally in this model, as much of the interaction goes through them, and because respeaking must be understood by all for it to be fully embedded into the event. However, as figure 9.2 illustrates, I believe this model can be applied more broadly. Whether a BSL interpreter, audio describer or STTR were to provide access, this same need for awareness and support would exist if it is to be embedded smoothly into the event.

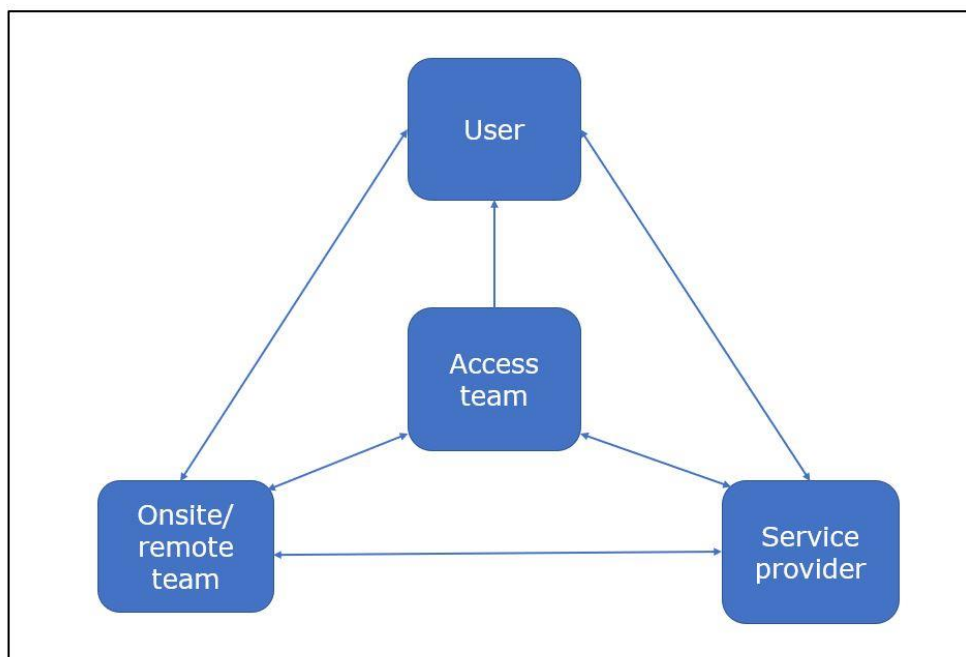


Fig. 9.2: Broader application of the model for participatory engagement

This is true regardless of the type of access, and consideration should be given to all forms of access – be it sensorial, linguistic or cognitive access that might be incorporated into an event, or how the event space is used and managed. Relaxed performances may

be organised¹⁰⁶, flexibility over ticketing may be introduced, not only in terms of cost, which itself is an important issue, but also in how tickets are booked and used; flexibility-ticketed performances, or a portion of flexibly-ticketed seats, may enable someone who cannot attend a tour one day to transfer their attendance to another. For many, the need for socially distanced performances may continue for a while to come and, as awareness of the need for environmental sustainability grows, action towards this could also be built into this accessible approach¹⁰⁷.

In this way, an unbroken chain of access (Greco *et al.*, 2012) that runs through the provision and co-ordination of access is embedded within this idea. I see the potential of this model functioning as a tool or toolkit for encouraging discussions about access, and how it can be provided in different settings and contexts. Rather than providing a set of solutions or an ABC checklist, its purpose is to stimulate thinking, encourage a review of what provision is currently in place, and see what opportunities changes may bring. Some questions will be specific to the access service(s) in use, whilst others will necessarily and very importantly be more general. The questions I have included are designed as a starting point, for finding even better questions to ask as this provision evolves, and awareness grows of different needs that may exist. Returning to the theoretical perspective from which this study began, this model also allows the space for intersectionality, social and epistemic justice and the social model of accessibility to be incorporated.

With its flexible approach, intersectional needs can be considered and accommodated, incorporating the principles of universality and personalisation of poetic design, inherent in the social model of accessibility (Greco, 2019a; 2019b). The participatory nature of this model corresponds to the fifth principle of participation, since there is encouragement

¹⁰⁶ For example, Depot in Lewes say their relaxed screenings offer a more relaxed experience for people who find ordinary cinema visits difficult; the sound is lower, the lighting is not as dark and people are free to come and go during the screening (Depot, 2017). Further discussion can be found in (Fryer and Cavallo, 2021: 24).

¹⁰⁷ During the pandemic, Attitude Is Everything (in CMU, 2021) and #We Shall Not Be Removed *et al.* (2021) provided further guidance on how to reopen venues in an inclusive way and work safely through Covid-19.

for all stakeholders to be involved in the process. Similarly, the principles of universality and personalisation are addressed, as access is flexible and designed to respond to those who attend, rather than being prescribed in advance. By drawing the provision of access forward proactively (Greco, 2018; 2019a), more events will be made accessible, in line with the notion of distributive justice, and there is likely to be more understanding and awareness towards those who may require them, thus embedding hermeneutical justice into the process. In order to fully address testimonial (in)justice, and the principle of epistemic inclusivity of the poietic design inherent in the social model of accessibility, increased involvement of a range of people will be required, people with different needs and different experiences, in line with the 'Nothing about us without us' approach. For some venues, this may begin in-house; for others, it may necessitate wider conversations and partnerships so that those with experience can be involved in the discussion and, as I discuss further in section 9.5, awareness and more representation at a higher societal and governmental level will most undoubtedly be required, for more widespread changes to be seen.

9.4 A starting point for future research

As I had hoped, by using action research, I was able to find many practical answers to the questions I posed and to refine the resources I created so that they are ready to use. Nevertheless, this research has also led to many new questions, and highlighted new areas that are worthy of further exploration.

9.4.1 Extending the range of the events

The events I focused on in Cycles Two and Three were fairly traditional in their nature, albeit very varied. Approaching event design in this way was a fundamental part of my research as I wanted to ensure that the guidelines that were created would be widely applicable. I do, however, see these eight research events as a starting point for future research.

It would be of great benefit to explore different genres and event types in more detail. University lectures, museum tours, Q&A panels, presentations, endless variety exists, and how respeaking is handled in each scenario will depend greatly on the peculiarities of the setting and on the individual content of any event, as well, of course, as on those who attend.

Today, the use of multimodality is increasing. This could be an interesting direction for the research into respeaking at live events. How could respoken subtitles be displayed in a variety of ways? How do we get more creative in the use of respeaking? To an extent, I have paved a way forward for this in the NERLE model and the Model of Participatory Engagement, so there is a framework for beginning this discussion; the pathways are set, and the artistic ideas and collaboration are needed. This move from traditional forms of a modality towards the more creative, is one that is mirrored in, for example, pre-recorded subtitling, and the live event setting offers many possibilities for it to be explored. And to pick up on Desblache's (2019) metaphor, making more "minor" events accessible (Chapter Two), the work of disabled artists and smaller organisations, would complement the range of access on offer and help to increase the representation that is so greatly needed in society.

9.4.2 Wider audience research

I have worked very closely with the audience during this project and the participants who took part came with a range of lived experiences. However, as I have highlighted, the audience is diverse. Further work with different audience groups would be very beneficial. That might mean people of different ages, based in different locations, it might involve considering the different intersections that exist between groups, working with people with a range of different hearing statuses, including different levels of hearing as I have done, but also people who are deafblind or who have tinnitus or hyperacusis, groups who are less often referred to in media accessibility research. This might include learning disabled people or dyslexic people who were mentioned as potential users in

the course of this research, or doing the more extensive research that is certainly justified with NNE, which I began in this study.

Much more consideration of how BSL users respond to both pre-recorded and live subtitles in all settings is another area for further research. This could be a very good interdisciplinary meeting point between media accessibility, Deaf studies and British Sign Language interpreting.

How the audience engaged with the events is an area I was only able to touch on lightly, but it would merit extended research, something which could be coupled with the exploration of different genres of events, and the way that live subtitles, creative or otherwise, are displayed and integrated in different ways. This could lead neatly into interdisciplinary research, which would be well-suited to accessibility studies.

9.4.3 Respeaker training

The respeaker training programme I outlined in Chapter Six is focused on training professional respeakers, working intralingually and moving from television to in-person events in the UK. There are a number of ways that it could be extended in the future.

Additional training could be delivered to tailor it to more specific event settings, be it genres such as educational lectures or museum tours, or working with different platforms to provide respeaking at fully remote or hybrid, online events. Whilst the current training would offer the basics in these areas, further adaptation would be required and this is an area of growing importance both as a result of the pandemic and as our understanding of accessibility grows.

Similarly, this training would provide a good grounding for professional interlingual respeakers to work at live events. Further refinement of modules would be required to adapt to country- and language-specific scenarios, but many of the broader skills here are likely to be widely applicable.

For respeakers beginning their careers in the live event setting, more adaptation would be required. Training on respeaking skills, such as that outlined in Chapter Three, would be required to complement the content provided here and more training resources would

need to be found to allow the practice that new trainees require. At the same time, since new trainees would be less familiar with the different audience groups that use subtitles and access, more general grounding in the principles of accessibility would be a useful addition to the training, in line with Greco's (2019b) critical model.

The SMART Project¹⁰⁸ provides a useful model for how the respeaker training programme might evolve; In SMART, one of our deliverables will be a training programme where professionals from different backgrounds, with different levels of experience, can join the programme and access the specific content they need to train as interlingual respeakers. Adapting the live event programme to a similar format might be a way to accommodate these different entry points and reflect the range of genres and options that could be included.

9.4.4 Process research

Related to this is the idea of process research in media accessibility (Jankowska, 2021). Here, the mental processes at work as a respeaker respeaks would be considered more closely. I addressed aspects of this in the current study, but was not able to go into as much depth as process research would allow. In the case of respeaking, screen recording with key logging in combination with think aloud protocols might give interesting insights into the processes of editing, monitoring and correction in particular, which might inform approaches to respeaking and future training. Similarly, using EEGs to further explore crisis points in the live event setting would build on the work of Szarkowska et al. (2016, 2017) and be very informative.

9.4.5 Links with other access modalities

An area I am very interested in exploring further, in line with the broader application of access, is how different access modalities may complement each other in their provision. I think respeaking and audio description offer a great potential for future research, and

¹⁰⁸ See footnote 33 for further information about SMART.

see many parallels between them – the use of audio and subtitle introductions suggested above, and the fact that both necessitate the use of the spoken word mean that the technical preparation and location on site, plus the terminological research required for both may overlap. Accessible filmmaking has addressed the need to embed different forms of access into a single recorded project and I think the same need exists at live events, when so many different people might attend. BSL interpreters are often present at events, and BSL is likely to be the provision of choice for many people. I would expect the presenter training to be of use to sign language interpreters working at events, since they may well be respoken during it, but respeakers may also benefit from learning more about sign language interpreting and what they could do to support the sign language interpreters as they work.

In the past there has been a great deal of tension between STTRs and respeakers, for understandable reasons, as the areas they have worked have overlapped. I have tried hard to explain the purpose of my research as being to add rather than replace access, but I am aware that concerns may remain. Seeking out more collaboration in the future may be of benefit to both groups and especially to the audience, as ultimately the aim for all involved is for the best possible access to be provided.

9.4.6 The question of quality

Quality remains an area which we talk about frequently in translation studies and media accessibility and yet it is one that can be understood in so many different ways. All too often, there is a tendency to focus on accuracy and latency, yet other aspects also affect quality, and this seems particularly true in the live event setting. In addition, what are we asking about when we talk of quality? Is it subtitles? Is it access? Is it art or cultural content? All three are likely to lead to very different discussions, both in terms of how to mark quality, and if we actually want to – in the case of art, for example. The multidimensional framework offers interesting possibilities for focused cases, for example subtitling (Greco and Moores, 2021), but how could this be expanded more widely?

The qualitative responses to the online questionnaire and following the events, gave some initial direction to what a sense of quality might mean for the audience, in particular in relation to access, but deserves more detailed exploration, in particular as it was not a question I directly asked, but a sense that was revealed through a collection of responses.

9.4.7 Further exploration of the NERLE and subtitle introductions

Finally, in Chapter Eight, the NERLE was introduced. It outlined a new way of approach to assessing the quality of live events, which aligns the NER process more closely with communicative models such as the IRA (Eugeni, 2017) and WIRA (Eichmeyer-Hell, 2020, 2021) and also introduces new elements which could be added to live events, such as the subtitle introduction.

As a new model, the NERLE would benefit from further testing. Can it be applied in its current form to a wider range of events, or are there new features that need to be accounted for and further adaptations required? Will there be a high interrater agreement? Can it be used for different types of speech-to-text output, for example STTR or even automatic subtitles? And how could it be adapted for interlingual respeaking? In the case of the proposed subtitle introduction, the focus groups suggested there was real need for the content it offers to be included in some form at live events and it would be very interesting to take this back to those who attended some of events to get further feedback from them, and to explore it more widely with future audiences.

9.5 Wider reflections

9.5.1 The spread of respeaking

This is a very timely study as respeaking is entering new territories. Its usefulness is being recognised, and its use is growing. In countries such as Poland where respeaking was already used at events, it is now increasingly used on television; in the UK, its entry into the live event setting was spurred on by the advent of Covid-19, which highlighted the need for a fast move to online provision. Respeakers across Europe have seen an

increase in demand for their work, and requests for access provision through respeaking have come from the UK as well. Benefitting from the increased understanding of live event provision that this study has provided, I have worked on a number of online events myself. This study has been a contributory factor towards this growth, both through the findings I have shared and my involvement in a number of the projects I refer to below.

Even before the pandemic began, advances had been made to facilitate this. For example, the LiRICS (Romero-Fresco *et al.*, 2019) certification programme, referred to above, was created which established respeaking as a profession that was eligible for funding from the Disability Allowance scheme. This means that universities in the UK can now apply for respeaking to be used at lectures. In 2019, I was invited to Stockholm to advise on a project similar to my own, where a service for providing intralingual respeaking at university lectures was being established; for the first time, SLIs and stenographers, alongside conference interpreters, were trained so that their aptitude towards respeaking could be evaluated¹⁰⁹.

Research into interlingual respeaking has also come on with great pace. The first online course is being run through the University of Vigo and the ILSA project¹¹⁰ developed an online, accessibility-based training course. 2020 saw the completion of the first PhD in the subject (Dawson, 2020) and the start of the SMART project¹¹¹ at the University of Surrey, where we are investigating how language professionals can be trained in interlingual respeaking.

Since the demand for interlingual access is so broad and extends across many languages, research is also being conducted into how this access can be achieved by combinations of different modalities, including, for example, a simultaneous interpreter in combination with an intralingual respeaker (Eugeni, 2020a; Romero-Fresco and

¹⁰⁹ *Ökad tillgänglighet för studenter med funktionsnedsättning genom taligenkänning* [Increased accessibility for students with disabilities through speech recognition] is a joint project by Stockholm University and the Royal Technical College, funded by Post-och telestyrelsen.

¹¹⁰ See footnote 26 for further information about the ILSA project.

¹¹¹ See footnote 33 for further information about the SMART project.

Alonso-Bacigalupe, 2022). The possibilities are endless and respeaking is an exciting field to be involved in.

