



The
University
Of
Sheffield.

Multisensory input video-therapy for young children with cleft palate-related speech sound disorder: Examining service delivery and interactional characteristics

By:

Samantha Calladine

A thesis submitted in partial fulfilment of the requirements for the degree of
Doctor of Philosophy

The University of Sheffield
Faculty of Medicine, Dentistry & Health
Department of Human Communication Sciences

Apr 2020

Abstract

Background

Multisensory input video-therapy (MSIVT) is an early intervention for cleft palate-related speech sound disorder. It is an input-based approach to support the development of accurate internal templates for speech sounds and create opportunities for output practice. Activities are videoed and watched by parents and child at home. Little is known about characteristic features of MSIVT.

Aims

1. Describe the delivery of MSIVT as it is implemented as an episode of care in the NHS
2. Analyse the features of therapist-child interaction during MSIVT sessions
3. Examine interaction over the course of an episode of care

Methods

Data were collected on 29 MSIVT sessions featuring five children aged 1;6-2;11 and three speech and language therapists. It included 573 minutes of video data. Descriptive methods examined the delivery of MSIVT. Therapist-child interaction was analysed with conversation analysis to examine features of therapist turns, actions they fulfilled, and the consequences for the child.

Results

Episodes of care involved up to nine monthly therapy sessions. Characteristic features include multiple speech sound stimuli and modified articulations within skilfully designed activities. Therapist-child interactions fit within four broad kinds of action: a) Demonstration to the camera; b) Invitation to attend to the stimulus; c) Invitation to participate in the stimulus routine; and d) Invitation to produce the stimulus. Therapists use creative methods to demonstrate stimuli and facilitate verbal and nonverbal engagement in activities.

Discussion and Implications

This is the first study to analyse characteristic features of MSIVT. The nature of interaction differs from output-based interventions due to the nonverbal responses that the therapist invites from the child. This new evidence extends knowledge of how speech and language therapy works as a real-time interactional process and how very young children are engaged in speech intervention. Results will inform clinical implementation, training of therapists, and the development of future research to evaluate the impact of MSIVT.

Acknowledgements

I would like to use this space to acknowledge the contributions and support of some very special people.

Dr Anne Harding-Bell, for inspiring my interest in the intervention approach at the centre of my thesis, and for all our inspiring and reflective discussions since then. Professor Joy Stackhouse, for supporting me through my first research experience in 2007-09 and the early part of my PhD. Dr Hilary Gardner, whose PhD thesis has been an inspiration to my research and clinical practice.

Professor Ray Wilkinson and Dr Sarah Spencer, my current supervisors who got me to the end! Ray, your expertise and teaching in conversation analysis has enriched the quality of my research. Sarah, for keeping me focused on my aims, providing broad perspective, and for stoking the fire in my belly that got me to submission!

My colleagues in Leeds, for your patience and encouragement, and for enabling me to step back for some big final pushes at the end! My managers, past and present, for your support to take extra leave when I needed it. Past colleagues in Trent, for believing in me, and for supporting me to complete my MSc research and get this journey started at the beginning.

Dr Blanca Schaefer, you introduced me to Eckhart Tolle's *The Power of Now* when my PhD got caught up in a difficult chapter of my life. It gave me perspective when I needed it and helped me to get back on track – thank you.

My amazing PhD peers Tanja, Helen, Kim, Rachael and Alex. Such a long and challenging journey has been made easier because of you. I will miss our coffee breaks, treats, chat and reflection, but look forward to sharing new things with you!

I would like to say thank you to all the families and therapists who have selflessly taken part in this study and in doing so made it possible. You have let me in to your therapy world and I hope I have done justice to the amazing work you do with each other. I would also like to thank Health Education England Yorkshire and Humber for funding my PhD tuition fees and making this viable.

I would like to show my gratitude to my amazing family and dear friends for their nurturing support and patience. You make me feel very lucky. Dad, I am sad you will not get to see me finish – you are always in my thoughts and I miss you dearly. Laurie, you have been my rock of support from day one. My time has been stretched thin for a long time and the selfless way you have helped me through warms my heart. I am looking forward to bringing some free time back into our lives.

Publications and Presentations

Calladine, S. (2013) Developing a methodology to evaluate a psycholinguistic-based intervention for two-year-old children with cleft-related speech sound difficulties. Poster presentation at the University of Sheffield Postgraduate Research Conference; Sheffield, UK.

Britton, L., **Calladine, S.**, Extence, H., Phippen, G. & Pinkstone, M. (2017) Inequities in speech and language therapy provision for children born with cleft palate. *Royal College of Speech and Language Therapists Bulletin*; October. (pp. 12-15).

Calladine, S. (2018) Uncovering the distinctive characteristics of multisensory input video-therapy for toddlers with cleft palate related speech impairment. Presentation at the University of Sheffield Postgraduate Research Conference; Sheffield, UK.

Calladine, S. & Vance, M. (2019) A psycholinguistic approach to therapy in very young children born with cleft palate. In Harding-Bell, A. (Ed.) *Case Studies in Cleft Palate Speech: Data Analysis and Principled Intervention*. Chichester: J&R Press.

Calladine, S., Spencer, S. & Wilkinson, R. (2019) The distinctive interactional characteristics of multisensory input video-therapy. Presentation at the Child Language Symposium; Sheffield, UK.

Calladine, S., Spencer, S. & Wilkinson, R. (2019) Doing speech sound intervention with two-year-olds: Interaction is key. Poster presentation at the Royal College of Speech and Language Therapists Conference; Nottingham, UK.

Calladine, S. (2019) Multisensory input video-therapy for young children born with cleft palate: Interaction is key. Presentation at the Cleft and Craniofacial Clinical Excellence Network Study Day; Birmingham, UK.

Disclaimer

Some of the text in Chapter 2 of this thesis has been reproduced from a book chapter co-written by the author: Calladine, S. & Vance, M. (2019) A psycholinguistic approach to therapy in very young children born with cleft palate. In A. Harding-Bell (Ed.), *Case Studies in Cleft Palate Speech: Data Analysis and Principled Intervention*. Chichester: J&R Press.

Table of Contents

Abstract	<i>i</i>
Acknowledgements	<i>iii</i>
Publications and Presentations.....	<i>iv</i>
Disclaimer	<i>iv</i>
List of Tables	<i>xii</i>
List of Figures	<i>xii</i>
List of Appendices.....	<i>xiii</i>
List of Abbreviations.....	<i>xiv</i>
Chapter 1 : Introduction.....	<i>1</i>
1.1 What is multisensory input video-therapy?.....	<i>1</i>
1.2 What is cleft palate-related speech sound disorder?.....	<i>2</i>
1.3 Aim and rationale for the study.....	<i>3</i>
1.4 Research questions and methods.....	<i>4</i>
1.5 Significance of the study.....	<i>5</i>
1.6 Personal motivation for conducting the study.....	<i>6</i>
1.7 Structure of the thesis.....	<i>6</i>
Chapter 2 : A Literature Review of Cleft Palate-Related Speech Sound Disorder and the Multisensory Input Video-Therapy Approach.....	<i>11</i>
2.1 Section 1: Cleft palate and related speech sound disorder.....	<i>11</i>
2.1.1 Prevalence of cleft palate and speech sound disorder	<i>11</i>
2.1.2 Prevalence of otitis media with effusion and hearing loss	<i>11</i>
2.1.3 Nature of speech sound disorder in children born with cleft palate.....	<i>12</i>
2.1.4 A psycholinguistic perspective on speech development	<i>13</i>
2.1.5 Development of speech processing skills	<i>13</i>
2.1.6 Cleft palate-related speech sound disorder from a psycholinguistic perspective.....	<i>16</i>
2.1.6.1 Effect of cleft palate on emerging speech processing skills	<i>16</i>
2.1.6.2 Classification of cleft palate-related speech sound disorder (CPSSD)	<i>17</i>
2.1.7 Considerations in the differential diagnosis of CPSSD.....	<i>19</i>

2.2	Section 2: Early remedial therapy for children with CPSSD.....	21
2.2.1	Naturalistic approaches	22
2.2.2	Multisensory input therapy.....	26
2.2.3	Multisensory input modelling with output	27
2.2.3.1	Aims and goals of therapy and therapy techniques	29
2.2.3.2	Stimulus design and selection.....	32
2.3	Conclusions.....	33
Chapter 3 : Pilot Study.....		35
3.1	Methods.....	35
3.1.1	Design.....	35
3.1.2	Procedure.....	35
3.1.2.1	Ethical approval and consent	35
3.1.2.2	Data collection.....	35
3.1.3	Materials	36
3.1.4	Participants	37
3.1.4.1	Therapist participant.....	37
3.1.4.2	Child participants.....	37
3.1.5	Data analysis	40
3.1.5.1	Descriptive analysis of MSIVT episodes of care.....	40
3.1.5.2	Descriptive analysis of therapist behaviours and child responses.....	41
3.1.5.3	Quantitative analysis of therapist behaviours and child responses.....	42
3.2	Results	43
3.2.1	Descriptive analysis of MSIVT episodes of care	43
3.2.1.1	Nature of target selection and stimulus design.....	43
3.2.1.2	Child stimulus productions	45
3.2.2	Descriptive analysis of therapist behaviours and child responses	46
3.2.2.1	Nature and frequency distribution of therapist behaviours.....	46
3.2.2.2	Nature and frequency distribution of child responses.....	48
3.2.3	Quantitative analysis of therapist behaviours and child responses	49
3.2.3.1	Sam-James.....	49
3.2.3.2	Sam-Holly.....	50
3.2.3.3	Methodological considerations.....	51
3.3	Discussion	52
3.3.1	Nature of target selection and stimulus design	53
3.3.2	Child stimulus productions.....	54
3.3.3	Profiling therapist behaviours and child responses in MSIVT	54

3.3.4	Associations between behaviours and responses	55
3.4	Conclusions and implications for the main study.....	56
Chapter 4 : A Literature Review of Clinical and Non-Clinical Interaction		57
4.1	Section 1: Clinical interaction in SLT.....	57
4.1.1	Introduction.....	57
4.1.2	Therapist-child interaction in SLT: An overview	57
4.1.3	Structure of the therapy session	59
4.1.4	Therapist and child roles and the therapeutic alliance	59
4.1.5	Structure and pattern of therapist-child interaction.....	61
4.1.6	Nature of therapist turns.....	63
4.1.6.1	Speech acts in phonology therapy.....	64
4.1.6.2	Turn design, action and consequence.....	65
4.1.6.3	Multimodality.....	68
4.1.7	Child-led interventions to support early language development: Nature, structure and patterns of interaction.....	70
4.1.8	Making therapy practice explicit: Lessons from studying aphasia therapy sessions.....	72
4.2	Section 2: Non-clinical interaction: Insights from CA research.....	73
4.2.1	Response mobilisation.....	73
4.2.2	Clustered actions and response relevance	75
4.2.3	Longitudinal studies of interaction	76
4.2.4	Frames and footing.....	78
4.3	Conclusions.....	78
4.4	Aims and rationale for the study.....	79
4.4.1	Research questions.....	80
Chapter 5 : Methods		83
5.1	Study design	83
5.2	Procedures	84
5.2.1	Ethical approval	84
5.2.2	Identification of eligible sites.....	84
5.2.3	Research and Development (R&D) approvals	85
5.2.4	Identification and recruitment	86
5.2.5	Informed consent	86
5.2.6	Data collection.....	87
5.3	Materials.....	89

5.3.1	Summary of the data.....	89
5.4	Participants.....	90
5.4.1	Child participants	90
5.4.1.1	Inclusion and exclusion criteria.....	90
5.4.1.2	Characteristics of child participants.....	91
5.4.2	Speech and language therapist (SLT) participants	95
5.4.2.1	Inclusion and exclusion criteria.....	95
5.4.2.2	Characteristics of SLT participants.....	95
5.5	Data analysis	96
5.5.1	Phase 1: Delivery of MSIVT in the NHS	96
5.5.1.1	Descriptive analysis using metadata.....	96
5.5.2	Phase 2: Interactional features of MSIVT.....	99
5.5.2.1	Sampling the data for analysis.....	99
5.5.2.2	Conversation analysis	100
5.5.3	Phase 3: MSIVT interactions over an episode of care	102
5.5.3.1	Sampling the data for analysis.....	103
5.5.3.2	A mixed methods approach combining CA and quantitative analysis.....	103
5.6	Summary.....	104
Chapter 6 : Delivery of MSIVT in the NHS (Phase 1).....		105
6.1	Introduction.....	105
6.2	Features of MSIVT episodes of care	105
6.2.1	Structure of MSIVT episodes of care	105
6.2.2	Nature of MSIVT activities.....	107
6.2.3	Nature of target selection and stimulus design	109
6.2.3.1	Target selection.....	111
6.2.3.2	Stimulus design.....	112
6.2.3.3	Linguistic level.....	113
6.3	Variation in the delivery of MSIVT.....	114
6.3.1	Structure of MSIVT episodes of care	114
6.3.2	Nature of MSIVT activities.....	114
6.3.3	Nature of target selection and stimulus design	115
6.3.3.1	Linguistic level.....	118
6.4	Child stimulus productions.....	119
6.5	Conclusions.....	121

Chapter 7 : Interactional Features of MSIVT (Phase 2)	123
7.1 Introduction	123
7.2 A. Demonstration to the camera	126
7.2.1 A1. Demonstration for later	126
7.3 B. Invitation to attend to the stimulus	129
7.3.1 B1. Demonstration for attention	129
7.3.2 B2. Verbal request for attention.....	134
7.4 C. Invitation to participate in the stimulus routine	138
7.4.1 C1. Gesture to elicit participation.....	138
7.4.2 C2. Verbal request for participation	141
7.5 D. Invitation to produce the stimulus	145
7.5.1 D1. Demonstration for imitation	145
7.5.2 D2. Incomplete demonstration	151
7.5.3 D3. Nose-closing to elicit production	154
7.5.4 D4. Verbal request for production	156
7.6 Conclusions	159
Chapter 8 : MSIVT Interactions over an Episode (Phase 3)	163
8.1 Introduction	163
8.2 Profile of therapist action in Video 1	163
8.3 Profile of therapist action in Video 7	165
8.4 Comparative analysis of Videos 1 and 7	167
8.4.1 Nature and relative frequencies of action	167
8.4.2 Practices for initiating and responding to the same action	168
8.4.2.1 A1. Demonstration for later.....	168
8.4.2.2 C2. Verbal request for participation.....	172
8.4.2.3 D1. Demonstration for imitation.....	176
8.4.3 Initiation-Response-Evaluation (IRE) and repair sequences.....	182
8.5 Conclusions	187
Chapter 9 : Discussion	189
9.1 Introduction	189
9.2 Pilot Study	189

9.3	Phase 1: Delivery of MSIVT in the NHS.....	189
9.3.1	Structure of MSIVT episodes of care.....	190
9.3.2	Nature of MSIVT activities.....	191
9.3.3	Nature of speech sound targets and stimuli.....	192
9.3.3.1	Target selection.....	192
9.3.3.2	Stimulus design.....	193
9.3.3.3	Linguistic level.....	193
9.3.4	Child stimulus productions.....	195
9.4	Phase 2: Interactional features of MSIVT.....	195
9.4.1	Structure of MSIVT sessions and activities.....	196
9.4.2	Therapist and child roles.....	197
9.4.3	Structure and pattern of therapist-child interaction in MSIVT.....	199
9.4.4	The stimulus routine: Layers and multimodality.....	200
9.4.5	Nature of therapist-child interaction in MSIVT.....	201
9.4.5.1	Multimodality in therapist initiations.....	201
9.4.5.2	Nature of therapist action: Design features and consequences.....	201
9.4.5.3	Nature of child responses.....	206
9.4.6	Therapist evaluation of child responses.....	207
9.5	Phase 3: MSIVT interactions over an episode of care.....	209
9.5.1	Evolution in action and practices.....	209
9.5.2	Orientation to the layered and multimodal stimulus routine.....	211
9.5.3	Insight into a therapist and young child’s speech repair practices.....	212
9.5.4	Camera as an interactional tool.....	213
Chapter 10 : Implications and Conclusions.....		215
10.1	Introduction.....	215
10.2	How do therapists deliver MSIVT as an episode of care within the NHS: what is the structure of an episode of care; what speech sounds do therapists target; and what types of activities, materials and speech sound stimuli do they use?.....	215
10.3	What are the interactional features of MSIVT, and in what ways do therapists establish a child’s attention and stimulate their awareness and production of speech sounds?.....	216
10.4	How does therapist-child interaction change over the course of an episode of care; (comparing first and last sessions)?.....	218
10.5	Implications for clinical practice and teaching.....	219