At this point, one more question and modality must also be addressed: the use of automatic translation and the extent to which its capabilities compare and contrast with respeaking, both from a task-based perspective and when considering the quality of access it provides for the audience. I remember a conversation as I was beginning my PhD, where I was told that it was in ASR that the future lay. Is that really the case? Is it the right solution for access? As Wooton (2020) said at the most recent Symposium on Live Subtitling and Accessibility, ASR is not only inevitable, but it is already happening, at least in some television contexts¹¹². For example, over 50% of the media that Red Bee Media receives goes through some sort of ASR process and their live automatic captioning solution, ARC, is currently being used for “regulated live content news” for customers in the US (*ibid.*); automatic captions may also allow increased provision of caption content in unregulated areas, such as digital and streaming platforms, and provide a solution when outages occur. So how does the quality of ASR compare to that of respeaking? At Red Bee Media, NER scores for the news have reached 98.5%, though certain programme features mean that the scores for other genres are more varied; in particular, challenges remain with names, places and domain-specific vocabulary, high volume ambient noise, music, people talking over each other and very strong accents (*ibid.*). Fresno and Romero-Fresco (2021) compared output from ASR and respeaking/stenography in television programmes broadcast in the UK and US in Spanish and English. They also found that news was the most accurate genre for ASR, with an average accuracy score of 96.67% for ASR compared to 98.32% for respeaking/stenography. In talks shows and sports, the ASR accuracy score dropped to 95.41% and 95.24% respectively. It is this lower boundary of quality in ASR that is still prone to fluctuation (Wooton, 2020) and that demands that the discussion around its use continues. What of the live event setting? Would ASR be suitable there? My own

¹¹² Live real-time automatic captioning was showcased at the 2019 NAB Show (Red Bee Media, 2019).

response to this, drawn from all I have seen in this study and from using automatic captions on platforms such as Zoom and Teams, is that using ASR alone, as it is currently, will not be a viable solution. Whilst ASR does offer improved latency, it is likely that this will be accompanied by a drop in quality in comparison to respeaking and it is quality and reliability that the audience are asking for. The features of live events tally far more with those that remain a challenge for ASR than those scenarios where more acceptable results are seen. What the NERLE has shown is the importance of the interpersonal interaction in achieving better quality subtitles and better access overall, something that would be lost in an automated solution. As MyClearText (2021), the live subtitle providers for Stagertext write, “The difference between our captions and autocaptions is...the difference between listening to a person or a machine”. Whilst automated solutions may be powerful, they are also limited (*The Economist*, 2020) and it may be that the live event setting will prove to be an example of the latter.

To what extent ASR does spread as a form of access provision in this setting remains to be seen, yet it is partly because of its spread in society that makes the timing of this study so apt. Our world is one where speech recognition, in its different forms, is an undeniable part of our lives, from Siri and Cortana on our phones, to Alexa in our home. In the anti-Brexit marches of March 2019, one sign read, “Alexa, cancel Brexit!”, showing – if only in desire! - just how integrated our interaction with SR has become.



Fig. 9.3: Anti-Brexit placards

The link between automated and adaptive technology is also growing, as the range of accessibility features incorporated within products continues to be refined (Medcalf, 2019). Adverts for Alexa on television feature its use by a blind woman (Farey-Jones, 2019), and now, when connected to the Echo show, the automatic captions created make it more accessible for DH users as well. It certainly feels like awareness of accessibility is increasing.

9.5.2 Awareness of accessibility

Recent media headlines place much attention on issues surrounding deafness: the UK Government's failure to provide a BSL interpreter at its coronavirus briefings being ruled illegal by the High Court (O'Dell, 2021), the finding against the Little Mix concert promoter, who was deemed to have discriminated against three Deaf mothers by failing to provide a BSL interpreter, a finding which prompted Wembley Stadium (2021) – and perhaps other venues - to commit to BSL interpretation of all concerts. There was a Deaf contestant on *Strictly Come Dancing* (2021), Rose Ayling-Ellis, who went on to win the Glitterball (Saner, 2021; *BBC News*, 2021a), and there are more deaf and disabled storylines on television. Yet in amongst these headlines come the stories of a blind woman being refused entrance into a restaurant with her guide dog (Petherick, 2021), the cut to the universal credit supplement that will affect so many families (Butler, 2021a, 2021b), and the MacTaggart Lecture at the Edinburgh Television Festival (2021), where the scriptwriter Jack Thorne explained how he believes the industry has failed disabled people “utterly and totally”. Yes, awareness is growing, but the societal and structural sea change that is so desperately needed has not yet happened. Marginalisation and discrimination are rife and equity of access, be it to arts and cultural events, or fundamental requirements like food and drink or heating and medical attention is far from present. However, awareness that change is needed is slowly seeping through.

The events of the last year and a half, from the pandemic itself to the killing of George Floyd in the USA (Silverstein, 2021; *BBC News*, 2021c) and the prominence of Black Lives Matters (Campbell, 2021) have, in my opinion, made the UK wake up and confront

aspects of its past that it might prefer to ignore. To fully address all the inequalities and inequities that exist in society is a huge task and it demands change at all levels, but gradually signs of that change can be seen; they now need to be extended further, so that those with the power and authority to legislate for change share this understanding and begin to act upon it. I hope that, if only in a small way, this study is a contribution towards this and that the disciplines of translation studies and accessibility studies can contribute further to driving this change.

9.5.3 The future of translation studies and accessibility studies

In Chapter Two I described TS and AS as ubiquitous, with language and access being relevant to all we do; as such, they potentially have a significant role to play in increasing this awareness and implementation of access in the society around us. This may be particularly true of AS, as an emerging discipline and with its broader scope that could be incorporated into and through any discipline. As well as discrete courses in AS, why not incorporate this understanding into any course offered? So that students learning any subject from architecture to web design, from law to medicine also learn about the need for access and accessibility, and the importance of effective translation or interpretation? Those attending may not learn how to provide these skills themselves, but recognising how essential they are might ensure that they follow this good practice as they move forwards in their careers, a little like the stealth access approach I referred to and which I have tried to embed in the Model for Participatory Engagement. Once the need for access is understood and tools for embedding it are known, the logical move is to start to include it. And perhaps from there, as well as incorporating access, people will begin to call out those instances where it is absent. The principle of 'Nothing about us without us' and user involvement is an essential part of this and should be prioritised.

9.5.4 Be the change that is needed

In Chapter One, I wrote of the journey that led me to where I am today and how I came to write a PhD on accessibility and translation studies. As I have carried out this doctoral

research project, I have had many opportunities to question and explore who I am and where I stand in relation to these questions of access that are unfolding around me. What place do I hold, as a person and as a researcher, and how do my actions contribute to inequalities that exist or act as a catalyst for change when change is needed? These are not questions to which I always have a ready answer, but they are ones that I carry with me in all I do. As I said at the start, “I am as much a creation of (this study), as it is something I have created”.

If I wanted to try and capture the effect this journey has had on me, I might refer to the words on a button pin that I discovered recently. In September 2021, I was at Pride with friends and as we roamed amongst the selection of pins and stickers there, I saw one that read, “Take up space”. For such a simple phrase, the way it resonated with me was striking. As a researcher, lecturer, trainer, it is perhaps expected of me that I am someone who can stand up, speak and share ideas, in other words, take up space. Yet, taking up space is something I have often struggled with or not wanted to do on a more personal level. Whether afraid of what people might say, or simply not wishing to show up, I have often stepped back and remained silent. Today, I step forward far more. In the course of this work and research, I have found my voice and learned to use it. I am able to show up and be heard on both a personal and professional level, be heard as myself. And, increasingly, it feels like using the voice we have and taking action is an essential part of calling out inequalities and contributing to the allyship and advocacy that is so greatly needed.

However, alongside this sense of embodying the space I occupy, I have also understood that this space is not always mine to take. Understanding when and how to speak up is also what the principle of ‘Nothing about us without us’ demands of us – as well as being present at the debate, there must be space for everyone’s voice to be heard. Knowing when it is or is not my place to speak is as important as having the confidence of speaking up when I need to and it is this exploration of exactly where this balance lies, and what words and actions are required in situations where change is needed, that I take with me as I submit this thesis.

In the epigraph to this chapter, I refer to the words of Peter Tatchell: “Don't accept the world as it is, dream of what the world could be - and then help make it happen.” Perhaps I am dreaming big, but I do dream of a world where the difference between people is more accepted, different needs are understood and incorporated as a matter of course, and where social injustice and epistemic injustice eventually come to an end. I also firmly believe that to achieve this, there is a need for solidarity and a recognition of intersectionality in the approach we take. As Faye (2021) demands in her argument for trans liberation and justice, what is required is “a manifesto for change, and a call for justice and solidarity between all marginalised people and minorities” (Penguin, 2021); that is how all groups will be heard and none will be forgotten. This is the change that is required – and a change that can be achieved. The way we interact with people, the language we use and the fast judgements that we all make are ingrained, but can also evolve.

A stark reminder of this came recently, as I watched *Subnormal: A British Scandal* (Shannon, 2021), a documentary on how “in the 1960s and 1970s, hundreds of black children in Britain were labelled as “educationally subnormal” and sent to schools for pupils who were deemed to have low intelligence” (*BBC News*, 2021b). As I sat watching with incredulity, unaware of this part of British history, I heard the term ‘slow at learning’ being used, the very same term I remember being used to describe my brother; another term we used, ‘slightly mentally handicapped’ also came to mind. The effect was jarring; the inadequacy of both terms, their failure to capture who my brother was, or the support he might have benefitted from, and the fact that black children were routinely labelled and treated in this way. How much our language has changed; now our perceptions and actions must follow. These injustices are gradually being called out and it is our responsibility to join this cause. We are going to make mistakes and use the wrong labels, we may say or do something that is ableist without realising, or speak when it really is not our place, but we can also learn from our mistakes, show an openness towards being corrected, and call out those moments where something different is

required. Being able to show up with integrity and honesty and self-reflection is also part of what we learn as we judge when to take up space.

No one of us experiences the world in exactly the same way. Our needs will differ, yet we can all work together towards creating a society which is more open and understanding and in which every one of us is able to participate and engage.

Glossary

Access: The idea of being able to enter, use or interact with a particular service or environment, but more importantly being able to fully engage and participate with(in) it. Like translation, I consider this to be ubiquitous (Blumcynski, 2016), and to often involve both a product and process (Munday, 2012:8; section 2.3.1).

Accessibility Studies: An emerging discipline which explores matters relating to accessibility across a range of fields (section 2.3.1.1).

Action research: A research methodology which involves “learning by doing” (O’Brien, 2001; section 4.1.2).

Antenna delay: An intentional delay when live material is broadcast on television; this can also be referred to as broadcast delay (section 8.6.1).

As-live: i) I use this term to refer to the respeaking I carried out at test events, which was not transmitted to the audience (section 4.2.2). ii) This term can also refer to blocked subtitles that are prepared in advance and cued out by the respeaker as the programme is broadcast (section 31). See also semi-live.

Audience: This term is used in different ways throughout the study. The focus group research I conducted included two audience groups, d/Deaf, deafened and hard of hearing participants and non-native speakers of English (Chapter Five). This term also refers to the people who attended each of the research events, and to people who might attend events in the future.

Audiovisual Translation: The academic discipline that initially incorporated areas of translation involving image and sound (Díaz Cintas and Remael, 2007: 9) including subtitling, dubbing and audio description (section 2.3.11).

Captions: i) the pre-prepared captions that are cued out live at the theatre, for example by Stagertext (section 3.1). ii) the term used to refer to same-language subtitles in the context of second language learners (section 3.5.2).

Cognitive/empirical turn in audiovisual translation: This turn has led to a closer consideration of the mental processes of the translator and the audience's response (Chaume, 2018: 53-54; section 2.3.1.1).

Discriminatory normative assumptions: Beliefs about how things are or should be; often built on what we are familiar with, they are discriminatory in their nature as they are used as a conscious or unconscious background to judge others (section 2.1.2). See also normative assumptions.

Epistemic justice: This form of justice demands a shift in our values and understanding with respect to people who are different to us. There are two types, testimonial and hermeneutical (Fricker, 2007; section 2.2.2).

The Equality Act, 2010: The legislative framework in the UK that governs the provision of access at live events, through the principle of reasonable adjustments being made (section 3.3).

Fidelity: A concept within translation studies that relates to the faithful relationship between the source and target texts (section 2.3.1.2). See also loyalty.

Human rights model of disability: This model is sometimes thought of as a model of disability policy, rather than of disability per se. It states that impairments cannot be used as a reason for denying disabled people any human right (Maastricht University, 2019; section 2.2.3.1).

Hybrid: i) Hybrid programmes contain sections of blocked semi-live subtitles and sections which are respoken. ii) Hybrid event scenarios include those where some people attend online whilst others attend in person.

Inclusion: The principle and practice of making sure that everyone in society is able to access resources and opportunities (Collins Online Dictionary, 2021a).

Intersectionality: "A way of understanding the complexity in the world, in people and in human experiences" by acknowledging that "people's lives and the organisation of power in a given society are better understood as being shaped not by a single axis of social division, be it race or gender or class, but by many axes that work together and influence each other" (Collins & Bilge, 2016:2; section 2.2.1).

Live events: An unscripted, or partially scripted event, happening in real-time, where the audience attend in-person and which is not watched in its entirety through a screen; examples include Q&A sessions, discussions, public talks or museum tours (section 3.4).

Live subtitling: The process of creating subtitles in real-time. The subtitles may be created using different techniques, for example, respeaking or STTR (section 3.1).

Loyalty: A concept within translation studies, where the translator is equally committed to the source and target sides (Nord, 1997: 125; section 2.3.1.2). See also fidelity.

Media accessibility: “Access to media and non-media objects, services and environments through media solutions, for any person who cannot or would not be able to, either partially or completely, access them in their original form (Greco, 2019a: 18; section 2.3.1.1).

Media access quality: This refers to a series of considerations which have been proposed when discussing what quality refers to in the context of media access (Greco and Jankowska, 2019). Of particular relevance in this study are ‘quality of experience’ for the audience or user, and ‘quality of service’ that is provided to them (p.8; section 6.2.2).

Medical model of disability: A model of disability where it is the individual who is considered as the ‘problem’ and the fix for their disability comes through it being “prevented, cured or contained” (Inclusion London, 2015: 6; section 2.2.3.1).

NER: This is the name of the model used to analyse respeaking quality (Romero-Fresco and Martínez, 2015); it is taken from the calculation used within it, where N stands for the total number of words, E for the penalty applied for edition errors, and R for the penalty applied for recognition errors.

NERLE: An adapted version of the NER model for use at Live Events (LE) (section 8.6).

Non-native speakers of English: I use this term to refer to my second audience group, people who are either living, working or visiting the UK and who took part in this study, or who might attend future events made accessible through live subtitles, who would name the language they grew up with as one other than English or a sign language (section 5.3.4). In using it, I am aware that although the term ‘native speaker’ feels

familiar, there are many different ways that it can be interpreted and that some debate exists around its use (see for example (Marek, 2017) .

Normative assumptions: Beliefs about how things are or should be, which are often built on what we are familiar with. (section 2.1.2). See also discriminatory normative assumptions.

Ofcom: The regulator for the communications services in the UK, who set quotas for the amount of accessible content across television and On Demand channels and who conducted a study into the quality of live subtitling on television (Ofcom, 2013).

Poietic design: “An access-oriented methodology which embeds the fundamental features of accessibility within the design process” and which enshrines the responsibility that “we all have collectively and individually as co-creators and agents” within this process (Greco, 2019a: 24; section 2.2.3.2).

Presenter: The person or people who speak(s) at events or present or lead them.

Pure respeaking: I use this term to refer to scenarios where respeaking is used for the duration of a programme or event, without subtitles being cued out live for any portion of it.

Respeaking: The production of live subtitles by a person using speech recognition software (section 3.1).

Respect: A notion I suggest as most appropriate for this study, rather than loyalty or fidelity which are commonly discussed in translation studies. Here, the relationship between parties is one built on mutual respect and trust (section 2.3.1.2).

Semi-live: This term refers to blocked subtitles that are prepared in advance in order to be cued out live, often alongside respoken passages, for example during the news (section 3.5). See also as-live.

Stagetext: “A deaf-led charity that makes arts and culture accessible to deaf, deafened and hard of hearing people” (<https://www.stagetext.org/about-us/about-stagetext/>).

Streamtext: The presentation software used to display live subtitles at the museum tour events in this study (section 6.4.4). See also Text on Top.

Speech-to-text: Speech-to-text is an umbrella term for a range of different modalities, including respeaking alongside keyboard-based techniques and automatic speech recognition, which are used to create live intralingual or interlingual subtitles (section 3.1).

Speech-to-text reporter/reporting: In contrast to respeaking, a speech-to-text reporter uses a specialist phonetic keyboard to create live subtitles (Stagetext, 2021a).

Social justice: The demand for a fairer and more equitable world with respect to how wealth, opportunities and privileges are distributed within society (LegalDictionary Content Team, 2016; section 2.2.2.).

Social model of accessibility: A model which considers access from the perspective of the specific needs that every person will have in different situations; it incorporates within it a universalist account of access alongside user-centred and proactive approaches to access (section 2.2.3.2). It can only be fully understood when combined with the principles of poietic design.

Social model of disability: A model of disability where people are disabled by barriers within society, rather than being 'victims' of their impairments or conditions. (Graeae Theatre Company, 2016; section 2.2.3.1).