10.6	Limitations of the study.....	221
10.7	Implications for future research.....	223
10.8	Conclusions.....	225
	<i>References</i>	227
	<i>Appendices</i>	247

List of Tables

Table 3.1: Summary of the complete dataset	36
Table 3.2: Characteristics of child participants	38
Table 5.1: Participant identification names	87
Table 5.2: The complete dataset.....	90
Table 5.3: Characteristics of child participants	94
Table 5.4: The video data for Phase 2 analysis	100
Table 6.1: Structure of MSIVT episodes of care	106
Table 6.2: Types of activities in the MSIVT sessions	108
Table 6.3: Speech sound targets and stimuli	110
Table 6.4: Position of stimulus in real word level productions.....	113
Table 6.5: Sequence of target and stimuli selection in the five episodes of care.....	116
Table 6.6: Number and proportion of activities containing child stimulus production.....	119
Table 7.1: Summary of therapist actions, turn design features and the nature of subsequent turns	125
Table 8.1: Activities and symbolic representations in Video 1	164
Table 8.2: Nature and frequency of action in Video 1	164
Table 8.3: Activities and symbolic representations in Video 7	165
Table 8.4: Nature and frequency of action in Video 7	166
Table 8.5: Turn design features of C2 initiations and second pair part responses.....	173
Table 8.6: Analysis of D1. Demonstration for imitation in Video 7.....	181

List of Figures

Figure 2.1: Speech processing model by Stackhouse & Wells (1997)	14
Figure 6.1: Number of activities containing child stimulus productions.....	120
Figure 8.1: Nature and frequencies of action in Videos 1 and 7.....	167
Figure 9.1: Structure of MSIVT episodes of care	190
Figure 9.2: Structure of MSIVT sessions and activities.....	196

List of Appendices

Appendix A:	Pilot Study: Confirmation of ethics exemption	247
Appendix B:	Pilot Study: Participant Information Sheet	248
Appendix C:	Pilot Study: Parent/Carer Consent Form	251
Appendix D:	IRAS (Integrated Research Application System) application form	252
Appendix E:	Research Protocol	282
Appendix F:	WoSRES ethics approval letter	297
Appendix G:	NHS sponsor R&D approval letter	301
Appendix H:	Table showing eligible sites	305
Appendix I:	R&D approval letter Site 1	306
Appendix J:	R&D approval letter Site 2	310
Appendix K:	Letter of Access form Site 1	312
Appendix L:	Letter of Access form Site 2	315
Appendix M:	Invitation letter to therapist participants	318
Appendix N:	Therapist Information Sheet	319
Appendix O:	Invitation letter to parent/carers of child participants	323
Appendix P:	Parent/Carer Information Sheet	324
Appendix Q:	Therapist Consent Form	328
Appendix R:	Therapist Consent Form: Videos	329
Appendix S:	Parent/Carer Consent Form	330
Appendix T:	Parent/Carer Consent Form: Videos	331
Appendix U:	Child Data Collection Form	332
Appendix V:	Therapist Data Collection Form	333
Appendix W:	Transcription conventions	334

List of Abbreviations

CA	conversation analysis
CPSSD	cleft palate-related speech sound disorder
CRANE	Cleft Registry and Audit Network database
CSAG	Clinical Standards Advisory Group
CV	consonant and vowel sequence
(dyad-V#-A#)	video number-activity number
ExtIPA	Extended International Phonetic Alphabet
IPA	International Phonetic Alphabet
IRE	initiation-response-evaluation
IW	invented word
MRC	Medical Research Council
MSIM	multisensory input modelling
MSIM+O	multisensory input modelling with output
MSIVT	multisensory input video-therapy
MSIM+/-VT	multisensory input modelling with/without video-therapy
<i>n</i>	number of participants
OME	otitis media with effusion
RRE	request-response-evaluation
RW	real word
SLT	speech and language therapist
S	single sound in isolation
SS	sound sequence
V	vowel
VC	vowel and consonant sequence
WF	word final position
WI	word initial position
WM	word medial position

Chapter 1: Introduction

This chapter introduces the intervention at the focus of this study, multisensory input video-therapy, and the nature of cleft palate-related speech sound disorder. I introduce the aim of the study and explain my rationale. I introduce my research questions and the methodological approaches I used. I explain the significance of the study for clinical practice and research in the area of speech and language therapy for young children born with cleft palate, and share my personal motivation for undertaking the study. The chapter concludes with an overview of the structure of the thesis.

1.1 What is multisensory input video-therapy?

In 2000, Addenbrooke's NHS Trust and East & North Herts NHS Trust published a video produced by two speech and language therapists (SLTs), Dr Anne Harding and Alison Bryan: *The use of multi-sensory input modelling (MSIM) to stimulate speech output processing: A teaching and demonstration video*. The video described a novel input-based intervention approach based on Stackhouse and Wells' (1997) speech processing model. The therapist provides the child with frequent and salient models of innovative sounds, sound sequences and eventually assembled real words. The aim is to stimulate reflection and generation of new sounds and revision of existing sounds in order to modify the child's speech output (Harding & Bryan, 2000). As an input-based approach that does not depend on output practice, it grew favour in the UK as an early intervention for young children born with cleft palate.

Harding and Bryan (2000) describe the production of 'therapy videos' to enable continuation of input beyond the therapy session, stimulate silent and private rehearsal, and demonstrate techniques to family members. In this study, I use multisensory input video-therapy (MSIVT) to refer to the hybrid of MSIM and the production and provision of therapy videos (video-therapy; VT).

Since 2000, conference and study day meetings in and outside of the UK have hosted presentations and professional discussions of MSIM+/-VT for children born with cleft palate (Britton, 2001a; Britton, 2008; Calladine, 2009; Dive, 2001; Extence, 2017; Harding, 2001; Waldron, Roberts, Latham & Gibb, 2019). Specialist SLT services incorporate MSIM into parent babble groups, a one-off session that typically takes place when babies are six to 12 months of age (Britton, 2001b; Extence, 2017; Lane & Wren, 2018). It is also used in one to one therapy with young children from around 18 months to three years of age, up to the point when a more

traditional output-based approach is introduced (Calladine, 2010).

Multisensory input modelling with/without video-therapy is a popular approach in the UK, but it is not manualised. Descriptions exist in the form of the original video (now available in DVD format) and more recently as chapters in two textbooks, Treslove (2014) and Calladine and Vance (2019). These descriptions draw on the underlying theoretical basis of MSIM+/-VT and empirical evidence from clinical case studies and one non-experimental case study (Calladine, 2009) to describe components of the intervention. However, very little research has been carried out on MSIM+/-VT so its key components are not fully understood and there is very little evidence of its impact. As such, it is a novel intervention with only clinical level evidence (Bessell et al., 2013).

1.2 What is cleft palate-related speech sound disorder?

The anatomical features of cleft palate directly affect a baby's early sound-making experiences (Chapman & Willadsen, 2011). Studies show that at 12 months of age, babies born with cleft palate produce fewer oral sounds and more nasal and non-oral sounds than babies without cleft palate (Chapman, Hardin-Jones, Schulte & Halter, 2001; Chapman & Willadsen, 2011). They also produce fewer alveolar and palatal sounds, made at the front of the mouth, and more non-oral sounds, made in the throat or nose (Lohmander, Olsson & Flynn, 2011). These characteristics are directly related to the unrepaired structure. For many babies, effective palate surgery at six to 12 months of age alleviates these effects and facilitates typical speech development that is free of cleft palate-related speech characteristics (Britton et al., 2014; CRANE, 2018). For other babies, early and/or persisting effects will continue to impact, directly or indirectly, on speech development into early and later childhood (CRANE, 2018). If there are continuing anatomical restrictions impacting on sound-making, for example, palate surgery has not provided an adequately functioning palate mechanism, speech may have a nasal quality and speech sounds may be weak or realised as nasal sounds. Even when palate surgery has been effective, the restrictions prior to surgery may have impacted the child's language system in such a way that sound substitutions or distortions continue, for example, alveolar sounds may be produced as velar sounds. Such cleft palate-related speech characteristics can be diagnosed in children from around 18 months of age (Barrett & Extence, 2010; Bowden, Harland & Sommerlad, 1997). In this study, I use the term cleft palate-related speech sound disorder (CPSSD) to refer to the persistence of cleft palate-related speech characteristics beyond 18 months.