Sociological turn in audiovisual translation: This turn has seen increased consideration being given to the role of translator and to where power lay in the process of creating audiovisual texts (Chaume, 2018: 52; section 2.3.1.1).

Subtitles: I use this term to refer to the subtitles that appear on television, which may be live (created by respeaking) or pre-recorded (prepared in advance) (section 3.1).

Surtitles: The term used for the pre-prepared text displayed above the stage at the opera, and which offers a translation of what is said and sung (section 3.1).

SmartMics: Specialised microphones that fit over the mouth (and nose) and muffle the voice of the respeaker, allowing them to work in the main event room (section 6.4.2).

Text on Top: The presentation software used to display subtitles at the seated research events in this study (section 6.4.4.). See also Streamtext.

Translation: Translation involves language, yet also moves beyond it in order to fully capture the vital and complex whole that is created when different media and modalities interact in different contexts in the world around us. In this way, as well as being a product and a process (Munday, 2012: 8), translation can be considered ubiquitous (Blumcynski, 2016; section 2.3.1).

Translation studies: The wider academic discipline of translation, which could be seen as including audiovisual translation (section 2.3).

Universal Design: “The design of products and environments to be usable by all people, to the greatest extent possible, without the need for adaptation or specialised design” (Mace, 1998; section 2.2.3.2.)

Universalist account of MA: “The universalist account defines MA as concerning access to media and nonmedia objects, services and environments through media solutions, for any person who cannot or would not be able to, either partially or completely, access them in their original form. The universalist account focuses on the processes involved in the interaction between users’ specificities, the particular contexts in which they act or are placed, and the means to address such specificities in those contexts” (Greco, 2019b: 27; section 2.2.3.2).

Vulnerability: i) This term may refer to people who are weak or without protection (Collins Online Dictionary, 2021b) (Collins Online Dictionary, 2021), or who may be seen to depend on access; ii) In the context of subtitling (Díaz Cintas and Remael, 2007: 55), vulnerability refers to the idea that a difference exists between the spoken original and written translation, which may be visible to some who use it (section 2.3.1.2).

References

- Abercrombie, N. and Longhurst, B. (1998) *Audiences: A Sociological Theory Of Performance And Imagination*. London: Sage.
- Adelman, C. (1993) 'Kurt Lewin And The Origins Of Action Research', *Educational Action Research*, 1(1), pp. 7–24. doi:10.1080/0965079930010102.
- Ahmed, S. (2012) *On Being Included Racism And Diversity In Institutional Life*. Durham and London: Duke University Press.
- Aho, T. and Alter, G. (2018) "'Just Like Me, Just Like You": Narrative Erasure As Disability Normalization In Children's Picture Books', *Journal Of Literary & Cultural Disability Studies*, 12(3), pp. 303–319. doi:10.3828/jlcds.2018.24.
- Ai-Media (no date a) 'Ai-Live Style Guide'. Available at: <https://kb.ai-media.tv/wp-content/uploads/2018/08/Ai-Live-Style-Guide.pdf>. (Accessed: 27 December 2021).
- Ai-Media (no date b) *The Magic Of Live Captions For Students With ASD*. Available at: <https://blog.ai-media.tv/blog/magic-live-captions-students-asd> (Accessed: 21 November 2021).
- Ai-Media (2021) *Disabled Students' Allowance (DSA) And Ai-Media, Ai-Media Creating Accessibility, One Word At A Time*. Available at: <https://www.ai-media.tv/industry/disabled-students-allowance-and-ai-media/> (Accessed: 27 December 2021).
- Alcoff, L.M. (2010) 'Epistemic Identities', *Episteme*, 7(2), pp. 128–137. doi:10.3366/E1742360010000869.
- All Sign Images (2021) *Toilets Disabled Men/Women*.
- Al-Othman, H. (2015) *Starbucks Tells Woman She Is 'Not Disabled Enough' To Use Toilet*. Available at: <https://www.standard.co.uk/news/uk/hold-waiting-for-additional-comments-starbucks-apologises-for-telling-woman-she-was-not-disabled-enough-to-use-the-toilet-a2916486.html> (Accessed: 9 November 2021).

- Althique, A. (2016) *Transnational Audiences: Media Reception On A Global Scale*. Cambridge: Polity Press.
- Anderson, E. (2012) 'Epistemic Justice As A Virtue Of Social Institutions', *Social Epistemology*, 26(2), pp. 163–173. doi:10.1080/02691728.2011.652211.
- AOHL (2015) *Hearing Matters*. Available at: <https://rnid.org.uk/wp-content/uploads/2020/05/Hearing-Matters-Report.pdf> (Accessed: 21 November 2021).
- AOHL (2018) *Understanding Tinnitus*. Available at: <https://rnid.org.uk/wp-content/uploads/2020/05/Understanding-tinnitus-leaflet.pdf> (Accessed: 21 November 2021).
- Apone, T., Brooks, M. and O'Connell, T. (2010) 'Subtitle Accuracy Metrics Project. Subtitle Viewer Survey: Error Ranking Of Real-Time Subtitles In Live Television News Programmes. WGBH National Center for Accessible Media, Boston.' WGBH National Center For Accessible Media, Boston. Available at: http://ncam.wgbh.org/invent_build/analog/subtitle-accuracy-metrics (Accessed: 1 July 2015).
- Arumí Ribas, M. and Romero-Fresco, P. (2008) 'A Practical Proposal For The Training Of Respeakers 1', *Journal Of Specialised Translation*, 10, pp. 106–127. Available at: https://jostrans.org/issue10/art_arumi.php (Accessed: 31st December 2021).
- Atewologun, D. and Mahalingam, R. (2018) 'Intersectionality As A Methodological Tool In Qualitative Equality, Diversity And Inclusion Research', in Booyen, L., Bendl, R. and Pringle, J. (eds.) *Handbook Of Research Methods In Diversity Management, Equality And Inclusion At Work*. Edward Elgar Publishing Ltd.
- autisticality (2014) 'Neurodiversity, Language, And The Social Model', *autisticality*, 27 October. Available at: <https://autisticality.com/2014/10/27/neurodiversity-language-and-the-social-model/> (Accessed: 23 July 2020).
- Barker, M.J. and Iantaffi, A. (2019) *Life Isn't Binary: On Being Both, Beyond, And In-Between*. London and Philadelphia: Jessica Kingsley Publishers.

- Barnard, H. (2021) 'COVID-19 Day Of Reflection: Injustices Must Not Be Forgotten', *JRF*, 19 March. Available at: <https://www.jrf.org.uk/blog/covid-19-day-reflection-injustices-must-not-be-forgotten> (Accessed: 14 November 2021).
- Baum, W.M. (2000) 'Being Concrete About Culture And Cultural Evolution', in Tonneau, F. and Thompson, N.S. (eds.) *Perspectives In Ethology: Evolution, Culture, And Behavior*. Boston, MA: Springer US (Perspectives In Ethology), pp. 181–212. doi:10.1007/978-1-4615-1221-9_7.
- BBC (no date) *Access Services At The BBC*. Available at: <https://www.bbc.com/aboutthebbc/whatwedo/publicservices/bbc.com/aboutthebbc/whatwedo/publicservices/accessservices> (Accessed: 21 November 2021).
- BBC (2021) *BBC Subtitle Guidelines Version 1.1.9*. Available at: <https://bbc.github.io/subtitle-guidelines/> (Accessed: 20 December 2021).
- BBC News (2020a) 'Coronavirus Confirmed As Pandemic By World Health Organization', 11 March. Available at: <https://www.bbc.co.uk/news/world-51235105> (Accessed: 2 November 2021).
- BBC News (2020b) 'Coronavirus: Daily Downing Street Press Conference Scrapped', 23 June. Available at: <https://www.bbc.com/news/uk-politics-53155905> (Accessed: 23 July 2020).
- BBC News (2021a) 'Strictly Come Dancing 2021: Rose Ayling-Ellis Becomes First Deaf Winner', 18 December. Available at: <https://www.bbc.com/news/entertainment-arts-59599740> (Accessed: 18 December 2021).
- BBC News (2021b) 'Subnormal: A British Scandal', 20 May. Available at: <https://www.bbc.co.uk/news/av/uk-57178342> (Accessed: 14 December 2021).
- BBC News (2021c) 'Who Was George Floyd And What Happened To Derek Chauvin?', 24 May. Available at: <https://www.bbc.com/news/world-us-canada-56270334> (Accessed: 14 December 2021).
- BDA (no date) *Help & Resources, British Deaf Association*. Available at: <https://bda.org.uk/help-resources/> (Accessed: 21 November 2021).

- de Beauvoir, S. (1949) *Le Deuxième Sexe*. Paris: France Loisirs. Available at: http://elfege.com/bibliotheque/Beauvoir/Le-deuxieme-sexe-tome-1-Simone-de-Beauvoir-FRENCHPDF.COM_.pdf (Accessed: 10 November 2021).
- Berthaud, S. and Mason, S. (2018) 'Embedding Reflection Throughout The Postgraduate Translation Curriculum: Using Communities Of Practice To Enhance Training', *The Interpreter And Translator Trainer*, 12(4), pp. 388–405. doi:10.1080/1750399X.2018.1538847.
- Betteley, C. (2020) 'Coronavirus: Mask Wearing "Risks Isolating" Deaf People', *BBC News*, 14 May. Available at: <https://www.bbc.com/news/uk-wales-52659083> (Accessed: 23 July 2020).
- Bickenbach, J.E., Chatterji, S., Badley, E.M. & Üstün, T.B. (1999) 'Models Of Disablement, Universalism And The International Classification Of Impairments, Disabilities And Handicaps', *Social Science & Medicine*, 48, pp. 1173–1187.
- Block, D. and Corona, V. (2016) 'Intersectionality In Language And Identity Research', in Preece, S. (ed.) *The Routledge Handbook Of Language And Identity*, pp. 507–522.
- Blumczynski, P. (2016) *Ubiquitous Translation*. Oxford: Routledge.
- Blumenfeld, H.K., Marian, V. and Kaushanskaya, M. (2007) 'The Language Experience And Proficiency Questionnaire (LEAP-Q): Assessing Language Profiles In Bilinguals And Multilinguals', *Journal Of Speech Language And Hearing Research*, 50(4), pp. 940–967.
- Boéri, J. and Maier, C. (eds.) (2010) *Compromiso Social Y Traducción/Interpretación - Translation/Interpreting And Social Activism*. Granada, Spain: ECOS Traductores e Intérpretes por la Solidaridad.
- Bogucki, L. (2004) 'The Constraint Of Relevance In Subtitling', *The Journal Of Specialised Translation*, 1, pp. 71–88. Available at: https://jostrans.org/issue01/art_bogucki_en.php (Accessed: 31 December 2021).

- Bogucki, Ł. (2010) 'The Applications Of Action Research To Audiovisual Translation', in Bogucki, Ł. and Krzystof, K. (eds.) *Perspectives In Audiovisual Translation*. Frankfurt am Main: Peter Lang (Łódz Studies in Language, 20), pp. 7–18.
- Bowker, L. (2002) *Computer-Aided Translation Technology*. Canada: University of Ottawa Press.
- Brannen, J. (1992) *Mixing Methods: Qualitative And Quantitative Research*. Avebury: Ashgate Publishing Ltd.
- Braun, V. and Clarke, V. (2006) 'Using Thematic Analysis In Psychology', *Qualitative Research In Psychology*, 3(2), pp. 77–101. doi:10.1191/1478088706qp063oa.
- Braun, V. and Clarke, V. (2013) *Successful Qualitative Research: A Practical Guide For Beginners*. London: SAGE.
- Braun, V. and Clarke, V. (2019) 'Reflecting On Reflexive Thematic Analysis', *Qualitative Research In Sport, Exercise And Health*, 11(4), pp. 589–597.
- Bricker, S. (2015) 'What Sounds Would You Miss With Hearing Loss?', 10 September. Available at: <https://www.starkey.com/blog/articles/2015/09/sounds-missed> (Accessed: 22 November 2021).
- Briggs, S.R. and Cheek, J.M. (1986) 'The Role Of Factor Analysis In The Development And Evaluation Of Personality Scales', *Journal Of Personality*, (54), pp. 106–48.
- British Tinnitus Association (no date) *What Is Tinnitus?*, *British Tinnitus Association*. Available at: <https://www.tinnitus.org.uk/what-is-tinnitus> (Accessed: 21 November 2021).
- Burnes, B. (2008) 'Complexity Theories', in *The SAGE Dictionary Of Qualitative Research Management*. London: SAGE Publications Ltd. Available at: <https://dx.doi.org/10.4135/9780857020109>.
- Butler, C. (1985) *Statistics In Linguistics*. Oxford: Blackwell.
- Butler, P. (2021a) 'Universal Credit £20 Top-Up Will End This Autumn, MPs Are Told',

- The Guardian*, 7 July. Available at: <https://www.theguardian.com/society/2021/jul/07/universal-credit-20-top-up-will-end-this-autumn-mps-are-told> (Accessed: 14 December 2021).
- Butler, P. (2021b) 'Universal Credit Cut Will Lead To More UK Children In Care – Study', *The Guardian*, 26 September. Available at: <https://www.theguardian.com/society/2021/sep/26/universal-credit-cut-will-lead-to-more-uk-children-in-care-study> (Accessed: 14 December 2021).
- Campbell, A. (2021) 'What Is Black Lives Matter And What Are The Aims?', *BBC News*, 12 June. Available at: <https://www.bbc.com/news/explainers-53337780> (Accessed: 14 November 2021).
- Campion, J. (1993) *The Piano*.
- Carbonell, K.B., Stalmeijer, R.E., Könings, K.D., Segers, M., & van Merriënboer, J.J.G. (2014) 'How Experts Deal With Novel Situations: A Review Of Adaptive Expertise', *Educational Research Review*, (12), pp. 14–29.
- Carr, L. (no date) 'Ouch! Liz Carr: Silent Witness Star Reveals Film Role'. (Ouch!). Available at: <https://www.bbc.co.uk/sounds/play/p082bvs4>.
- Carr, L. and Darke, P. (2012) 'The Principles And Practice Of Disability Equality Training (DET)', in Carr, L., Darke, P. and Kuno, K. (eds.) *Disability Equality Training - Action For Change*. Sdn Bhd, Malaysia: MPH Publishing.
- Casualty* (2020). BBC One. 11 July 2020, 20:25.
- CCAC (2016a) *Support For Live Event Captioning: Apply For CCAC Grant*. Available at: <http://ccacaptioning.org/ccac-sponsorship-for-cart-a-new-captioning-advocacy-program> (Accessed: 8 April 2019).
- CCAC (2016b) *What Is CART And FAQs – Find A CART Captioning Provider*. Available at: <http://ccacaptioning.org/faqs-cart> (Accessed: 8 April 2019).
- Chandra (2018) 'Neural Networks For Dummies: A Comprehensive Guide', *upGrad blog*, 6 February. Available at: <https://www.upgrad.com/blog/neural-networks-for-dummies-a-comprehensive-guide/> (Accessed: 18 November 2021).

- Channel 4 (2013) *Channel 4 Response To Ofcom's Consultation On The Quality Of Live Subtitling*. Available at: https://www.ofcom.org.uk/__data/assets/pdf_file/0028/45892/channel_4.pdf (Accessed: 21 November 2021).
- Charlton, J.I. (1998) *Nothing About Us Without Us*. University of California Press. doi:10.1525/california/9780520207950.003.0001.
- Chaume, F. (2018) 'An Overview Of Audiovisual Translation: Four Methodological Turns In A Mature Discipline', *Journal Of Audiovisual Translation*, 1(1), pp. 40–63. doi:10.47476/jat.v1i1.43.
- Choi, B.C.K. and Pak, A.W.P. (2007) 'Multidisciplinarity, Interdisciplinarity, And Transdisciplinarity In Health Research, Services, Education And Policy: 2. Promotors, Barriers, And Strategies Of Enhancement', *Clinical & Investigative Medicine*, 30(6), pp. E224–E232. doi:10.25011/cim.v30i6.2950.
- Chon, M. (2000) 'Erasing Race?: A Critical Race Feminist View Of Internet Identity-Shifting', *Journal Of Gender, Race And Justice*, 3(2), pp. 439–474.
- Clason, D. (2021) *Mild Hearing Loss, Healthy Hearing*. Available at: <https://www.healthyhearing.com/report/7733-Living-with-mild-hearing> (Accessed: 11 November 2021).
- Clayton, M. and Williams, A. (2004) *Social Justice*. Malden, Mass.; Oxford: Blackwell Publishing.
- CMU (2021) 'Attitude Is Everything Launches Accessible Reopening Checklist For Venues And Festivals.' 11 June. Available at: <https://completemusicupdate.com/article/attitude-is-everything-launches-accessible-reopening-checklist-for-venues-and-festivals/>. (Accessed: 20 November 2022).
- Cock, M., Bretton, M., Fineman, A., France, R., Madge, C. & Sharpe, M. (2018) *State Of Museum Access 2018: Does Your Museum Website Welcome And Inform Disabled Visitors?* Available at: <https://vocaleyes.co.uk/state-of-museum-access-2018/> (Accessed: 2 November 2021).

- Cock, M., Pountney, M., Sharpe, M., Taylor, N. & Thom, J. (2019) *State Of Theatre Access 2019*. Available at: <https://vocaleyes.co.uk/state-of-theatre-access-2019/> (Accessed: 2 November 2021).
- Cock, M., Madge, C. and Sharpe, M. (2020) *Museum And Heritage Access 2020 Survey*. Available at: <https://vocaleyes.co.uk/museum-and-heritage-access-survey-2020/> (Accessed: 1 November 2021).
- Cohen, L. and Manion, L. (1985) *Research Methods In Education*. London: Croom Helm.
- Collins Online Dictionary (2021a) 'Inclusion', *Collins Online Dictionary*. Available at: <https://www.collinsdictionary.com/dictionary/english/inclusion> (Accessed: 2 November 2021).
- Collins Online Dictionary (2021b) 'Vulnerable'. Collins Online Dictionary (Accessed: 2 November 2021).
- Collins, P.H. (1993) 'Toward A New Vision: Race, Class, And Gender As Categories Of Analysis And Connection', *Race, Sex & Class*, 1(1), pp. 25–45.
- Collins, P.H. and Bilge, S. (2016) *Intersectionality*. Cambridge: Polity Press.
- Committee on Disability in America and Institute of Medicine (2007) *The Future Of Disability In America*. Edited by M.J. Field and A.M. Jette. Washington, DC: National Academies Press.
- Convention On Rights Of Persons With Disabilities* (2007). Available at: <https://www.un.org/development/desa/disabilities/convention-on-the-rights-of-persons-with-disabilities/convention-on-the-rights-of-persons-with-disabilities-2.html> (Accessed: 31 December 2021).
- Cook, T. (2009) 'The Purpose Of Mess In Action Research: Building Rigour Through A Messy Turn', *Educational Action Research*, 17(2), pp. 227–291.
- Costa, R.M. (2011) 'A Selectionist Account Of De Novo Action Learning', *Current Opinion in Neurobiology*, 21(4), pp. 579–586. doi:10.1016/j.conb.2011.05.004.
- Cravo, A. (1999) *Promoção Da Autonomia Na Aprendizagem Da Língua Alemã Para Fins Específicos – Estudo De Casos No Âmbito Da Gramática*. Masters Dissertation. University Of Nottingham.

- Cravo, A. and Neves, J. (2007) 'Action Research In Translation Studies', *Journal of Specialised Translation*, 7, pp. 92–107. Available at: https://jostrans.org/issue07/art_cravo.php (Accessed: 31 December 2021).
- Crenshaw, K. (1991) 'Mapping The Margins: Intersectionality, Identity Politics, And Violence Against Women Of Color', *Stanford Law Review*, 43(6), pp. 1241-1299. doi:10.2307/1229039.
- Criado Perez, C. (2019) *Invisible Women Exposing Data Bias In A World Designed For Men*. London: Chatto & Windus.
- Crossick, G. and Kaszynska, P. (2016) *Understanding The Value Of Arts & Culture | The AHRC Cultural Value Project*. AHRC. Available at: <https://ahrc.ukri.org/documents/publications/cultural-value-project-final-report/> (Accessed: 31 December 2021).
- CRTC (2019) *Broadcasting Notice Of Consultation CRTC 2019-9, Canadian Radio-Television And Telecommunications Commission*. Available at: <https://crtc.gc.ca/eng/archive/2019/2019-9.htm> (Accessed: 21 December 2021).
- Daily Mail Reporter (2009) 'Care Home Boss "Told Police 'S*** Happens' After Autistic Resident Choked On Ham Sandwich"', *Daily Mail*, 20 April. Available at: <https://www.dailymail.co.uk/news/article-1172128/Care-home-boss-told-police-s--happens-autistic-resident-choked-ham-sandwich.html> (Accessed: 31 January 2020).
- Dainty, S. (2020) 'Casualty Made History With Jade Lovall Episode', *Digital Spy*, 11 July. Available at: <https://www.digitalspy.com/soaps/casualty/a33274893/casualty-jade-lovall-episode-deaf-representation-behind-the-camera/> (Accessed: 11 November 2021).
- Danan, M. (2004) 'Captioning And Subtitling: Undervalued Language Learning Strategies', *Meta*, 49(1), pp. 67–77. doi:10.7202/009021ar.
- Dangerfield, K. (2018) 'The Value Of Difference In Media Accessibility Quality.' *Understanding Media Accessibility Quality (UMAQ)*, Barcelona, 4 June.

- Dangerfield, K. (2022) *Within Sound And Image: The Practice And Theory Of Accessible Filmmaking*. Unpublished doctoral dissertation. The University of Roehampton.
- Davies, H.J. (2021) 'Channel 4 Subtitles To Remain Unavailable Until Mid-November', *The Guardian*, 19 October. Available at: <https://www.theguardian.com/media/2021/oct/19/channel-4-outage-subtitles-remain-unavailable-mid-november> (Accessed: 26 December 2021).
- Davis, L.J. (1995) *Enforcing Normalcy - Disability, Deafness And The Body*. London; New York: Verso.
- Davis, M. (2018) 'Speaker-Dependent vs. Speaker-Independent Software', *RFgen*, 2 May. Available at: <https://www.rfgen.com/blog/what-is-the-difference-between-speaker-dependent-software-and-speaker-independent-software/> (Accessed: 18 November 2021).
- Davitti, E., Moores, Z., Dawson, H. and Fryer, L. (2020) 'Real-Time Interlingual Speech-To-Text Via Speech Recognition'. *7th Live Subtitling and Accessibility Symposium*, Universitat Autònoma de Barcelona, 5 November. Available at: <https://www.youtube.com/watch?v=ZLTsN4lb7nQ> (Accessed: 17 November 2021).
- Davitti, E. and Sandrelli, A. (2020) 'Embracing The Complexity: A Pilot Study On Interlingual Respeaking', *Journal of Audiovisual Translation*, 3(2), pp. 103–139. doi:10.47476/jat.v3i2.2020.135.
- Dawson, H. (2019) 'Feasibility, Quality And Assessment Of Interlingual Live Subtitling: A Pilot Study', *Journal Of Audiovisual Translation*, 2(2), pp. 36–56. doi:10.47476/jat.v2i2.72.
- Dawson, H. (2020) *Interlingual Live Subtitling: A Research-Informed Training Model For Interlingual Respeakers To Improve Access For A Wide Audience*. Unpublished doctoral dissertation. The University of Roehampton.
- Deaf Umbrella (no date) 'Deaf Awareness Tips For The Workplace'. Available at: https://www.deafumbrella.com/_files/ugd/69a578_b4aeec95ca134dd9a5b2066d5926f08c.pdf (Accessed: 21 December 2021).

- Degener, T. (2014) 'A Human Rights Model Of Disability'. Available at: https://www.researchgate.net/publication/283713863_A_human_rights_model_of_disability (Accessed: 31 December 2021).
- Dehghani, M., Khooshabeh, P., Nazarian, A. & Gratch, J. (2015) 'The Subtlety Of Sound: Accent As A Marker For Culture', *Journal Of Language And Social Psychology*, 34(3), pp. 231–250. doi:10.1177/0261927X14551095.
- Depot (2017) *Relaxed - Lewes Depot*. Available at: <https://lewesdepot.org/access/relaxed> (Accessed: 14 December 2021).
- Desblache, L. (2019) *Music And Translation New Mediations In The Digital Age*. London: Palgrave Macmillan.
- Di Giovanni, E. (2018) 'Participatory Accessibility: Creating Audio Description With Blind And Non-Blind Children', *Journal Of Audiovisual Translation*, 1(1), pp. 155–169. doi:10.47476/jat.v1i1.50.
- Di Giovanni, E. and Gambier, Y. (2018) *Reception Studies And Audiovisual Translation* (1 online resource (xii, 353 pages) vol). Amsterdam ; John Benjamins Publishing Company (Benjamins translation library, volume 141). Available at: <http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=1824700> (Accessed: 2 November 2021).
- Díaz Cintas, J. (2003) *Teoría Y Práctica De La Subtitulación Inglés-Español*. Barcelona: Ariel.
- Díaz Cintas, J. (2004) 'Subtitling: The Long Journey To Academic Acknowledgment', *The Journal Of Specialised Translation*, 1, pp. 50–67. Available at: https://jostrans.org/issue01/art_diaz_cintas.php (Accessed: 31 December 2021).
- Díaz Cintas, J. and Remael, A. (2007) *Audiovisual Translation: Subtitling*. Manchester, UK ; Kinderhook, NY: St Jerome Publishing.
- Díaz Cintas, J. and Remael, A. (2020) *Subtitling: Concepts And Practices*. London: Routledge. doi:10.4324/9781315674278.

- Díaz Cintas, J., Remael, A. and Orero, P. (2007) *Media For All: Subtitling For The Deaf, Audio Description And Sign Language*. Amsterdam: Rodopi.
- Dumouchel, P., Boulianne, G. and Brousseau, J. (2011) 'Measures For Quality Of Closed Captioning', in Şerban, A., Matamala, A., and Lavour, J.-M. (eds.) *Audiovisual Translation In Close-Up: Practical And Theoretical Approaches*. Bern: Peter Lang, pp. 161–172.
- EASIT (2019) *Intellectual Output 2 - Final Report EASIT (Easy Access For Social Inclusion Training)*, pp. 1–78.
- EASIT (2020) *Intellectual Output 4 - Final Report EASIT (Easy Access For Social Inclusion Training)*. Available at: <https://core.ac.uk/download/pdf/323197193.pdf>.
- EastEnders* (2020). BBC One.1 June 2020, 20:10.
- Edinburgh Television Festival (2021) *MacTaggart Lecture: Jack Thorne | w/ Captions & BSL | Edinburgh TV Festival 2021*. Available at: <https://www.youtube.com/watch?v=TaxwlpbJbbg> (Accessed: 14 December 2021).
- Eichmeyer, D. (2017) 'Speech-To-Text Interpreting: Barrier-Free Access To Universities For The Hearing Impaired', *Proceedings Of The 1st Swiss Conference On Barrier-Free Communication*, Winterthur: ZHAW, pp. 6–15.
- Eichmeyer, D. and Kurch, A. (2017) 'Live Subtitling And Speech-To-Text-Interpreting: Differentiation And Quality Assessment.' *Conference on Barrier-Free Communication*, 15-16 September 2017.
- Eichmeyer-Hell, D. (2020) *WIRA: Model For Qualitative Assessment Of A Speech-To-Text Interpreting Service, Taking Into Account The User Perspective*. Extract from PhD Thesis.
- Eichmeyer-Hell, D. (2021) 'Speech Recognition (Respeaking) Vs. The Conventional Method (Keyboard): A Quality-Oriented Comparison Of Speech-To-Text Interpreting Techniques And Addressee Preferences', *Proceedings Of The 3rd Swis Conference On Barrier-Free Communication (BfC 2020)*, Winterthur (online), 29th June-4th July 2020, pp. 209–215.

- Elcessor, E. (2015) 'Blurred Lines: Accessibility, Disability, And Definitional Limitations | First Monday', *First Monday*, 20(9). Available at: <https://firstmonday.org/ojs/index.php/fm/article/view/6169> (Accessed: 15 November 2021).
- Elliott, J. (1991) *Action Research For Educational Change*. Buckingham: Open University Press.
- Elliott, J. (2007) *Reflecting Where the Action Is: The Selected Works of John Elliott*. Abingdon: Routledge.
- Emiliani, P. (2009) 'Perspectives On Accessibility: From Assistive Technologies To Universal Access And Design For All.', in Stephanidis, C. (ed.) *The Universal Access Handbook*. Boca Raton, FL: CRC Press, pp. 2.1-2.18.
- Eugeni, C. (2006) 'Introduzione Al Rispeakeraggio Televisivo', in *TRAlinea. Online Translation Journal*, Special Issues: Respeaking. Available at: https://www.intralea.org/index.php/specials/article/Introduzione_al_rispeakeraaggio_televisivo (Accessed: 17 November 2021).
- Eugeni, C. (2008) 'A Sociolinguistic Approach To Real-Time Subtitling: Respeaking Vs. Shadowing And Simultaneous Interpreting', *English In International Deaf Communication*, (72), pp. 357–382.
- Eugeni, C. (2009) 'Respeaking The BBC News: A Strategic Analysis Of Respeaking On The BBC', *The Sign Language Translator And Interpreter*, 3(1), pp. 29–68.
- Eugeni, C. (2015) 'Teaching Live Subtitling For The Deaf Through Respeaking - The Potential Of Blind Respeakers And Paraplegic Live Editors'. *5th International Symposium on Live Subtitling and Accessibility*, Rome, 12 June.
- Eugeni, C. (2017) 'La Sottotitolazione Intralinguistica Automatica: Valutare La Qualità Con IRA', *Come Studi Di Comunicazione E Medializzazione Linguistiche E Culturale*, 2(1), pp. 102–113.
- Eugeni, C. (2020a) 'Human-Computer Interaction In Diamesic Translation. Multilingual Live Subtitling', in Dejica, D., Eugeni, C., and Dejica-Carțiș, A. (eds) *Translation Studies And Information Technology - New Pathways For Researchers*,

- Teachers And Professionals. Translation Studies Series.* Politehnica University Timișoara: Editura Politehnica, pp. 19–31.
- Eugeni, C. (2020b) 'Respeaking : Aspects Techniques, Professionnels Et Linguistiques Du Sous-Titrage En Direct', *ESSACHESS. Journal For Communication Studies*, 13(1 (25)), pp. 21–35.
- EUR-LEX (2019) *Directive (EU) 2019/882 Of The European Parliament And Of The Council Of 17 April 2019 On The Accessibility Requirements For Products And Services (Text with EEA relevance)*, 151. Available at: <http://data.europa.eu/eli/dir/2019/882/oj/eng> (Accessed: 21 November 2021).
- Farey-Jones, D. (2019) *Amazon Alexa Helps Blind Woman In RNIB-Approved Ad.* Available at: https://www.campaignlive.co.uk/article/amazon-alexa-helps-blind-woman-rnib-approved-ad/1595228?utm_source=website&utm_medium=social (Accessed: 14 December 2021).
- Faye, S. (2021) *The Trans-Gender Issue An Argument For Justice.* UK: Penguin.
- Federal Communications Commission (no date) *ECFS Filings Results.* Available at: [https://www.fcc.gov/ecfs/search/filings?limit=25&proceedings_name=RM-11848&q=\(proceedings.name:\(\(05%5C-231*\)\)%20OR%20proceedings.description:\(\(05%5C-231*\)\)\)&sort=date_disseminated,DESC](https://www.fcc.gov/ecfs/search/filings?limit=25&proceedings_name=RM-11848&q=(proceedings.name:((05%5C-231*))%20OR%20proceedings.description:((05%5C-231*)))&sort=date_disseminated,DESC) (Accessed: 21 December 2021).
- Festival of the Spoken Nerd (2014) *Festival Of The Spoken Nerd Technobabble.* Available at: <https://www.youtube.com/watch?v=3paxngyySwo&feature=youtu.be&t=2m9s>.
- Filmer, D. (2019) 'Voicing Diversity? Negotiating Italian Identity Through Voice-Over Translation In BBC Broadcasting', *Perspectives*, 27(2), pp. 299–315. doi:10.1080/0907676X.2018.1449871.
- Foddy, W. (1993) *Constructing Questions For Interviews And Questionnaires Theory And Practice In Social Research.* UK: Cambridge University Press.