A cleft palate also gives rise to higher levels of otitis media with effusion, which is associated with conductive hearing loss (Bruce et al., 2015; Flynn & Lohmander, 2014). This affects early auditory

experiences and may impact on emerging speech and language development. Some of the speech characteristics that are associated with hearing loss are also associated with cleft palate and, specifically, an inadequate palate mechanism (Purdy, Kim & Harding-Bell, 2019). In the UK, national standards of care and a systematic audit process are in place to ensure specialist cleft lip and palate services provide high quality care and achieve best possible clinical outcomes (Britton et al., 2014; Ness et al., 2017; Sandy et al., 1998). Speech is one of the primary outcomes of care. Specialist SLT services are audited on their adherence to a series of process and outcome standards, including: 1) 100% of children will be offered a specialist SLT assessment by age 27 months (process); and 2) 60% of children will have normal speech by age 5;11 (outcome) (Britton et al., 2014; CRANE, 2018). Analyses of regional audit data have shown that SLTs can reliably identify cleft palate-related speech characteristics at the early SLT assessment, typically carried out at 18-24 months of age. Systematic analysis of national audit data shows that speech outcomes in the UK are improving and standards are rising (Britton et al., 2014; CRANE, 2016; 2017; 2018).

1.3 Aim and rationale for the study

The overall aim of the study was to examine the nature of MSIVT as it is implemented in the NHS for children with CPSSD aged 18 months to three years, in order to identify and describe service delivery and interactional characteristics.

There has been very little research into MSIM+/-VT. A review of the existing literature (Chapter 2) found that MSIM+/-VT has a coherent theoretical basis but the empirical evidence underpinning current descriptions is limited. Despite this, there is anecdotal preference within the UK cleft SLT community for using MSIM+/-VT to treat CPSSD in young children under three. This may reflect the demand for early intervention for children born with cleft palate as well as the unsuitability of alternative speech-based approaches for children at this age. It is important that we examine and develop evidence for MSIM+/-VT to support clinical practice and service delivery to ensure best possible use of finite NHS resources.

In 2007-2009 I carried out a preliminary study to investigate the impact of MSIM+/-VT on speech output processing in four two-year-old children with CPSSD. The findings have contributed to current descriptions of the approach (e.g. Treslove, 2014) and were recently extended by Calladine and Vance (2019) using data from clinical case studies. Current descriptions illustrate that MSIM+/-VT is a complex intervention as it has multiple interacting components and is used with a heterogeneous clinical population (Craig et al., 2008; Medical Research Council, 2000). As

such, it is important that we can define it fully in order to evaluate it. Furthermore, whilst ever there is no manual for MSIM+/-VT there is likely to be significant variation in how it is delivered from one SLT to another.

A review of existing descriptions of MSIM+/-VT suggests it is different to other speech-based approaches. As an input-based approach the focus is on the therapist rather than the child producing speech sound stimuli whilst also creating opportunities for the child to engage in output if they wish. The unusual nature of MSIM+/-VT supported a preliminary pilot study to explore methods for examining the features of therapy (Chapter 3). The findings informed a review of the literature on clinical interaction (Chapter 4).

Current descriptions of MSIM+/-VT describe it as therapist-led but with a non-directive and non-corrective style of interaction. My review of the literature on clinical interaction found that this style of interaction is more characteristic of naturalistic child-led approaches, which do not typically target speech as the primary goal of therapy. This strengthened the case for a detailed examination of the nature of MSIM+/-VT to provide more insight into how this unusual method of therapy is enacted. Of particular interest was the way therapists manage their interaction with the child in order to stimulate their awareness and production of speech sounds. Interactional studies of aphasia therapy show this is an important area of research for helping to identify key components of therapy.

1.4 Research questions and methods

I collected data from three specialist SLT services in the NHS where MSIVT had been provided as part of routine clinical care. Data consisted of therapy videos of MSIVT sessions. The pilot study used data from one of these services where I was the SLT who provided the therapy. The therapy videos featured two children with CPSSD, aged 1;8 to 2;7. The main study used data from all three services. The two external services provided 29 therapy videos totalling 573 minutes of video data. Videos featured three SLTs and five children, aged 1;6 to 2;11.

The study was made up of three phases and each phase addressed a specific research question:

Phase 1. How do therapists deliver MSIVT as an episode of care within the NHS: what is the structure of an episode of care; what speech sounds do therapists target; and what types of activities, materials and speech sound stimuli do they use?

Phase 2. What are the interactional features of MSIVT, and in what ways do therapists establish a child's attention and stimulate their awareness and production of speech

sounds?

Phase 3. How does therapist-child interaction change over the course of an episode of care; (comparing first and last sessions)?

Phase 1 used all of the data. I used a descriptive method to examine the delivery of MSIVT in the five episodes of care. Phases 2 and 3 used subsets of the data to allow more detailed analysis. In Phase 2 I drew on conversation analysis (CA) to examine the practices SLTs use during MSIVT and the consequences they have for the child (Sidnell, 2010). In Phase 3 I used an innovative approach, combining CA with a quantitative method, to examine therapist-child interaction longitudinally over the course of an episode of care.

1.5 Significance of the study

This study makes an important contribution to current research knowledge because it provides the most explicit description to date of the distinctive nature and characteristic features of this novel intervention. The findings will support consistent labelling of MSIVT, which will support meaningful conversations about the approach in clinical, supervision and teaching environments, and may lead to less variability in clinical practice.

Phase 1 describes the delivery of five MSIVT episodes of care as they were implemented in the NHS by three specialist SLTs at two cleft lip and palate centres. In respect of target selection, the findings reveal alignment between current practice and well-established principles for this clinical population. However, in respect of speech sound stimuli, there was variation between therapists in the levels they targeted, and practice did not fully align with existing descriptions.

Phases 2 and 3 describe MSIVT as an interactional process. The findings make visible the way therapists interact with the young child, establishing their attention to salient demonstrations of targeted speech stimuli, their participation in activities and their production of target stimuli. The findings suggest the characteristic stimulus routine and nature and pattern of therapist-child-camera interaction are key components of MSIVT.

The findings of this study have implications for the considerations therapists make when planning and delivering MSIVT, and will support conscious and reflective therapy practice. However, more research is now needed to explore the influence of key components on the outcomes of therapy and further expand the evidence base. Importantly, this study has a facilitative role in progressing research in this area because the findings can be used as the basis for a research protocol to evaluate MSIVT.

The study is the first known study to apply the highly detailed and systematic approach of CA to the study of therapy talk involving two-year-old children with CPSSD, and to the study of therapy talk in the context of an input-based approach. As such, the interactional insights have potential to inform the development of research and clinical practice more broadly in respect of interventions for young children with other speech, language and communication needs.

1.6 Personal motivation for conducting the study

My personal motivation for conducting this study has come from my clinical experiences as a specialist SLT in two regional cleft lip and palate centres and my early research experiences as an MSc student. I have worked with children with CPSSD since I began my professional career in 2002 and in specialist roles since 2004. Over the last 15 years I have been part of the national network of specialist SLTs who use MSIM+/-VT, challenged by the lack of evidence to support clinical practice and service provision in the area of early intervention and specifically the MSIM+/-VT approach. My own research enquiry began with my MSc research in 2007-09, which I carried out under the supervision of Professor Joy Stackhouse, and has progressed into my PhD studies.

The current study has benefited from the individual and collective expertise and experiences of six academic supervisors representing four different supervisory teams:

- 2011-2014: Professor Joy Stackhouse and Dr Tom Muskett
- 2014-2016: Dr Blanca Schaefer and Dr Hilary Gardner
- 2016-2017: Dr Sarah Spencer and Dr Hilary Gardner
- 2017-2019: Dr Sarah Spencer and Professor Ray Wilkinson

Professor Joy Stackhouse's work and expertise supported and informed the psycholinguistic considerations in this study. Dr Hilary Gardner's research experience supported and informed the focus on interaction and, since he joined my supervisory team in 2017, Professor Ray Wilkinson has guided me through the CA process. This marked the start of a steep learning curve for me personally, but the highly detailed nature of analysis supported my research questions and built upon my analytical strengths.

1.7 Structure of the thesis

Here I provide an overview of the structure of the thesis to assist navigation. The thesis is constructed of a further nine chapters.

Chapter 2: A Literature Review of Cleft Palate-Related Speech Sound Disorder and the Multisensory Input Video-Therapy Approach

In this chapter I provide background to the clinical population (children with CPSSD) and intervention (MSIVT) at the focus of this study. I draw on the literature to provide an explanation of the nature of CPSSD and the theoretical and empirical basis of MSIVT. I outline existing descriptions of the design and delivery procedure of MSIM+/VT and compare these to other approaches used to treat CPSSD in young children. In my conclusion I state the overall aim of the study and explain my rationale for investigating the nature of MSIVT.

Chapter 3: Pilot Study

This chapter describes the methods I used in the pilot study to examine my own clinical practice with two children. This involved the analysis of 30 MSIVT sessions. I present eight categories of therapist behaviour and four categories of child response generated from the analysis and describe a quantitative procedure I used to examine associations between them. I present and discuss the findings and explain how they contributed to the development of the main study.

Chapter 4: A Literature Review of Clinical and Non-Clinical Interaction

In this chapter, I review the existing literature on clinical interaction in speech and language therapy involving children. I discuss the following themes that have emerged from previous research: structure of a therapy session; roles of the therapist and child; structure and patterns of therapist-child interaction; nature of therapist turns; and multimodality as a feature of the interaction between therapist and child. Within this, I will discuss the relative contributions from two methodological approaches to the study of discourse: discourse analysis and conversation analysis (CA). Additional insights from the CA literature outside of speech and language therapy are also discussed. I conclude the chapter with a reorientation to the overall aim of the study and the rationale, and present the research questions addressed in each of the three phases.

Chapter 5: Methods

In this chapter, I provide details of the study design, the procedure, including the process of obtaining NHS ethical approval, identification and recruitment of participants, and characteristics of the three SLT and five child participants. I describe the nature of data I collected, how I collected it, and how it was sampled for analysis in each phase of the study. This leads on to a detailed description of the methods used to analyse the data in each phase and rationale for choosing them. This includes a description of the process used in Phase 1 to produce metadata

from the therapy datasets as a way of organising the data and facilitating within-session and session-by-session analysis. I explain the process used in Phases 2 and 3 to produce detailed written transcripts of therapy talk from the video recordings. I describe the CA method drawn on to analyse the data and its fundamental principles. I also refer to the literature on longitudinal studies of social interaction in my explanation of the approach used to analyse MSIVT interactions at two moments in time.

Chapter 6: Delivery of MSIVT in the NHS (Phase 1)

Chapter 6 presents the findings from within-session and session-by-session analysis of the 29 therapy datasets that represent the five MSIVT episodes of care. It gives details of the structure of episodes in relation to the number, frequency and duration of sessions within them, and the number and duration of activities that took place in the MSIVT sessions. From the analysis, I identify characteristic features of the delivery of these episodes of care, and highlight aspects of delivery where there was variation in practice.

Chapter 7: Interactional Features of MSIVT (Phase 2)

Chapter 7 presents the results of systematic analysis of 120 minutes of video recordings featuring four SLTs and four children. I use written extracts of therapy talk extracted from written transcripts of video data to illustrate different types of action displayed in therapists' turns, their design and delivery, and the consequences they had for the child and continuing trajectory of talk.

Chapter 8: MSIVT Interactions over an Episode of Care (Phase 3)

In Chapter 8, I present the findings from analysis of 2x10minutes of video recordings from the first and last sessions of one episode of care. I use charts to illustrate quantitative differences in the type and distribution of therapist action at these two moments in time, and written extracts of therapy talk, with analysis, to illustrate qualitative differences.

Chapter 9: Discussion

In this chapter, I discuss the findings from Phases 1, 2, and 3 in relation to the existing literature on CPSSD and MSIVT, and critically reflect on how features explicated by the analysis align with the theoretical basis and principles of MSIVT. I discuss how characteristic features of the MSIVT interaction compare to other therapy approaches, considering distinctive features and key components of the approach.

Chapter 10: Implications and Conclusions

In this final chapter, I outline the implications for clinical practice, teaching and further research, and specifically how the findings will inform the design of future studies. I consider the benefits and challenges of CA methods, limitations of the study, and make concluding comments.

Chapter 2: A Literature Review of Cleft Palate-Related Speech Sound Disorder and the Multisensory Input Video-Therapy Approach

This chapter is organised in two sections. In Section 1, I use the literature to explain the nature and prevalence of cleft palate-related speech sound disorder (CPSSD). In Section 2, I review the literature on early intervention for young children with CPSSD and describe the approach at the focus of this study, multisensory input video-therapy (MSIVT).

2.1 Section 1: Cleft palate and related speech sound disorder

2.1.1 Prevalence of cleft palate and speech sound disorder

Cleft lip and/or palate are among the most common congenital conditions, affecting approximately one in 700 babies every year in the UK (CRANE, 2015). Clefting occurs when there is disruption to the embryological processes involved in formation of these structures. A cleft palate affects the structure and function of the palate and velopharynx; these are important parts of the speech mechanism and the consequences for speech development can persist beyond surgical repair. Although children's surgical histories differ across studies depending on the protocols at the clinical centres, studies have consistently reported a higher prevalence of speech difficulties in children born with cleft palate than in children without cleft palate. This has been demonstrated at three (Klintö et al., 2014; Willadsen et al., 2018) and five years of age (Bercow, 2008; Britton et al., 2014; Johnson et al., 1999; Klintö, Falk, Wilhelmsson, Schönmeyr & Becker, 2018). Britton et al. (2014) reported the speech outcomes of 1110 British children at age five. Fifty two % of children had cleft palate-related features in their speech.

2.1.2 Prevalence of otitis media with effusion and hearing loss

Otitis media with effusion (OME) is the presence of fluid within the middle ear and may be caused by a poorly functioning Eustachian tube. It is particularly common in young children because of the flatter angle of the Eustachian tube in early child development (Purdy et al., 2019). Studies assessing the prevalence of OME consistently report higher levels in babies with cleft palate up to five years of age than in babies without cleft palate (Bruce et al., 2015; Flynn & Lohmander, 2014; Flynn, Möller, Jöhnsson, & Lohmander, 2009). This is because the muscles that regulate the function of the Eustachian tube are the same muscles affected by cleft palate.

However, there is no strong evidence that OME resolves as a result of palate surgery.

Hearing loss commonly presents with OME. Flynn et al. (2009) assessed ear status and hearing levels in 22 children with unilateral cleft lip and palate (UCLP) and 21 children without cleft palate (non-cleft group) at 1, 1.5, 3 and 5 years of age. Data were analysed by ears. Across the four data collection points, data were available on 322/344 ears (162 children with UCLP and 160 non-cleft group). Prevalence of OME was significantly higher in children with UCLP at all four ages studied. It was prevalent in 121/162 (75%) ears of the children with UCLP compared to 31/160 (19%) ears in the non-cleft group. The type and level of hearing loss associated with OME is typically mild to moderate, fluctuating and conductive in nature (Broen et al., 1996; Bruce et al., 2015; Sundman, Flynn, Tengroth & Lohmander, 2016). In a review of 90 infants under two months of age with cleft palate +/- lip, Viswanathan, Vidler and Richard (2008) found the incidence of hearing loss to be as high as 82%. Flynn et al. (2009) and Flynn and Lohmander (2014) assessed hearing levels in children up to five years of age. Both studies report more prevalent hearing loss in children born with cleft palate compared to children without cleft palate, but hearing levels did improve with age. Levels of OME remained high and Flynn and Lohmander (2014) found that hearing in the high frequencies did not significantly improve. A recent study by Sell et al. (2017) of 268 British five-year-old children with unilateral cleft lip and palate found a statistically significant association between poor speech outcomes and a history of hearing loss.