- Fresno, N. and Romero-Fresco, P. (2021) 'Automatic Live Subtitling In English: Taking The Bad With The Good'. *14th International Languages and The Media Conference*, Online, 20 September.
- Fresno, N. and Sepielak, K. (2020) 'Subtitling Speed In Media Accessibility Research: Some Methodological Considerations', *Perspectives* [Preprint]. doi:<https://doi.org/10.1080/0907676X.2020.1761841>.
- Fricker, M. (2007) *Epistemic Injustice - Power And The Ethics Of Knowing*. Oxford: Oxford University Press.
- Fryer, L. (2015) 'Missing Data Analysis'. University College London.
- Fryer, L. (2016) *An Introduction To Audio Description*. London and New York: Routledge.
- Fryer, L. and Cavallo, A. (2021) *Integrated Access In Live Performance*. London and New York: Routledge.
- Gallant, N. (2012) 'Ni Gallant: The Craze For Pimping Your Hearing Aids!', *The Limping Chicken*, 11 June. Available at: <https://limpingchicken.com/2012/06/11/ni-gallant-everyones-pimping-their-hearing-aids/> (Accessed: 23 July 2020).
- Gerber, E. (2007) 'Seeing Isn't Believing: Blindness, Race And Cultural Literacy', *The Senses And Society*, 2(1), pp. 27–40.
- Getz, D. (2007) *Event Studies Theory, Research And Policy For Planned Events*. Amsterdam, Boston, Heidelberg, London, New York, Oxford, Paris, San Diego, San Francisco, Singapore, Sydney, Toronto: Elsevier.
- Gillieron, R. and Robson, C. (eds.) (2015) *Counterculture UK - A Celebration*. Twickenham: Supernova Books.
- Given, L. (ed.) (2012) 'Action Research', in *The SAGE Encyclopaedia Of Qualitative Research Methods*. SAGE Publications, Inc. Available at: <https://dx.doi.org/10.4135/9781412963909> (Accessed: 5 January 2020).
- Goffman, E. (1963) *Stigma - Notes On The Management Of Spoiled Identity*. London: Penguin.
- Goodley, D. (2014) *Dis/ability Studies Theorising Disablism And Ableism*. London and New York: Routledge.

- goodreads (no date) *A Quote By Xun Kuang*. Available at: <https://www.goodreads.com/quotes/7565817-tell-me-and-i-forget-teach-me-and-i-may> (Accessed: 6 December 2021).
- Gottlieb, H. (2012) 'Subtitles - Readable Dialogue?', in Perego E. (ed.) *Eye Tracking In Audiovisual Translation*. Aracne, pp. 37–81.
- Gough, H. (2020) 'Qualitative Vs Quantitative Research: What Is It And When Should You Use It?'. Available at: <https://www.qualtrics.com/blog/qualitative-research/> (Accessed: 20 April 2020).
- Government of Canada (2019) *Broadcasting Regulatory Policy CRTC 2019-308 English-Language Closed Captioning Mandatory Quality Standards Relating To The Accuracy Rate For Live Television Programming*. Available at: <https://crtc.gc.ca/eng/archive/2019/2019-308.htm> (Accessed: 21 December 2021).
- Graeae Theatre Company (2016) *A Guide For Journalists On The Use Of Language*. Available at: <https://graeae.org/wp-content/uploads/2016/10/Graeae-Media-Language-Guide.pdf> (Accessed: 31 December 2021).
- Greco, G.M., Pedone, L., Monsellato, E., Rizzo, B. & Spinelli, E. (2012) *Guida Per Eventi Accessibili Ed Ecosostenibili' [Guidelines For Accessible And Ecosustainable Events]*. Social Sound, Puglia Sounds. Apulia Regional Government, Italy.
- Greco, G.M. (2013) 'The Normative Dimension On Cultural Accessibility.' *First Italian Workshop On Cultural Accessibility, Lecce, Italy*.
- Greco, G.M. (2016) 'On Accessibility As A Human Right, With An Application To Media Accessibility', in Matamala, A. and Orero, P. (eds.) *Researching Audio Description: New Approaches*. London: Palgrave Macmillan UK, pp. 11–33. Available at: <https://ebookcentral.proquest.com/lib/roehampton-ebooks/reader.action?docID=4719968#> (Accessed: 11 November 2021).
- Greco, G.M. (2018) 'The Nature Of Accessibility Studies', *Journal Of Audiovisual Translation*, 1(1), pp. 205–232. doi:10.47476/jat.v1i1.51.

- Greco, G.M. (2019a) 'Accessibility Studies: Abuses, Misuses And The Method Of Poietic Design', in Stephanidis, C. (ed.) *HCI International 2019 – Late Breaking Papers*. Cham: Springer International Publishing (Lecture Notes in Computer Science), pp. 15–27. doi:10.1007/978-3-030-30033-3_2.
- Greco, G.M. (2019b) 'Towards A Pedagogy Of Accessibility. The Need For Critical Learning Spaces', *Linguistica Antverpiensia, New Series: Themes In Translation Studies* (18), pp. 23–46.
- Greco, G.M. and Jankowska, A. (2019) 'Framing Media Accessibility Quality', *Journal Of Audiovisual Translation*, 2(2), pp. 1–10. doi:10.47476/jat.v2i2.114.
- Greco, G.M. and Jankowska, A. (2020) 'Media Accessibility Within And Beyond Audiovisual Translation', in *The Palgrave Handbook Of Audiovisual Translation And Media Accessibility*. Springer International Publishing, pp. 57–81.
- Greco, G.M. and Moores, Z. (2021) 'Investigating A Multidimensional Framework For Subtitling Quality'. *Media For All 9 Sketching Tomorrow's Mediascape*, Barcelona, 27 January.
- Griffiths, M. (2012) 'Action Research For/As/Mindful Of Social Justice', in *The SAGE Handbook Of Action Research*. SAGE Publications Ltd. Available at: <https://dx.doi.org/10.4135/9780857021021>.
- Guffey, E. (2018) *Designing Disability Symbols, Space, And Society*. London: Bloomsbury.
- Hall, S. and du Gay, P. (eds.) (1996) *Questions Of Cultural Identity*. London ; Thousand Oaks; New Delhi: SAGE Publications.
- Hammond, L., Spence, R. and Simpson, L. (2018) *Still No Idea*. London: Oberon Books Ltd.
- Hamraie, A. (2017) *Building Access Universal Design And The Politics Of Disability*. Minneapolis, London: University of Minnesota Press.
- Hamscha, S. (2016) 'Crip Humor', in *Gender - Laughter*. Macmillan Interdisciplinary Handbooks, pp. 349–362.

- Harris, D. (2001) 'From The Kennedy Commission To The Combahee Collective: Black Feminist Organizing, 1960-80', in Collier-Thomas, B. and Franklin, V.P. (eds.) *Sisters In The Struggle - African American Women In The Civil Rights-Black Power Movement*. New York and London: New York University Press.
- Hatim, B. (2001) *Teaching And Researching Translation*. Harlow: Pearson Education Limited.
- Health and Safety Executive (no date) *Protect Homeworkers, The Health And Safety Toolbox: How To Control Risks At Work*. Available at: <https://www.hse.gov.uk/toolbox/workers/home.htm> (Accessed: 10 November 2021).
- Hewison, R. (2020) 'How The UK Government Failed To Save The Arts', *Artreview*, 9 September. Available at: <https://artreview.com/how-the-uk-government-failed-to-save-the-arts/> (Accessed: 18 November 2021).
- Hill, A. (2018) 'Media Audiences And Reception Studies', in Di Giovanni, E. and Gambier, Y. (eds.) *Reception Studies And Audiovisual Translation*, pp. 3–19.
- Ho, A. (2004) 'To Be Labelled, Or Not To Be Labelled: That Is The Question', *British Journal Of Learning Disabilities*, 32, pp. 86–92.
- Holliday, A. (2007) *Doing And Writing Qualitative Research*. Second. London: SAGE.
- Hughes, C. (2012) 'Seeing Blindness In Children's Picture Books', *Journal Of Literary & Cultural Disability Studies*, 6(1), pp. 35–51.
- Hunt, E. and Guardian Readers (2017) "'I feel Like A Second-Class Citizen": Readers On Navigating Cities With A Disability', *The Guardian*, 22 September. Available at: <https://www.theguardian.com/cities/2017/sep/22/second-class-citizen-readers-navigating-cities-disability> (Accessed: 27 July 2020).
- Hutchinson, D. (2001) 'Identity Crisis: Intersectionality, Multidimensionality And The Development Of An Adequate Theory Of Subordination', *Michigan Journal Of Race And Law*, 6(2), pp. 285–318.
- Ibbetson, R. (2019) *Barbie Launches Disabled Dolls In Wheelchairs And With Prosthetic Legs*, *Mail Online*. Available at: <https://www.dailymail.co.uk/news/article-428>

6695287/Barbie-launches-disabled-dolls-wheelchairs-prosthetic-legs-Mattel-expands-range.html (Accessed: 27 July 2020).

IFHOH and Blau, A. (2021) 'Position Paper Hearing Accessibility Guidelines'. Available at:

https://www.ifhoh.org/_files/ugd/4e728a_37a8c04995604f5291aa300dc4459a66.pdf (Accessed: 3 December 2021).

Imrie, R. (2012) 'Universalism, Universal Design And Equitable Access To The Built Environment', *Disability And Rehabilitation*, 34(10), pp. 873–882. doi:10.3109/09638288.2011.624250.

Inclusion London (2015) *Factsheet The Social Model of Disability*. Available at: <https://www.inclusionlondon.org.uk/disability-in-london/social-model/the-social-model-of-disability-and-the-cultural-model-of-deafness/> (Accessed: 23 July 2020).

International Symbol of Access (2012). Available at: https://en.wikipedia.org/wiki/International_Symbol_of_Access#/media/File:MUT_CD_D9-6.svg (Accessed: 11 November 2021).

ITC (1999) 'ITC Guidance On Standards For Subtitling'. Available at: https://webarchive.nationalarchives.gov.uk/ukgwa/20090208182649/http://www.ofcom.org.uk/static/archive/itc/itc_publications/codes_guidance/standards_for_subtitling/index.asp.html (Accessed: 31 December 2021).

Jankowska, A. (2021) 'Audio Describing Films: A First Look Into The Description Process', *Journal Of Specialised Translation*, 36a, pp. 26–52. Available at: https://jostrans.org/issue36/art_jankowska.php (Accessed: 31 December 2021).

JCWI (2017) *Windrush Scandal Explained*, *Joint Council For The Welfare Of Immigrants*. Available at: <https://www.jcwi.org.uk/windrush-scandal-explained> (Accessed: 14 November 2021).

Jivanda, T. (2014) 'Chinese New Year: BBC Subtitle Blunder Reads 'Welcome To The Year Of'', *The Independent*, 3 February. Available at: <https://www.independent.co.uk/news/media/tv-radio/chinese-new-year-bbc->

- blunder-welcome-to-the-year-of-the-whores-9102117.html (Accessed: 18 November 2021).
- Jones, M. (2002) 'Deafness As Culture: A Psychosocial Perspective', *Disability Studies Quarterly*, 22(2), pp. 51–60. doi:10.18061/dsq.v22i2.344.
- Jones, O. (2020) 'Coronavirus Is Not Some Great Leveller: It Is Exacerbating Inequality Right Now', *The Guardian*, 9 April. Available at: <https://www.theguardian.com/commentisfree/2020/apr/09/coronavirus-inequality-managers-zoom-cleaners-offices> (Accessed: 9 November 2021).
- Jordan, S. (2012) 'Participatory Action Research', in Given L. (ed.) *The SAGE Encyclopaedia Of Qualitative Research Methods*. SAGE Publications, Inc. Available at: <https://dx.doi.org/10.4135/9781412963909> (Accessed: 5 January 2020).
- Joyce, R. and Xu, X. (2019) *Inequalities In The Twenty-First Century*. Institute for Fiscal Studies / Nuffield Foundation.
- Juang, B.H. and Rabiner, L.R. (2005) *Automatic Speech Recognition - A Brief History Of The Technology Development*. Santa Barbara, 1 p.67: Georgia Institute of Technology. Atlanta Rutgers University and the University of California. Available at: https://folk.idi.ntnu.no/gamback/teaching/TDT4275/literature/juang_rabiner04.pdf (Accessed: 17 November 2021).
- Kahneman, D. (2011) *Thinking, Fast And Slow*. London: Penguin.
- Katan, D. (2009) 'Translation As Intercultural Communication', in Munday, J. (ed.) *The Routledge Companion To Translation Studies*. London: Routledge, pp. 74–92.
- KDH Research & Communication (2012) *1. How Hearing Works, The Cochlear Implant School Toolkit*. Available at: <http://cischooltoolkit.com/q/teachers/a-preparing-for-school-entry-1/1---how-hearing-works/familiar-sounds-audiogram.php> (Accessed: 22 November 2021).
- Kelly, D. (2005) *A Handbook For Translator Trainers*. Manchester, UK & Northampton, MA: St Jerome Publishing.

- Kiraly, D. (2000) *A Social Constructivist Approach To Translator Education – Empowerment From Theory To Practice*. Manchester: St Jerome Publishing.
- Kiversal (2019) 'Prelingual And Postlingual Hearing Loss: Communication Problems', *Blog Of Kiversal*, 14 February. Available at: <https://blog.kiversal.com/en/prelingual-hearing-loss/> (Accessed: 21 November 2021).
- Kleege, G. (2016) 'Audio Description Described', *Representations* 135 (1), 89-101.
- Kleege, G. (2017) *More Than Meets The Eye: What Blindness Brings To Art*. Oxford: Oxford University Press.
- Kleege, G. and Wallin, S. (2015) 'Audio Description As A Pedagogical Tool', *Disability Studies Quarterly* 35(2).
- Koshy, E., Koshy, V. and Waterman, H. (2010) *Action Research In Healthcare*. SAGE. Available at: <https://sk.sagepub.com/books/action-research-in-healthcare> (Accessed: 15 November 2021).
- Kusters, A., De Meulder, M. and O'Brien, D. (2017) 'Innovations In Deaf Studies: Critically Mapping The Field', in Kusters, A., De Meulder, M. and O'Brien, D. (eds.) *Innovations In Deaf Studies: The Role Of Deaf Scholars*. Oxford: Oxford University Press, pp. 1–53.
- Kvale, S. and Brinkmann, S. (2009) *Interviews: Learning The Craft Of Qualitative Research Interviewing*. 2nd ed. Los Angeles: Sage Publications.
- Ladd, P. (2003) *Understanding Deaf Culture - In Search of Deafhood*. Clevedon: Multilingual Matters Ltd.
- Lambourne, A. (2006) 'Subtitle Respeaking - A New Skill For A New Age', in *TRAlinea. online Translation Journal*, Special Issue: Respeaking. Available at: http://www.intralea.org/specials/article/Subtitle_respeaking (Accessed: 17 November 2021).
- Laukka, P., Linnman, C., Åhs, F., Pissiota, A., Frans, Ö., Faria, V., Michelgård, Å., Appel, L., Fredrikson, M. & Furmark, T. (2008) 'In A Nervous Voice: Acoustic Analysis

- And Perception Of Anxiety In Social Phobics' Speech', *Journal Of Nonverbal Behavior*, 32(4), pp. 195–214. doi:10.1007/s10919-008-0055-9.
- Lave, J. and Wenger, E. (1991) *Situated Learning: Legitimate Peripheral Participation*. Cambridge: Cambridge University Press.
- Lawson, A. and Beckett, A.E. (2020) 'The Social And Human Rights Models Of Disability: Towards A Complementarity Thesis', *The International Journal Of Human Rights*, pp. 1–32. doi:10.1080/13642987.2020.1783533.
- Lazard, C. (2019) *Accessibility In The Arts: A Promise And A Practice*. Available at: <https://promiseandpractice.art/> (Accessed: 26 December 2021).
- LegalDictionary Content Team (2016) 'Distributive Justice - Definition, Examples, Cases, Processes', *Legal Dictionary*, 16 July. Available at: <https://legaldictionary.net/distributive-justice/> (Accessed: 11 November 2021).
- Leitch, R. and Day, C. (2000) 'Action Research And Reflective Practice: Towards A Holistic View', *Educational Action Research*, 8(1), pp. 179–193.
- Lewin, K. (1946) 'Action Research And Minority Problems', in *Resolving Social Conflicts*. New York: Harper & Row.
- Lewis, H. (2017) 'Deaf Liberation Theology And Social Justice', *Religions*, (8), p. 232. doi:<https://dx.doi.org/10.3390/rel8100232>.
- Link, B.G. and Phelan, J.C. (2001) 'Conceptualizing Stigma', *Annual Review Of Sociology*, 27(1), pp. 363–385. doi:10.1146/annurev.soc.27.1.363.
- Lomax, H. (2020) 'Eastenders Aired "Ground-Breaking" Episode To Reflect Character's Hearing Loss', *British Deaf News*, 5 June. Available at: <https://www.britishdeafnews.co.uk/eastenders-hearing-loss/> (Accessed: 11 November 2021).
- Louise, S. (2017) '10 Ways To Be A Good Ally To Disabled People', *she might be*, 25 June. Available at: <https://shemightbe.co.uk/10-ways-good-ally-disabled-people/> (Accessed: 31 January 2020).