2.1.3 Nature of speech sound disorder in children born with cleft palate

A medical perspective on the nature of speech sound disorder in children describes underlying causes in terms of anatomical, physiological and genetic aetiology, e.g. Shriberg et al. (2005). A linguistic perspective describes the symptoms of speech sound disorder in terms of phonetic and phonological features, e.g. Grunwell (1988). A psycholinguistic perspective describes the underlying cognitive processes that are impaired, giving rise to the manifesting symptoms, e.g. Stackhouse and Wells (1997). The majority of the literature on cleft palate-related speech sound disorder (CPSSD) uses medical and linguistic perspectives to describe the cause (altered anatomy and function) and effects (phonetic and phonological features) of cleft palate on speech development (Grunwell & Sell, 2001; Harding & Grunwell, 1996; 1998). Surgical protocols for treating cleft palate vary from centre to centre both within and across countries. The protocol in the UK is a one-stage procedure when the infant is between 6 and 12 months of age. Routine two-stage protocols, with soft palate repair in infancy followed by delayed hard palate closure, still exist in some centres, mostly outside of the UK. Although these protocol differences across studies do affect comparability, there is common agreement on how the cleft palate anatomy

directly affects an infant's sound-making ability.

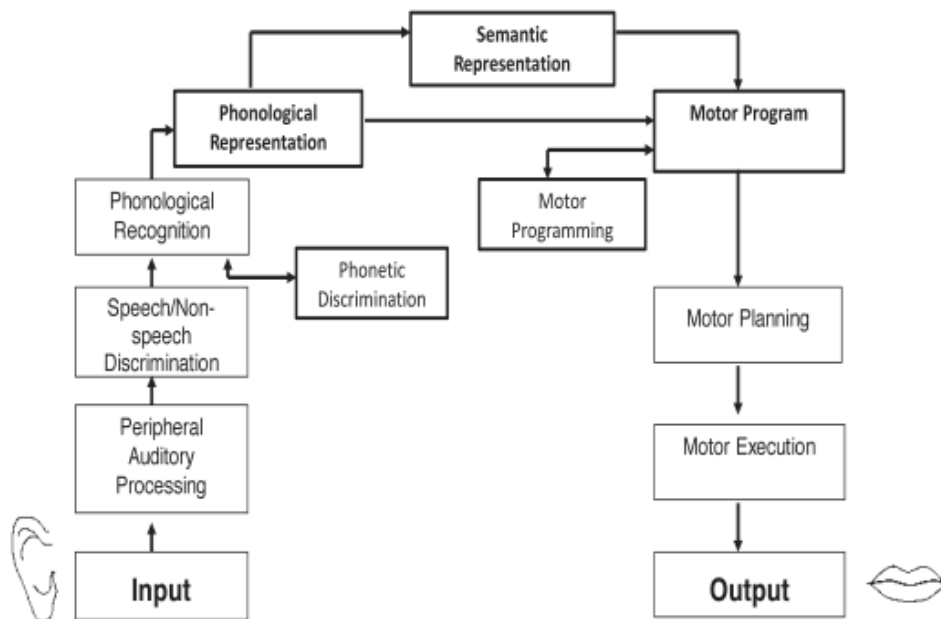
Studies show that babies with unrepaired cleft palate vocalise as much as babies without cleft palate, suggesting they have similar opportunities to experiment in sound-making (Chapman, Hardin-Jones, Schulte & Halter, 2001; Scherer, Williams & Proctor-Williams, 2008; Willadsen & Albrechtsen, 2006). However, they exhibit differences in their phonetic inventories, which may affect the quality of their early sound-making experiences. For example, at one year of age babies with unrepaired cleft palate have phonetic inventories with fewer oral consonants and more non-oral and nasal consonants, and produce fewer alveolar and palatal consonants and more glottal consonants (Chapman et al., 2001; Lohmander et al., 2011; Willadsen & Albrechtsen, 2006). These characteristics are directly related to the abnormal structure and function associated with cleft palate. It may be significant that babies born with cleft palate can take longer to reach the stage of canonical babbling and the diversity of their early canonical babble may be limited by the lower incidence of oral consonant production (Chapman et al., 2001; Chapman & Willadsen, 2011; Scherer et al., 2008a).

2.1.4 A psycholinguistic perspective on speech development

A psycholinguistic perspective on speech development conceptualises the psychological processes or components involved in processing incoming and outgoing speech (Baker, Croot, McLeod & Paul, 2001). In its simplest form, the speech processing chain is depicted as having three key levels: speech perception or input; storage of information or cognitive-linguistic representations; and speech production or output. Application of psycholinguistic models to speech disorder research has progressed descriptions of speech features into explanations of the underlying deficits within the speech processing system. Barbara Dodd and colleagues (Crosbie, Holm & Dodd, 2005; Dodd, 2005; Dodd & Bradford, 2000) used a psycholinguistic perspective to explain deficiencies that give rise to different types of speech sound disorder. For example, they make a distinction between consistent phonological disorder, caused by a deficit at the internal organisational level, and inconsistent phonological disorder, caused by a deficit in phonological planning. The capacity of a psycholinguistic perspective to encompass deficits anywhere in the speech processing chain aligns well with the nature of speech impairment in children born with cleft palate because of the impact of cleft palate on the hearing and speech mechanisms.

Here will follow a description of the theoretical speech processing model by Stackhouse and Wells (1997), since it is this model upon which MSIM+/-VT is based; see Figure 2.1.

Figure 2.1: Speech processing model by Stackhouse and Wells (1997)



The left side of the model represents input processing skills, which are involved in the perception, discrimination and recognition of the incoming speech signal. The top of the model represents lexical representations, the storage of previously processed information about individual words. The right side of the model represents output processing skills, which are involved in the retrieval of phonetic and phonological information and articulation of outgoing speech production.

Bottom-up processing of speech input begins with the reception of a sound signal by the ear and hearing mechanism (*peripheral auditory processing*). Pre-linguistic processing allows the signal to be recognised as speech rather than non-speech (*speech/non-speech discrimination*). At the next level of processing, the listener uses implicit linguistic knowledge to determine if they recognise the speech heard from a language they are familiar with (*phonological recognition*). In the early stages of language learning, off-line processing supports the recognition of phonetic distinctions that are new to the listener (*phonetic discrimination*).

Stored lexical representations allow the speech signal to be compared to existing *phonological representations*, which contain sufficient phonological information for a word to be recognised. Such recognition will trigger access to a *semantic representation*, containing information about the meaning of the word. The *motor program*, on the output side of the model, is considered to contain detailed specification of articulatory gestures required for production of known words. To learn to produce new words, new motor programs need to be created. This takes place within *motor programming*, which is conceived as a store of phonological units that can be assembled

into new combinations. It also allows existing inaccurate motor programs to be revised. Output processing requires selected motor programs to be assembled into the correct sequence with appropriate assimilations and prosody (*motor planning*). This triggers movement in the vocal mechanism and the production of speech (*motor execution*). Accurate motor execution depends on an accurate motor program.

Speech input processing requires bottom-up processing of the sound signal as well as top-down processing to enable word recognition. Although speech output processing is top-down, driven from access to a motor program to motor execution, proprioceptive and kinaesthetic feedback loops provide bottom-up feedback to modify speech output.

2.1.5 Development of speech processing skills

From a developmental perspective, the peripheral level of hearing is present from birth. Studies also indicate that infants can discriminate between different phonetic tokens (Werker, Yeung & Yoshida, 2012) suggesting speech/non-speech discrimination and phonetic discrimination are present. Although the neonate is using speech organs for feeding and crying, and later cooing and laughter, the sounds produced do not resemble speech, so motor execution within the context of speech processing is not yet developed. Anatomical changes from around four months of age give the infant increased laryngeal and articulatory control and the potential to make a broader range of sounds throughout the oral cavity (Kent, 1981). A period of vocal play follows as the infant discovers and explores their new capacity for sound-making (Stark, 1980).

By six months of age, speech-like sounds are being produced and some might be in the form of canonical babble utterances (Oller, 2000). Motor execution skills are now being used for speech. Production of such sound sequences suggests motor programming is active. Repeated babble strings constitute stored motor programs and sequences of babble strings indicate motor planning is beginning to develop. Recognition of familiar words marks the emergence of the first phonological representations (Stackhouse & Wells, 1997). At this stage, the phonological units available within motor programming are likely to be limited to those the child can articulate. Motor programs are not yet linked to semantic representations so they are not lexical in nature (Stackhouse & Wells, 1997).

By nine months of age, babble becomes more complex and takes on the characteristics of the language heard in the babies' environment (Kern, Davis & Zink, 2010). The store of units within motor programming is increasing and a wider range of patterns indicates further development of motor planning (Stackhouse & Wells, 1997). There is also evidence that babies recognise the

language spoken in their environment as opposed to a language they have never heard (Kuhl et al., 2006), suggesting the presence of some level of phonological recognition. The store of phonological representations is also increasing as the child begins to recognise more words. These become more specific over time (Sutherland & Gillon, 2007).