- lumencandela (no date) *Intersectionality | Introduction To Women Gender Sexuality Studies*. Available at: <https://courses.lumenlearning.com/suny-introwgss/chapter/intersectionality/> (Accessed: 23 July 2020).
- Maastricht University (2019) *The Inclusive Equality And The Human Rights Model Of Disability - Interview With Prof. Degener*. Available at: <https://www.youtube.com/watch?v=9ILQOUEAOW8> (Accessed: 15 November 2021).
- Mace, R. (1988) *Universal Design: Housing For The Lifespan Of All People*. The Center For Universal Design: Nort Carolina State University.
- Mackewn, J. (2008) 'Facilitation As Action Research In The Moment', in Reason, P. and Bradbury, H. (eds) *The SAGE Handbook Of Action Research Participative Inquiry And Practice*. London: SAGE Publications, pp. 615–628.
- Mann, B. (2014) *Equity And Equality Are Not Equal, The Education Trust*. Available at: <https://edtrust.org/the-equity-line/equity-and-equality-are-not-equal/> (Accessed: 10 April 2019).
- Manville, B. (1980) 'Solon's Law Of Stasis And Atimia In Archaic Athens', *Transactions Of The American Philological Association (1974-)*, 110, p. 213. doi:10.2307/284218.
- Marais, K. (2014) *Translation Theory And Development Studies A Complexity Theory Approach*. New York and London: Routledge Taylor & Francis Group.
- Marek (2017) *Of Native Speakers And Other Fantastic Beasts - TEFL Equity Advocates And Academy*. Available at: <https://teflequityadvocates.com/2017/05/13/of-native-speakers-and-other-fantastic-beasts/> (Accessed: 19 December 2021).
- Marrow, A.J. (1972) 'The Effects Of Participation On Performance', in Marrow, A.J. (ed.) *The Failure Of Success*. New York: Amacom, pp.90-102.
- Martí Ferriol, J.L. (2013) 'Subtitle Reading Speeds In Different Languages : The Case Of Lethal Weapon', *Quaderns. Revista de Traducció*, (20), pp. 201–210.

- Martínez, J.P., Romero-Fresco, P. and Linder, G. (2010) 'The Reception Of A New Display Mode In Live Subtitling'. *8th International Languages & The Media Conference*, Berlin, 6 October.
- Maslow, A. (1954) *Motivation And Personality*. New York: Harper & Row. Available at: http://pages.ucsd.edu/~nchristenfeld/Happiness_Readings_files/Class%207%20-%20Maslow%201954.pdf (Accessed: 11 September 2017).
- Max (2013) 'The Story Of The Wheelchair Barbie: Why Are There So Few Disabled Dolls? - Update Someone's Thought Of Everything!', *Blue Badge Style*, 30 December. Available at: <https://bluebadgestyle.com/2013/12/story-wheelchair-becky-barbies-disabled-dolls/> (Accessed: 23 July 2020).
- McCall, L. (2005) 'The Complexity Of Intersectionality', *Signs*, 30(3), pp. 1771–1800. doi:10.1086/426800.
- McCulloch, A. (2020) 'Landmark Ruling: Ethical Veganism Is A Philosophical Belief', *Personnel Today*, 3 January. Available at: <https://www.personneltoday.com/hr/tribunal-rules-that-ethical-veganism-is-a-protected-characteristic/> (Accessed: 27 July 2020).
- McDonald, C. (2009) 'The Importance Of Identity In Policy: The Case For And Of Children', *Children & Society*, 23(4), pp. 241–251.
- McIntyre, D., Moores, Z. and Price, H. (2018) *Respeaking Parliament: Using Insights From Linguistics To Improve The Speed And Quality Of Live Parliamentary Subtitles*. Huddersfield: Language Unlocked/Institute for Applied Linguistics.
- McNamara, J. and Kimyoncu, C. (2016) *Blue Pen*. Vital Xposure.
- McNiff, J. (2013) *Action Research Principles And Practice*. Third edn. London and New York: Routledge Taylor & Francis Group.
- McPolin, S. (2019) 'Woman "Not Disabled Enough" For Parking Space', *Kent Online*, 4 October. Available at: <https://www.kentonline.co.uk/gravesend/news/woman-not-disabled-enough-for-parking-space-213505/> (Accessed: 9 November 2021).
- Medcalf, L. (2019) 'Echo Show Accessibility Features', *Assistive Technology At Easter Seals Crossroads*, 6 February. Available at:

<https://www.eastersealstech.com/2019/02/06/echo-show-accessibility/>

(Accessed: 14 December 2021).

Meding, J. von, Forino, G., Gaillard, J.C. & Chmutina, K. (2017) 'Grenfell Tower Fire Exposes The Injustice Of Disasters', *The Conversation*, 20 June. Available at: <http://theconversation.com/grenfell-tower-fire-exposes-the-injustice-of-disasters-79666> (Accessed: 14 November 2021).

Meizel, K. (2011) 'A Powerful Voice: Investigating Vocality And Identity: Voice And Speech Review: Vol 7, No 1', *Voice and Speech Review*, 7. doi:<https://doi.org/10.1080/23268263.2011.10739551>.

Mele, A. (2018) 'Surtitling And The Audience: A Love-Hate Relationship', *Journal Of Specialised Translation*, 30, pp. 181–202. Available at: https://jostrans.org/issue30/art_mele.php (Accessed: 31 December 2021).

Merriam-Webster.com (no date) 'Identity politics', *Merriam-Webster*. Available at: <https://www.merriam-webster.com/dictionary/identity%20politics> (Accessed: 11 November 2021).

Merrick, A. (2015) *Designing For Disability*, *The New Yorker*. Available at: <https://www.newyorker.com/business/currency/designing-for-disability> (Accessed: 26 December 2021).

Meyer, J. (2000) 'Using Qualitative Methods In Health Related Action Research', *British Medical Journal*, 320, pp. 178–181.

Middleton, P. and Spinney, J. (2016) *Notes On Blindness*.

Miquel-Iriarte, M., Vilaró, A., Orero, P., Serrano, J. & Delgado, H. (2012) 'Entitling - A Way Forward For Accessibility', in Perego, E. (ed.) *Eye tracking in audiovisual translation*. Aracne, pp.259-276.

Mitterer, H. and McQueen, J.M. (2009) 'Foreign Subtitles Help But Native-Language Subtitles Harm Foreign Speech Perception', *PLoS ONE*. Edited by D.G. Pelli, 4(11), p. e7785. doi:10.1371/journal.pone.0007785.

- Mogensen, L. and Mason, J. (2015) 'The Meaning Of A Label For Teenagers Negotiating Identity: Experiences With Autism Spectrum Disorder', *Sociology Of Health & Illness*, 37(2), pp. 255–269. doi:10.1111/1467-9566.12208.
- Montero Perez, M., Van Den Noortgate, W. and Desmet, P. (2013) 'Captioned Video For L2 Listening And Vocabulary Learning: A Meta-Analysis', *System*, 41(3), pp. 720–739.
- Moore, Z. (2014) *Respeaking The Weather Accurately - Why The Nation's Favourite Topic Is A Respeaker's Nightmare*. Masters Dissertation. The University of Roehampton.
- Moore, Z. (2020a) 'Fostering Access For All Through Respeaking At Live Events', *Journal Of Specialised Translation*, 33, pp. 176–211. Available at: https://www.jostrans.org/issue33/art_moore.php (Accessed: 31 December 2021).
- Moore, Z. (2020b) 'Measuring The Quality Of Intralingual Respeaking At Live Events'. *7th Live Subtitling and Accessibility Symposium*, Universitat Autònoma de Barcelona, 5 November.
- Moore, Z. and Romero-Fresco, P. (2015) 'The Language Of Respeaking'. *5th International Symposium on Live Subtitling and Accessibility*, Rome, 12 June.
- Moore, Z. and Romero-Fresco, P. (2019) 'NER Assessment Flowchart'.
- Morash, K. (2017) *Bringing Europe To London: Legal Aliens' Poker Face*, *TheatreArtLife*. Available at: <https://www.theatreartlife.com/lifestyle/bringing-europe-london-legalaliens-poker-face/> (Accessed: 9 November 2021).
- Morgan, K.P. (1996) 'Describing The Emperor's New Clothes: Three Myths Of Educational (In-) Equity', in Diller A., Houston, B., Morgan, K.P. and Ayim, M. (eds.) *The Gender Question In Education: Theory, Pedagogy, & Politics*. Boulder, CO: Westview Press, pp. 105–22.
- Moser-Mercer, B. (1996) 'Quality In Interpreting: Some Methodological Issues', *The Interpreters' Newsletter*, 7, pp. 43–55.

- Moser-Mercer, B., Künzli, A. and Korac, M. (1998) 'Prolonged Turns In Interpreting: Effects On Quality, Physiological And Psychological Stress (Pilot study)', *Interpreting*, 3(1), pp. 47–64. doi:10.1075/intp.3.1.03mos.
- Munday, J. (2012) *Introducing Translation Studies - Theories And Applications*. 3rd edn. London and New York: Routledge Taylor & Francis Group.
- MyClearText (2021) 'The Difference Between Our Captions And Autocaptions Is... The Difference Between Talking To A Human Or Talking To A Machine'. Available at: <https://twitter.com/mycleartext/status/1470313806672928777>.
- National Theatre (2018) *Smart Caption Glasses*, *National Theatre*. Available at: <https://www.nationaltheatre.org.uk/your-visit/access/caption-glasses> (Accessed: 5 December 2021).
- Nertrial.com (no date) *Home Page - NER Trial*. Available at: <http://www.nertrial.com/> (Accessed: 20 December 2021).
- Nesta (2014) *Nesta Invests In Education Technology Set To Help Young People With Autism And Learning Difficulties*, *Nesta*. Available at: <https://www.nesta.org.uk/press-release/nesta-invests-in-education-technology-set-to-help-young-people-with-autism-and-learning-difficulties/> (Accessed: 26 December 2021).
- Neves, J. (2005) *Audiovisual Translation: Subtitling For The Deaf And Hard Of Hearing*. Unpublished doctoral dissertation. The University of Roehampton.
- Neves, J. (2012) 'Inclusion And Accessibility In Action Research Unit. When Research Actively Works Towards Change', *Design For All Institute Of India*, 1–7(8), pp. 169–183 .
- Neves, J. (2016) 'Action Research So Much To Account For', *Target*, 28(2), pp. 237–247.
- NHS (2017) *Hearing Loss*, *nhs.uk*. Available at: <https://www.nhs.uk/conditions/hearing-loss/> (Accessed: 21 November 2021).
- Nord, C. (1997) *Translating As A Purposeful Activity: Functionalist Approaches Explained*. Manchester: St Jerome Publishing.

- Nuance Communications (no date) *Nuance Online Store - Dragon Professional Individual, v15 Upgrade (from DPI 14)*, Nuance. Available at: https://shop.nuance.co.uk/store/nuanceeu/en_GB/pd/ThemeID.4857678000/productID.5063426100 (Accessed: 18 November 2021).
- Nunnally, J.O. (1978) *Psychometric Theory*. New York: McGraw-Hill.
- O'Brien, R. (2001) 'Um Exame Da Abordagem Metodológica Da Pesquisa Ação [An Overview Of The Methodological Approach Of Action Research]', in Richardson, R. (ed.) *Teoria E Prática Da Pesquisa Ação [Theory And Practice Of Action Research]*. João Pessoa, Brazil: Universidade Federal da Paraíba. Available at: <http://www.web.ca/~robrien/papers/arfinal.html> (Accessed: 20 January 2020).
- O'Brien, S. (2006) 'Eye-Tracking And Translation Memory Matches', *Perspectives: Studies on Translatology*, 14(3), pp. 185–205.
- O'Dell, L. (2021) 'Liam O'Dell: #Whereistheinterpreter – What Happens Next?', *The Limping Chicken*, 28 July. Available at: <https://limpingchicken.com/2021/07/28/liam-odell-whereistheinterpreter-what-happens-next/> (Accessed: 10 November 2021).
- Ofcom (no date) *The Quality Of Live Subtitling*, Ofcom. Available at: <https://www.ofcom.org.uk/consultations-and-statements/category-1/subtitling> (Accessed: 21 November 2021).
- Ofcom (2005) *Subtitling - An Issue Of Speed?* Available at: https://www.ofcom.org.uk/__data/assets/pdf_file/0018/16119/subt.pdf (Accessed: 24 November 2021).
- Ofcom (2013) *Measuring The Quality Of Live Subtitling Statement*. Available at: https://www.ofcom.org.uk/__data/assets/pdf_file/0017/51731/qos-statement.pdf (Accessed: 21 November 2021).
- Ofcom (2014a) *Measuring Live Subtitling Quality Results From The First Sampling Exercise*. Available at: https://www.ofcom.org.uk/__data/assets/pdf_file/0019/45136/sampling-report.pdf (Accessed: 24 November 2021).

- Ofcom (2014b) *Measuring Live Subtitling Quality Results From The Second Sampling Exercise*. Available at: https://www.ofcom.org.uk/__data/assets/pdf_file/0024/47625/qos_second_report.pdf (Accessed: 21 November 2021).
- Ofcom (2015a) *Measuring Live Subtitling Quality Results From The Fourth Sampling Exercise*. Available at: https://www.ofcom.org.uk/__data/assets/pdf_file/0011/41114/qos_4th_report.pdf (Accessed: 21 November 2021).
- Ofcom (2015b) *Measuring Live Subtitling Quality Results From The Third Sampling Exercise*. Available at: https://www.ofcom.org.uk/__data/assets/pdf_file/0022/40774/qos_3rd_report.pdf (Accessed: 21 November 2021).
- Ofcom (2015c) *Ofcom's Code On Television Access Services*. Available at: https://www.ofcom.org.uk/__data/assets/pdf_file/0016/40273/tv-access-services-2015.pdf (Accessed: 22 November 2021).
- Ofcom (2021a) *Ofcom's Code On Television Access Services*. Available at: https://www.ofcom.org.uk/__data/assets/pdf_file/0035/179954/tv-access-services-code.pdf (Accessed: 21 November 2021).
- Ofcom (2021b) *Update On Channel 4 Subtitles, Ofcom*. Available at: <https://www.ofcom.org.uk/news-centre/2021/update-on-channel-4-subtitles> (Accessed: 26 December 2021).
- O'Leary, Z. (2004) *The Essential Guide To Doing Research*. London: SAGE.
- Oncins, E., Eugeni, C. and Bernabé, R. (2019) 'The Future Of Mediators For Live Events: LTA Project - Academic And Vocational Training', *Cultus*, 12, pp. 129–153.
- O'Neill, D., McDonald, D. and Jones, S. (2018) 'Toying With Inclusivity', *BMJ*, p. k5193. doi:10.1136/bmj.k5193.
- O'Neill, T. (2016) 'Featured Interview - Tilly O'Neill', *Journal Of Specialised Translation*, 25. Available at: https://www.jostrans.org/issue25/int_oneill.php (Accessed: 15 December 2021).