From 12 months of age, there is rapid expansion in the number of words the child recognises (phonological representations), understands (semantic representations) and begins to produce (motor programs). The child's spoken words often do not match the adult version, reflecting the limitations of the child's motor programs and immature motor execution skills. The vocabulary spurt through the second year of life (Dapretto & Bjork, 2000) indicates increasing numbers of phonological representations and motor programs, and associated semantic representations. Development of the phonological system mostly occurs between the ages of two and five years (Grunwell, 1981) and the child's speech output (motor programs and motor execution) starts to more closely match the adult version.

2.1.6 Cleft palate-related speech sound disorder from a psycholinguistic perspective

2.1.6.1 Effect of cleft palate on emerging speech processing skills

In the UK, surgery to repair the palate is usually undertaken between six and 12 months of age, unless it is delayed for medical reasons. By this age, infants with an intact oral structure have already undergone a period of dramatic vocal exploration and motor programming for speech-like motor programs is underway (Stackhouse & Wells, 1997). For infants with cleft palate, the unrepaired structure will therefore impose limitations on emerging motor execution skills. The feedback loop from output to input means these limitations may have a knock-on effect to motor programming and resulting motor programs, and may limit early motor planning experiences. These restrictions on motor execution may mean an absence of motor programs for oral sounds in emerging babble strings, such as [b d], which feature highly in the canonical babble of infants without cleft palate (McCune & Vihman, 2001; Morgan & Wren, 2018). If an infant's first words emerge before surgery, the implementation of motor programs for words will also be affected. Hearing loss represents a deficit in peripheral auditory processing affecting the quality of the auditory signal infants receive, and a fluctuating loss affects the stability of this signal. This may affect the accuracy of emerging phonological representations (Stackhouse & Wells, 1997).

After palate repair, the phonetic inventory increases and oral consonants emerge (Chapman & Hardin, 1992; Chapman et al., 2001; Jones, Chapman & Hardin-Jones, 2003). These changes

reflect improvements in the structure and function of the speech mechanism. Effective palate repair resolves the restrictions at the level of motor execution and gives the infant capacity to produce the full range of speech sounds. Typical speech development requires that the full range of phonological units now potentially available for motor programming are utilised in updating existing motor programs and in the development of new motor programs (Stackhouse & Wells, 1997).

This normalisation in the developing speech processing system is not always spontaneous. Cleft palate-related features might persist in the infant's pre- and emerging speech development beyond palate repair (Chapman, Hardin-Jones & Halter, 2003; Jones et al., 2003; Russell & Grunwell, 1993). In some infants there are continuing mechanical restrictions at the level of motor execution after surgery. This may be due to a hole (fistula) in the palate, either because there has been a complication in wound healing or, as is sometimes the case with very wide clefts, it may have not been possible to close all of the palate (Smyth & Wu, 2019). It might also be due to insufficiency in the velopharyngeal mechanism, where the soft palate and surrounding structures do not close effectively during speech production. These mechanical restrictions may present a challenge to the developing speech processing system, e.g. give rise to weakened consonants and nasal realisations of oral pressure consonants (Harding & Grunwell, 1996; Sell & Harding-Bell, 2012).

For many infants, surgery provides the potential to develop adequate motor execution skills, but speech does not normalise because pre-existing motor programs do not update automatically. This may reflect a limited phonological system and reduced phonological units available within motor programming. If the infant's real word production is characterised by the same restricted range of consonants as their babble, we can hypothesise that the motor programs for babble strings which had emerged prior to surgery have integrated into lexical output processing. So, even though babies are seemingly at a pre-speech level prior to surgery, early word formulation can be affected by the unrepaired structure and its impact on the (fragile) emerging speech processing system.

2.1.6.2 Classification of cleft palate-related speech sound disorder (CPSSD)

Calladine (2009) reflected on the phonetic and phonological consequences of cleft palate to provide the first interpretative account of how a cleft palate can impact on a child's developing speech processing system (Treslove, 2014). This work has been extended by Calladine and Vance (2019), who propose a framework for classifying CPSSD. It makes a distinction between:

1. Impairment at the level of motor execution, resulting in inaccurate production
2. Impairment at the level of motor programming, resulting in inaccurate motor program

Impairment at the level of motor execution

This arises when the structure and function of the child's speech mechanism remains deficient following primary palate repair. There are on-going deficiencies in motor execution skills as a result of a palatal fistula or inadequate velopharyngeal function and, in older children, dental and occlusal anomalies. These will affect production or implementation of the motor program, such that all the elements of the motor program may or may not be realised, depending on how closely it matches the child's existing motor execution skills. For example, a significant palatal fistula may cause leakage of intra-oral pressure, giving rise to weakened consonants. Inadequate velopharyngeal function may result in coupling of the oral and nasal cavities such that the child is unable to produce voiced oral pressure consonants, e.g. /b d/, and they will be perceived by listeners as nasal consonants, e.g. /m n/. This type of passive process ('nasal realisation') compromises the child's ability to signal contrasts between oral and nasal sounds (Harding & Grunwell, 1998; Sell & Harding-Bell, 2012). A phonetic inventory consisting only of vowels and the consonants [m n w g] is strongly indicative of a persisting mechanical deficiency (Sell & Harding-Bell, 2012).

Low-tech methods to normalise the child's structure can sometimes help diagnose impairment in motor execution, such as occluding the child's nose during sound production. The findings from oral examination and evidence elsewhere in the child's speech output will further inform differential diagnosis. Whilst nasal realisations are a key indicator of an impairment in motor execution, typically requiring surgical intervention (Harding-Bell & Howard, 2011), the findings by Hardin-Jones and Chapman (2018) and case studies by Calladine and Vance (2019) and Hodson, Chin, Redmond and Simpson (1983) show that this is not always the case. They may instead represent inaccurate motor programs caused by impairment elsewhere in the child's speech processing system.

Impairment at the level of motor programming

This level of impairment affects the identity or accuracy of the motor program. Palate repair aims to provide the infant with potential to develop adequate motor execution skills. If babble patterns remain restricted after surgery, it is possible that there is potential for expansion of the phonetic repertoire but previously established motor programs for babble have become frozen (Bryan & Howard, 1992; Vance, Stackhouse & Wells, 2005). The range of potential phonological

units available to motor programming is therefore not updated or expanded. From 12 months, as babble progresses towards meaningful language, the range of phonetic motor programs usually includes [m b n d]. However, the restricted range of phonetic motor programs used by some infants born with cleft palate limits the range utilised during the phase of expanding phonotactic structures and in the construction of lexical motor programs. The restricted range of motor programs available to these infants is also likely to have been affected by inaccurate or 'fuzzy' phonological representations as a result of hearing loss. Impaired hearing may impact on the accuracy of phonological information extracted from auditory input, giving rise to inaccurate phonological representations and imprecise motor programs (Pascoe, Stackhouse & Wells, 2006).

'Backing to velar' and 'glottal articulation' are referred to as active cleft speech processes (Harding & Grunwell, 1998). They typically affect classes of sounds and may reduce contrastivity (Harding-Bell & Howard, 2011). Both have been associated with impairment in the structure and/or function of the speech mechanism. For example, glottal articulation has been referred to as a 'compensatory articulation' and attributed to velopharyngeal insufficiency (Chapman, 1991; Trost, 1981) and backing to velar has been linked to palatal fistulae (Lohmander, Persson & Owman-Moll, 2002). Their origin can be explained by initial restrictions on the child's motor execution skills before primary palate repair. This limits the child's ability to produce motor programs for oral pressure consonants, especially those with an anterior place of articulation, such as bilabial and alveolar. This leads to the development of compensatory motor programs that use velar and/or glottal articulations to mark some aspect of the consonant target, which become integrated into the child's phonological processing. This depicts the production of velar and glottal articulations as the result of faulty motor programming (Harding & Grunwell, 1998). We can therefore hypothesise that children who display backing and glottal processes after surgery have impairment at the level of motor programming (and not necessarily impairment at the level of motor execution): existing motor programs have not updated and/or motor programming remains limited in the range of phonological units available to create new motor programs.