- Opotow, S. (2016) 'Social Justice Theory And Practice: Fostering Inclusion In Exclusionary Contexts', in Hammack Jr., P.L. (ed.) *The Oxford Handbook Of Social Psychology And Social Justice*. Available at: 10.1093/oxfordhb/9780199938735.013.2 (Accessed: 23 July 2020).
- Orcutt, L. (2014) *Knowbrainer Review Of Naturallyspeaking Ver.13*. KnowBrainer. Available at: <https://www.knowbrainer.com/index.cfm/wiki/dragon-naturallyspeaking-13-review/> (Accessed: 1 June 2020).
- Orcutt, L. (2016) *Knowbrainer Review Of DPI (Dragon Pro Ver. 15)*. KnowBrainer. Available at: <https://www.knowbrainer.com/index.cfm/wiki/dragon-15-review/> (Accessed: 18 November 2021).
- Ozanne, J.L. and Anderson, L. (2010) 'Community Action Research', *Journal Of Public Policy & Marketing*, 29(1).
- Paas, F., Tuovinen, J.E. and van Gerven Tabbers, H. (2003) 'Cognitive Load Measurement As A Means To Advance Cognitive Load Theory', *Educational Psychologist*, 38(1), pp. 63–71.
- Pallant, J. (2016) *SPSS Survival Manual*. 6th edn. England: Open University Press.
- Pearce, N. and Paxton, W. (eds.) (2005) *Social Justice - Building A Fairer Britain*. Great Britain: ippr.
- Penguin (2021) *The Transgender Issue*. Available at: <https://www.penguin.co.uk/books/315349/the-transgender-issue/9780241423141> (Accessed: 14 December 2021).
- Penney, D. (2001) 'Equality, Equity And Inclusion In Physical Education And School Sport', in Laker, A. (ed.) *Sociology Of Sport And Physical Education: An Introduction*. ProQuest Ebook Central: Taylor & Francis Group, p. 19. Available at: <https://ebookcentral.proquest.com/lib/roehampton-ebooks/detail.action?docID=168578>. (Accessed: 7 May 2020).
- Perego, E., Del Missier, F., Porta, M. & Mosconi, M. (2010) 'The Cognitive Effectiveness Of Subtitle Processing', *Media Psychology*, 13(3), pp. 243–272. doi:10.1080/15213269.2010.502873.

- Pérez, J.M. (2012) 'Applying A Punctuation-Based Segmentation To A New Add-On Display Mode Of Respoken Subtitles', in Remael, A., Orero, P., and Carroll, M. (eds.) *Audiovisual Translation And Media Accessibility At The Crossroads*. Brill | Rodopi, pp. 133–146. doi:10.1163/9789401207812_009.
- Petherick, S. (2021) 'Yo Sushi Worker Refuses To Let Guide Dog In Until Owner Demands To Speak To Boss', *Metro*, 22 September. Available at: <https://metro.co.uk/2021/09/22/yo-sushi-worker-refuses-to-let-guide-dog-in-until-owner-demands-to-speak-to-boss-15299975/> (Accessed: 14 December 2021).
- Piller, I. and Takahashi, K. (2010) 'At The Intersection Of Gender, Language, And Transnationalism', in Coupland, N. (ed.) *The Handbook of Language and Globalization*. Oxford, UK: Wiley-Blackwell, pp. 540–554. doi:10.1002/9781444324068.ch24.
- Planet Dharma (no date) *Zen Teachings, Planet Dharma*. Available at: <https://www.planetdharma.com/zen-teachings/> (Accessed: 16 December 2021).
- Pöchhacker, F. (2020) '103: Profile Definition And Competences Of The Professional Ilser - Present And Future Of Interlingual Live Subtitling: Research, Training And Practice'. *Free ILSA Multiplier Event*, Antwerp, 7 June. Available at: <http://ka2-ilsa.webs.uvigo.es/dissemination/> (Accessed: 17 November 2021).
- Pražák, A., Loose, Z., Psutka, J.V., Radová, V. & Psutka, J. (2020) 'Live TV Subtitling Through Respeaking With Remote Cutting-Edge Technology', *Multimedia Tools And Applications*, (79), pp. 1203–1220.
- Price, K. (1983) 'Closed-Captioned TV: An Untapped Resource', *MATESOL Newsletter*, (1–2), pp. 1–8.
- Prime Minister's Office, 10 Downing Street and The Rt Hon Boris Johnson MP (2020) *Prime Minister's Statement On Coronavirus (COVID-19): 23 March 2020*, GOV.UK. Available at: <https://www.gov.uk/government/speeches/pm-address-to-the-nation-on-coronavirus-23-march-2020> (Accessed: 2 November 2021).

- Pring, J. (2020) *Coronavirus: BSL-Users Launch Mass Legal Action Over Government 'Discrimination'*, *Disability News Service*. Available at: <https://www.disabilitynewsservice.com/coronavirus-bsl-users-launch-mass-legal-action-over-government-discrimination/> (Accessed: 23 July 2020).
- Puchta, C. and Potter, J. (2004) *Focus Group Practice*. London: Sage.
- Pulver, A. (2020) 'BFI Diversity Standards "Failing To Tackle Race Discrimination In UK Film Industry"', *The Guardian*, 15 July. Available at: <https://www.theguardian.com/film/2020/jul/15/bfi-diversity-standards-failing-to-tackle-race-discrimination-in-uk-film-industry> (Accessed: 1 March 2021).
- QA (no date) *Managing, Leading & Personal Effectiveness Practice*. Available at: <https://www.qa.com/training/courses/leadership/> (Accessed: 15 December 2021).
- Quick Fix Theatre (2016) *Untitled*. Available at: <https://www.facebook.com/quickfixtheatre/photos/1826208864289074> (Accessed: 20 December 2021).
- Quick Fix Theatre (2017) *Untitled*. Available at: <https://www.facebook.com/quickfixtheatre/photos/1938570636386229>.
- Quinn, G., Degener, T., Bruce, A., Burke, C., Castellino, J., Kenna, P., Kilkelly, U. & Quinlivan, S. (2002) *Human Rights And Disability: The Current Use And Future Potential Of United Nations Human Rights Instruments In The Context Of Disability*. New York and Geneva: United Nations.
- Rajendran, D.J., Duchowski, A.T., Orero, P., Martínez, J. and Romero-Fresco, P. (2013) 'Effects Of Text Chunking On Subtitling: A Quantitative And Qualitative Examination', *Perspectives*, 21(1), pp. 5–21. doi:10.1080/0907676X.2012.722651.
- Rakić, T., Steffens, M.C. and Mummendey, A. (2011) 'Blinded By The Accent! The Minor Role Of Looks In Ethnic Categorization', *Journal Of Personality And Social Psychology*, 100(1), pp. 16–29. doi:10.1037/a0021522.

- Rand, E. (2014) *Reclaiming Queer: Activist And Academic Rhetorics Of Resistance*. 1st edn. Tuscaloosa: University Alabama Press.
- Rawls, J. (1971/1999) *A Theory Of Justice Revised Edition*. Cambridge, MA: Belknap Press of Harvard University Press.
- Rawls, J. (2004) 'On Justice As Fairness', in *Social Justice*. Oxford: Blackwell Publishing.
- Reason, P. and Bradbury, H. (eds) (2008) *The SAGE Handbook Of Action Research Participative Inquiry And Practice*. London: SAGE Publications.
- Red Bee Media (2019) *Live Automatic Captioning For The ABC*. Available at: <https://www.youtube.com/watch?v=klxML9KR3rU> (Accessed: 15 December 2021).
- Remael, A., Orero, P., Black, S. & Jankowska, A. (2019) 'From Translators To Accessibility Managers: How Did We Get There And How Do We Train Them?', *MonTI*, 11, pp. 131–154.
- Remael, A., Van Waes, L. and Leijtenet, M. (2016) 'Live Subtitling With Speech Recognition - How To Pinpoint The Challenges?', in *Media And Translation. An Interdisciplinary Approach*. London: Bloomsbury, pp. 120–147.
- Risku, H. (2016) 'Situated Learning In Translation Research Training: Academic Research As A Reflection Of Practice', *The Interpreter And Translator Trainer*, 10(1), pp. 12–28. doi:10.1080/1750399X.2016.1154340.
- RNID (no date) *Facts And Figures, RNID*. Available at: <https://rnid.org.uk/about-us/research-and-policy/facts-and-figures/> (Accessed: 21 November 2021).
- RNID (2018) *People With Deafness And Hearing Loss Still Feel Stigmatised, Shows Charity Poll, RNID*. Available at: <https://rnid.org.uk/about-us/media-centre/latest-press-releases/people-with-deafness-and-hearing-loss-still-feel-stigmatised-shows-charity-poll/> (Accessed: 11 November 2021).
- Roach, P. (2009) *English Phonetics And Phonology A Practical Course*. Fourth edn. UK: Cambridge University Press.

- Robert, I.S., Schrijver, I. and Diels, E. (2019) 'Live Subtitlers: Who Are They? A Survey Study', *Linguistica Antverpiensia, New Series – Themes In Translation Studies*, (18), pp. 101–129.
- Romero-Fresco, P. (2009) 'More Haste Less Speed: Edited Versus Verbatim Respoken Subtitles', *VIAL-VIGO International Journal Of Applied Linguistics*, (6), pp. 109–133.
- Romero-Fresco, P. (2010) 'Standing On Quicksand: Hearing Viewers' Comprehension And Reading Patterns Of Respoken Subtitles For The News', in Díaz Cintas, J., Matamala, A., and Neves, J. (eds.) *New Insights Into Audiovisual Translation And Media Accessibility*. Brill | Rodopi, pp. 175–194. doi:10.1163/9789042031814_014.
- Romero-Fresco, P. (2011) *Subtitling Through Speech Recognition: Respeaking*. Manchester, UK; Kinderhook, NY: St. Jerome Pub (Translation practices explained).
- Romero-Fresco, P. (2013) 'Accessible Filmmaking: Joining The Dots Between Audiovisual Translation, Accessibility And Filmmaking', *Journal Of Specialised Translation*, 20, pp. 201–223.
- Romero-Fresco, P. (2015) *The Reception Of Subtitles For The Deaf And Hard Of Hearing In Europe*. 1st edition. Bern, Berlin, Bruxelles, Frankfurt am Main, New York, Oxford & Wien: Peter Lang.
- Romero-Fresco, P. (2016) 'Accessing Communication: The Quality Of Live Subtitles In The UK', *Language And Communication*, 9, pp. 56–69.
- Romero-Fresco, P. (2018) 'Respeaking Subtitling Through Speech Recognition', in Pérez-González, L. (ed.) *The Routledge Handbook Of Audiovisual Translation*. London: Routledge, pp. 96–113.
- Romero-Fresco, P. (2019) *Accessible Filmmaking*. London and New York: Routledge.
- Romero-Fresco, P., Melchor-Couto, S., Dawson, H., Moores, Z., & Pedregosa, I. (2019) 'Respeaking Certification Bringing Together Training, Research And Practice',

- Linguistica Antverpiensia, New Series – Themes In Translation Studies*, 18.
doi:10.52034/lanstts.v18i0.514.
- Romero-Fresco, P. (2020) 'Negotiating Quality Assessment In Media Accessibility: The Case Of Live Subtitling', *Universal Access In The Information Society*.
doi:10.1007/s10209-020-00735-6.
- Romero-Fresco, P. (2021a) 'Creative Media Accessibility: Placing The Focus Back On The Individual', *Lecture Notes In Computer Science*.
- Romero-Fresco, P. (2021b) 'Creativity In Media Accessibility: A Political Issue', *Cultus*, 14 *Special Issue on Translation plus: The added value of the translator*.
- Romero-Fresco, P. and Alonso-Bacigalupe, L. (2022) 'An Empirical Analysis On Quality In Interlingual Live Subtitling: Horses For Courses', *XLinguae, Special Issue: Innovation in audiovisual translation: from production to reception*.
- Romero-Fresco, P. and Chaume, F. (2022) 'Creativity In Audiovisual Translation And Media Accessibility', *Journal Of Specialised Translation*, 38, pp.75-101. Available at: https://www.jostrans.org/issue38/art_romero.php (Accessed: 20 March 2022).
- Romero-Fresco, P. and Eugeni, C. (2020) 'Live Subtitling Through Respeaking', in Bogucki, Ł. and Deckert, M. (eds) *The Palgrave Handbook Of Audiovisual Translation And Media Accessibility*. Springer International Publishing, pp. 269–295.
- Romero-Fresco, P. and Fryer, L. (2016) 'The Reception Of Automatic Theatre Captions - Preferences, Perception And Engagement'. *Unlimited! International Symposium on Accessible Live Events*, Antwerp, 29 April.
- Romero-Fresco, P. and Martínez, J. (2015) 'Accuracy Rate In Live Subtitling: The NER Model', in *Audiovisual Translation In A Global Context: Mapping An Ever-Changing Landscape*. London & New York: Palgrave Macmillan, pp. 28–50.
- Romero-Fresco, P. and Pöchhacker, F. (2018) 'Quality Assessment In Interlingual Live Subtitling: The NTR Model', *Linguistica Antverpiensia, New Series – Themes In Translation Studies*, 16, pp. 149–167.
doi:<https://doi.org/10.52034/lanstts.v16i0.438>.

- Rosa, A. (2020) *Top Tips From A Top Crip - Amy Rosa*. Available at: <https://amyrosaliveart.com/Top-Tips-from-a-Top-Crip> (Accessed: 14 November 2021).
- Ryan, F. (2020a) 'Britain Has A Hidden Coronavirus Crisis – And It's Shaped By Inequality', *The Guardian*, 15 April. Available at: <https://www.theguardian.com/commentisfree/2020/apr/15/britain-coronavirus-crisis-inequality-hungry-domestic-violence> (Accessed: 27 March 2020).
- Ryan, F. (2020b) 'Chronically Ill People Are The Experts In This Strange New World', *The Irish Times*, 3 June. Available at: <https://www.irishtimes.com/life-and-style/health-family/chronically-ill-people-are-the-experts-in-this-strange-new-world-1.4268681> (Accessed: 14 November 2021).
- Ryan, F. (2020c) 'Covid Lockdown Opening Up World For People With Disabilities', *The Guardian*, 20 April. Available at: <https://www.theguardian.com/world/2020/apr/20/covid-lockdown-opening-up-world-for-people-with-disabilities> (Accessed: 14 November 2021).
- SAGE (no date) 'Focus Groups', *Methods Map*. Available at: <https://methods.sagepub.com/methods-map/focus-groups> (Accessed: 1 May 2020).
- Saldhana, G. and O'Brien, S. (2013) *Research Methodologies In Translation Studies*. Manchester: St Jerome Publishing.
- Sanchez, D. (2014) 'Reading Speed', *Red Bee Media*, 8 August. Available at: <https://www.redbeemedia.com/blog/reading-speed/> (Accessed: 23 November 2021).
- Sancho-Aldridge (1996) *Good News For Deaf People: Subtitling Of National News Programmes An ITC Research Publication*. ITC.
- Sandrelli, A. (2013) 'Reduction Strategies And Accuracy Rate In Live Subtitling Of Weather Forecasts: A Case Study'. *4th International Symposium on Live Subtitling. Live Subtitling with Respeaking and Other Respeaking Applications*, Barcelona, 12 March.

- Sandrelli, A. (2019) 'Training Respeakers: A Hybrid Profession For The (Near) Future?', *CULTUS*, (12), pp. 154–167.
- Sandrelli, A. (2020) 'Interlingual Respeaking And Simultaneous Interpreting In A Conference Setting: A Comparison', *inTRAlinea. online Translation Journal* [Preprint], (Special Issue: Technology In Interpreter Education And Practice). Available at: <https://www.intralinea.org/index.php/specials/article/2518> (Accessed: 23 November 2021).
- Saner, E. (2021) 'Strictly Exclusive! Winner Rose Ayling-Ellis On The Glitterball, Giovanni And The Joy Of Being Deaf', *The Guardian*, 19 December. Available at: <https://www.theguardian.com/lifeandstyle/2021/dec/19/strictly-exclusive-winner-rose-ayling-ellis-on-the-glitter-ball-giovanni-and-the-joy-of-being-deaf> (Accessed: 19 December 2021).
- Sapolsky, R.M. (2004) *Why Zebras Don't Get Ulcers*. Owl Books Edition. Available at: <https://books.google.co.uk/books?id=laRsnQAACAAJ>.
- SARAHSMITH76 (2013) 'Pimp My Hearing Aids And Cochlear Implants', *Pimp My Hearing Aids And Cochlear Implants*, 24 October. Available at: <https://pimpmyhearingaids.wordpress.com/> (Accessed: 27 July 2020).
- Saussure, F. de (1916) *Cours de linguistique générale*. Paris: Editions Payot.
- Schechner, R. (1988) *Performance Theory*. London: Routledge.
- Schultz, G. (2019) 'Plastic Straw Bans Are Not Fair To People With Disabilities, And Here's What We Can Do About It', *CreakyJoints*, 28 October. Available at: <https://creakyjoints.org/advocacy/plastic-straw-bans-bad-for-people-with-disabilities/> (Accessed: 23 July 2020).
- Schuman, A.J. (2014) 'Speed EHR Documentation With Voice Recognition Software: Incorporating Voice Recognition Technology Into Your EHR Will Help You Spend Less Time Documenting And More Time With Patients', *Contemporary Pediatrics*, 31(6), pp. 28–32.

- Sense (2021) *Deafblindness | Sense, For People With Complex Disabilities Or Who Are Deafblind*. Available at: <https://www.sense.org.uk/get-support/information-and-advice/conditions/what-is-deafblindness/> (Accessed: 21 November 2021).
- Shakespeare, T. (2013) 'The Social Model Of Disability', in Davis, L.J. (ed.) *The Disability Studies Reader*. Florence: Taylor & Francis Group, pp. 214–221. Available at: ProQuest Ebook Central (Accessed: 15 November 2021).
- Shannon, L. (2021) 'Subnormal: A British Scandal'. Available at: <https://www.bbc.co.uk/programmes/m000w81h>.
- Shape Arts (no date) *Social Model Of Disability, Shape Arts*. Available at: https://www.shapearts.org.uk/news/social-model-of-disability?gclid=Cj0KCQiA-K2MBhC-ARIsAMtLKRsqieFGrGF64OvdfMOtTzIGETPSxQN4zvw5qwPaMsMEI4K_Me9rDJkaAsDTEALw_wcB (Accessed: 10 November 2021).
- Sharma, S. (2021) 'Cinema Praised For "Going The Extra Mile" To Make Deaf People Feel Welcome', *ChronicleLive*, 12 December. Available at: <https://www.chroniclive.co.uk/news/north-east-news/whitley-bay-jam-jar-deaf-22403672> (Accessed: 16 December 2021).
- Shockey, L. (2003) *Sound Patterns Of Spoken English*. UK: Blackwell Publishing.
- SignHealth (no date) 'What Is The Difference Between Deaf And Deaf?', *SignHealth*. Available at: <https://signhealth.org.uk/resources/learn-about-deafness/deaf-or-deaf/> (Accessed: 3 December 2021).
- Silent Witness* (2013-2020) BBC One.
- Silverstein, J. (2021) 'The Global Impact Of George Floyd: How Black Lives Matter Protests Shaped Movements Around The World', *CBS News*, 4 June. Available at: <https://www.cbsnews.com/news/george-floyd-black-lives-matter-impact/> (Accessed: 14 December 2021).
- Sokoli, S. and Pedersen, J. (2021) 'Twitter Discussion On Creativity In AVT And MA'. Available at:

<https://mobile.twitter.com/JanPedersen1969/status/1366990128728727555> and
<https://mobile.twitter.com/StavroulaSokoli/status/1366750557768196098>.

Specsavers (2021) *Hearing Loss | Specsavers UK*. Available at:
<https://www.specsavers.co.uk/hearing/hearing-loss> (Accessed: 21 November 2021).

Stagetext (2011) *Captioning In Theatre*. Available at:
<https://www.youtube.com/watch?app=desktop&v=iflI2OYn7Zw> (Accessed 16 December 2021).

Stagetext (2013) *Getting More Theatre For Your Bucks*. Available at:
<https://www.youtube.com/watch?v=nMzm1yvtwKU> (Accessed: 16 December 2021).

Stagetext (2017a) 'Digital Subtitling Guidelines'. Available at:
https://www.stagetext.org/wp-content/uploads/2021/09/Digital_subtitling_Guidelines.pdf.

Stagetext (2017b) 'Stagetext Marketing Pack'. Available at:
<https://www.brightonfringe.org/files/00f3728dbab7867688e2d7e8d0906cda.pdf>
(Accessed: 22 November 2021).

Stagetext (2017c) 'Stagetext Provides Equal Access To The Arts With Captions And Subtitles'. Available at: https://stagetext.org/wp-content/uploads/2021/06/resources_st_Infographic_online.pdf (Accessed: 22 November 2021).

Stagetext (2018) 'Stagetext Captioning Style Guide - Theatre1'.

Stagetext (2019) *Trustee's Report And Accounts For The Year Ended 31 March 2019*. Available at: <https://register-of-charities.charitycommission.gov.uk/> (Accessed: 9 November 2021).

- Stagetext (2021a) *Live Subtitling For Talks And Tours*, Stagertext. Available at: <https://www.stagetext.org/for-venues/live-subtitling-for-talks-and-tours/> (Accessed: 15 December 2021).
- Stagetext (2021b) *Theatre Captioning*, Stagertext. Available at: <https://www.stagetext.org/for-venues/theatre-captioning/> (Accessed: 5 December 2021).
- Stagetext (2021c) *What's The Difference Between Captions And Subtitles?*, Stagertext. Available at: <https://www.stagetext.org/faqs/whats-the-difference-between-captions-and-subtitles/> (Accessed: 17 November 2021).
- Stiker, H.-J. (1999) *A History Of Disability*. USA: University of Michigan Press.
- Stinson, M. (2015) 'Speech-To-Text Interpreting', *Routledge Encyclopedia Of Interpreting Studies*. London and New York: Routledge Taylor & Francis Group.
- Strictly Come Dancing* (2021) BBC One.
- Szarkowska, A. (2011) 'Text-To-Speech Audio Description: Towards Wider Availability Of AD', (15), pp. 142–162.
- Szarkowska, A., Krejtz, I., Klyszejko, Z. & 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., Krejtz, K., Dutka, Ł. & Pilipczuk, O. (2016) 'Cognitive Load In Intralingual And Interlingual Respeaking', *Poznań Studies In Contemporary Linguistics*, 52(2), pp. 209–233. doi:10.1515/psicl-2016-0008.
- Szarkowska, A., Dutka, Ł., Pilipczuk, O. & Krejtz, K. (2017) 'Respeaking Crisis Points. An Exploratory Study Into Critical Moments In The Respeaking Process', in Deckert, M. (ed.) *Audiovisual Translation - Research and Use*. Bern: Peter Lang, pp. 179–201.
- Szarkowska, A., Krejtz, K., Dutka, Ł. & Pilipczuk, O. (2018) 'Are Interpreters Better Respeakers?', *The Interpreter And Translator Trainer*, 12(2), pp. 207–226. doi:10.1080/1750399X.2018.1465679.

- Szarkowska, A. (2018) *Report On The Results Of The SURE Project Study On Subtitle Speeds And Segmentation*. Available at: https://avt.ils.uw.edu.pl/files/2018/07/SURE_Project_Report.pdf (Accessed: 24 November 2021).
- Szarkowska, A. (2020) 'Research - What Can We Do Better In Media Accessibility Research? Q&A'. *LEAD-ME Winter Training School*, Online, 24 November.
- Szarkowska, A. and Gerber-Morón, O. (2018) 'Viewers Can Keep Up With Fast Subtitles: Evidence From Eye Movements', *PLOS ONE*. Edited by D. Orrego-Carmona, 13(6), p. e0199331. doi:10.1371/journal.pone.0199331.
- Szczygielska, M., Dutka, Ł., Szarkowska, A., Romero-Fresco, P., Pöchhacker, F., Tampir, M., Figiel, W., Moores, Z., Robert, I., Schrijver, I. & Haverhals, V. (2020) 'How To Implement Speech-To-Text Interpreting (Live Subtitling) In Live Events. Guidelines On Making Live Events Accessible'. ILSA Project. Available at: <http://ka2-ilsa.webs.uvigo.es/guidelines>.
- Szczygielska, M. and Dutka, Ł. (2017) 'How To Implement Live Subtitling On TV In 3 Days? A Case Study From Poland'. *Intermedia International Conference on Audiovisual Translation*, Poznan, 25 September.
- Tai, E. (2009) 'A Case For Distributive Justice In Design', in Brody, D. (ed.) *Design Studies A Reader*. London and Oxford: Berg Publishers.
- Tatchell, P. (2021) *Peter Tatchell | Human Rights, Democracy, Global Justice, LGBTI Freedom*. Available at: <https://www.petertatchell.net/> (Accessed: 14 December 2021).
- Taylor-Coleman, J. (2020) 'Coronavirus: Call For Clear Face Masks To Be "The Norm"', *BBC News*, 26 May. Available at: <https://www.bbc.com/news/world-52764355> (Accessed: 27 July 2020).
- Teixera, C.S.C. and O'Brien, S. (2017) 'Investigating The Cognitive Ergonomic Aspects Of Translation Tools In A Workplace Setting', *Translation Spaces*, 6(1), pp. 79–103.

- Text on Tap (no date) *Text On Tap - Live Subtitles Meets Live Events*. Available at: <https://text-on-tap.live/> (Accessed 31 March 2022).
- The Communications Act (2003) *Communications Act 2003*. Statute Law Database. Available at: <https://www.legislation.gov.uk/ukpga/2003/21/contents> (Accessed: 21 November 2021).
- The Economist* (2020) Autumn Is Coming: Technology Quarterly, 13 June.
- The Equality Act 2010 (2010) *Equality Act 2010, c15*. Statute Law Database. Available at: <https://www.legislation.gov.uk/ukpga/2010/15/contents> (Accessed: 10 November 2021).
- Thompson, I. (2020) *Guest Post: 'Let Us Look Into The UK's Equality And Discrimination Laws In The Age Of Black Lives Matter' By Ife Thompson, Equal Rights Trust*. Available at: <https://www.equalrightstrust.org/news/guest-post-let-us-look-uk%E2%80%99s-equality-and-discrimination-laws-age-black-lives-matter-ife-thompson> (Accessed: 14 November 2021).
- #ToyLikeMe (no date a) *Our Story*, #toylikeme. Available at: <https://www.toylikeme.org/about-us/our-story/> (Accessed: 23 July 2020).
- #ToyLikeMe (no date b) *Why Representation Matters*, #toylikeme. Available at: <https://www.toylikeme.org/about-us/why-representation-matters/> (Accessed: 23 July 2020).
- Trochim, W.M.K. (2021) *Descriptive Statistics*. Available at: <https://conjointly.com/kb/descriptive-statistics/> (Accessed: 16 December 2021).
- TUC (2017) *Personal protective equipment and women Guidance for workplace representatives on ensuring it is a safe fit*.
- Tymoczko, M. (2010) *Translation, Resistance, Activism. Essays On The Role Of Translators As Agents Of Change*. Amherst and Boston, MA: University of Massachusetts Press.

- UN General Assembly (1948) *Universal Declaration Of Human Rights, United Nations, 217 (III) A*. Available at: <https://www.un.org/en/about-us/universal-declaration-of-human-rights> (Accessed: 15 November 2021).
- United Nations (2012) *Human Rights Indicators: A Guide To Measurement And Implementation*. United Nations. doi:10.18356/58576336-en.
- United Nations (2016) 'Good Practices Of Accessible Urban Development - Making Urban Environments Inclusive And Fully Accessible To ALL'. Department of Economic and Social Affairs. Available at: https://www.un.org/disabilities/documents/desa/good_practices_in_accessible_urban_development_october2016.pdf.
- UPIAS (1975) *Fundamental Principles*. Available at: <https://disability-studies.leeds.ac.uk/wp-content/uploads/sites/40/library/finkelstein-UPIAS-Principles-2.pdf> (Accessed: 15 November 2021).
- Üstün, T.B., Chatterji, S., Bickenbach, J.E., Trotter, I.I. & RT, S.S. (2001) 'Disability And Cultural Variation: The ICIDH-2 Cross-Cultural Applicability Research Study'. *Disability And Culture: Universalism And Diversity*. Seattle: Hogrefe & Huber Publishers.
- Van Waes, L., Leijten, M. and Remael, A. (2013) 'Live Subtitling With Speech Recognition. Causes And Consequences Of Text Reduction', *Across Languages And Cultures*, 14(1), pp. 15–46.
- Vanderplank, R. (1988) 'The Value Of Teletext Sub-Titles In Language Learning', *ELT Journal*, 42(4), pp. 272–281. doi:10.1093/elt/42.4.272.
- 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 (New Language Learning & Teaching Environments).
- van der Veer, B. (2007) 'De Tolk Als Respeaker: Een Kwestie Van Training', *Linguistica Antverpiensia, New Series – Themes In Translation Studies*, (6). doi:<https://doi.org/10.52034/lanstts.v6i.194>.

- de Vries, K.M. (2015) 'Transgender People Of Color At The Center: Conceptualizing A New Intersectional Model', *Ethnicities*, 15(1), pp. 3–27. doi:10.1177/1468796814547058.
- Wallhagen, M.I. (2010) 'The Stigma Of Hearing Loss', *The Gerontologist*, 50(1), pp. 66–75. doi:10.1093/geront/gnp107.
- #We Shall Not Be Removed, Ramps On The Moon, Attitude Is Everything, Paraorchestra & Friends and What Next? (2021). *Working Safely Through Covid-19 Seven Inclusive Principles For Arts & Cultural Organisations*. Available at: <https://www.whatnextculture.co.uk/wp-content/uploads/2020/09/Seven-Inclusive-Principles-for-Arts-Cultural-Organisations-working-safely-through-COVID-19-1.pdf> (Accessed: 30 November 2022).
- Weale, S. (2021) 'Funding Cuts To Go Ahead For University Arts Courses In England Despite Opposition', *The Guardian*, 20 July. Available at: <https://www.theguardian.com/education/2021/jul/20/funding-cuts-to-go-ahead-for-university-arts-courses-in-england-despite-opposition> (Accessed: 18 November 2021).
- Wellcome Collection (no date) *Medicine Man, Wellcome Collection*. Available at: <https://wellcomecollection.org/exhibitions/Weoe4SQAAKJwjcDC> (Accessed: 20 December 2021).
- Wells, J.C. (2008) *Phonetic Symbols For English, Phonetic Symbols For English, With Unicode Hex Numbers*. Available at: https://www.phon.ucl.ac.uk/home/wells/phoneticsymbolsforenglish_hex.htm (Accessed: 18 November 2021).
- Wembley Stadium (2021) *Wembley Stadium Beomes UK's Biggest Venue To Offer British Sign Language At Every Concert*. Available at: <https://www.wembleystadium.com/news/2021/Sep/16/Wembley-Stadium-commits-to-BSL-interpretation-of-all-concertshttps://www.wembleystadium.com/news/2021/Sep/16/Wembley->

- Stadium-commits-to-BSL-interpretation-of-all-concerts (Accessed: 14 December 2021).
- WHO (2021) *Deafness And Hearing Loss*. Available at: <https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss> (Accessed: 3 December 2021).
- Williams, J. and Chesterman, A. (2002) *The Map - A Beginner's Guide To Doing Research In Translation Studies*. Manchester: St Jerome Publishing.
- Wooton, T. (2020) 'Automatic For The People - Captioning And Automation'. *7th Symposium on Live Subtitling and Accessibility*, Online, 4 December. Available at: <https://www.youtube.com/watch?v=NubQx3oSoRw> (Accessed: 10 December 2021).
- d'Ydewalle, G., Van Rensbergen, J. and Polley, J. (1987) 'Reading A Message When The Same Message Is Available Auditorily In Another Language: The Case Of Subtitling', in O'Regan, J.K. and Lévy-Schoen, A. (eds.) *Eye Movements: From Physiology to Cognition*. B.V. North Holland: Elsevier Science Publishers.
- Zimmerman, D. (2010) 'Increasing The Power Of Nonparametric Tests By Detecting And Downweighting Outliers', *J. Exp. Educ.*, 64(1), pp. 71–78.
- Zola, I.K. (1989) 'Towards The Necessary Universalizing Of A Disability Policy', *The Milbank Quarterly*, 67, *Supplement 2 Part 2*. Disability Policy: Restoring Socioeconomic Independence, pp. 401–428.

Filmography for the focus groups

Back In Time For The Weekend The 60s, Tuesday 9th February 2016, BBC Two, Kim Maddever.

Beowulf Return To The Shieldlands Episode 7, Sunday February 14th 2016, ITV, James Dormer.

Question Time, Thursday 18th February 2016, BBC One, Hayley Valentine.

Sunday Brunch, Sunday 7th February 2016, Channel 4, Presented by Tim Lovejoy and Simon Rimmer.