2.1.7 Considerations in the differential diagnosis of CPSSD

In a recent study by Hardin-Jones and Chapman (2018), 76% (26/34) of children with repaired cleft palate produced nasal substitutions in their real words after surgery, most commonly between 21 and 27 months of age. They were also seen in 35% (7/20) of children without cleft palate, most commonly at 21 months. Across both groups, the pressure consonants /b d/ were the most commonly substituted sounds. An important conclusion of the study was that nasal

substitutions observed in the children with repaired cleft palate were not always associated with an inability to produce the substituted sounds. None of the children had secondary palatal surgery during the follow-up period and, at 39 months of age, only 38% (10/26) demonstrated symptoms of a physical problem (inadequate velopharyngeal mechanism) requiring assessment or treatment. Similar observations are reported elsewhere in the literature (Calladine & Vance, 2019; Hodson et al., 1983). The authors suggest a strong developmental component to this speech phenomenon, reflective of the high variability seen in young children's word production at this pre-systematic stage of phonological development (Ferguson & Farwell, 1975; Grunwell, 1981). In children with repaired cleft palate, an interaction will exist between this characteristic developmental variability (unstable motor programs) and the lexical integration of motor programs (frozen motor programs).

Within Stackhouse and Wells' (1997) model, inaccurate phonological representations can also account for inaccurate motor programs, such as nasal realisations (Purdy et al., 2019; Stackhouse & Wells, 1997). Calladine and Vance (2019) observed inconsistent nasal realisations in Sophie's speech, a 37-month-old child with repaired cleft palate and a history of OME and mild bilateral conductive hearing loss. Sophie's production of /d/ in word-initial position was either [d] or [n]. Her [d] realisations were evidence that she had an accurate motor program for /d/ and the physical capacity to produce it. This led the authors to hypothesise that Sophie's [n] realisations were the consequence of an inaccurate phonological representation, causing her to utilise an inaccurate motor program in output processing. Given Sophie's hearing history, impairment in peripheral auditory processing may have compromised the accuracy of her phonological representations (Purdy et al., 2019). These nasal realisations were not targeted directly in Sophie's therapy but they later resolved in spite of this, thereby supporting the hypothesis that they were phonological in nature (inaccurate phonological representation and motor program). The subject of Hodson et al.'s (1983) study, Tim, who also produced nasal realisations, similarly had a history of recurrent OME and persisting mild conductive hearing loss. Purdy et al. (2019) draw on the acoustics of speech perception and production in an attempt to explain a possible association between phonological nasal realisations and history of OME and hearing loss. These are important considerations in the differential diagnosis process so that effective decisions are made about treatment. As such, this forms an important part of the rationale for early therapy intervention for children with cleft palate, OME and hearing loss.

2.2 Section 2: Early remedial therapy for children with CPSSD

The features of CPSSD can be identified in children as young as 18 months and risk factors even earlier (Barrett & Extence, 2010; Bowden et al., 1997; Chapman et al., 2003; Jones et al., 2003; Williams & van Eeden, 2019). In the UK, a national standard of care requires all children born with cleft palate to be seen for their first speech assessment by a specialist speech and language therapist (SLT) before 27 months of age (Lead SLT Forum, 2016); they are typically carried out at 18-24 months. This provides opportunity for early identification and in turn early intervention. Early intervention is justified by studies demonstrating the impact that persisting speech sound disorder can have on other aspects of a child's development and their life activities and participation in later childhood and adulthood (Bercow, 2018; Bishop & Adams, 1990; McCormack, Harrison, McLeod & McAllister, 2011). In respect of CPSSD, its early onset, prevalence and persistence present a strong case for early intervention before three years of age (Britton et al., 2014; Harding-Bell & Howard, 2011).

Speech outcomes in children born with cleft palate are assessed at five years of age and are reported and compared nationally (Britton et al., 2014; CRANE, 2018). In the Sell et al. (2017) study, children who had received preschool therapy intervention were three times more likely to have good speech outcomes at five. Poor speech outcomes were associated with palatal fistula and surgery for velopharyngeal insufficiency (VPI), which highlights the importance of adequate structure and function. In a recent study by Smyth and Wu (2019) of 410 five-year-old children with repaired cleft palate, children who had developed a fistula following primary palate repair were three times more likely to have a poor structural speech outcome. Since significant fistulae and VPI can limit a child's ability to change their consonant production (Harding & Grunwell, 1998; Henningson & Isberg, 1987), effective treatment in the preschool years is essential for enabling children to achieve optimal speech outcomes by the time they start school. Speech and language therapists have an important role in differential diagnosis, and early intervention facilitates this process and informs decisions about treatment (Enderby et al., 2009; Harding-Bell & Howard, 2011; Scherer & Louw, 2011; Sell & Harding-Bell, 2012).

There is a wide range of intervention approaches available to treat speech sound disorder in children (Williams, McLeod & McCauley, 2010). However, many are unsuitable for very young children because of the demands they place on a child's attention, cognition and language skills (Harding-Bell & Howard, 2011). Bessell et al. (2013) conducted a systematic review of speech and language therapy interventions for children born with cleft palate. The studies represented a range of intervention approaches: some articulation; others phonological; others a combination.

This reflects the underlying nature of CPSSD. The authors identified significant limitations in the evidence base and concluded that there is currently insufficient evidence to support any one specific model of intervention for this client group. Age at intervention varied across studies and only three interventions involved children younger than three years of age. One of these studies, Regan and Versaci (1977), evaluated the use of non-speech oral motor exercises, which do not have current support as an effective method for treating speech disorder in children (McCauley, Strand, Lof, Schooling & Frymark, 2009). The two other studies (Hardin-Jones & Chapman, 2008, and Scherer, D'Antonio & McGahey, 2008) examined outcomes in children who had received early intervention consisting of naturalistic approaches. Whilst the volume of published studies in this area remains small, with only one new study (Kaiser, Scherer, Frey & Roberts, 2017) since the Bessell et al. (2013) review, the majority have focused on similar approaches.

2.2.1 Naturalistic approaches

Naturalistic approaches, such as focused stimulation and enhanced milieu teaching (EMT), are language interventions to develop language and communication skills, whilst simultaneously promoting phonological development (Girolametto, Pearce & Weitzman, 1996; Kaiser & Hampton, 2017; Weismer, Venker & Robertson, 2016). Their design reflects a blend of principles from several different theoretical frameworks, including behavioural, social and developmental. From the behavioural perspective, the antecedent-behaviour-consequence paradigm creates opportunities for language learning, and imitation and production practice are considered key to learning. The adult 'teacher' uses strategies to prompt the child to imitate speech and language forms and initiate language, thus creating opportunity for them to engage in production practice (Kaiser & Hampton, 2017). The adult provides reinforcement through feedback and modelling in order to shape new learning. Social perspectives on language learning perceive that language is learned through meaningful interactions with others. Naturalistic approaches therefore use engaging play interactions and conversations as the context for stimulating, prompting and reinforcing speech and language forms (Kaiser & Hampton, 2017). The adult manipulates the play environment so there is opportunity to model selected speech and language targets with high levels of repetition, but activities are not adult-led; the adult follows the child's lead, responding to the child's focus of attention and interest. The adult does not directly elicit output from the child, instead she models speech and language targets and creates opportunities for the child to imitate and initiate. A developmental perspective on speech and language acquisition influences target selection. Targets are real words, rather than single sounds, are functional for the child, and reflect both vocabulary and phonological targets (Scherer & Kaiser, 2010). The behavioural perspective on acquisition, generalisation and maintenance as critical phases of learning also

informs the use of multiple contexts and multiple exemplars for target sounds and words. These principles suggest target words are likely to be known to the child and, where known, lexical representations will exist. This suggests naturalistic approaches target phonetic and phonological change from a linguistic top-down direction.

A large number of studies have demonstrated the efficacy and effectiveness of naturalistic approaches for children with delayed language development and autistic spectrum disorder (Kaiser & Hampton, 2017; Schreibman et al., 2015; Weismer et al., 2016). A small number of studies have explored efficacy for children with cleft palate. Scherer (1999) demonstrated increases in vocabulary and phonetic inventories in three two-year-old children with repaired cleft palate following EMT intervention. However, the nature of phonetic expansion may have reflected typical phonological development rather than resolution of cleft palate-related speech features. For example, all three children acquired voiceless counterparts to consonants that existed in their voiced form prior to intervention. In a larger study, Scherer et al. (2008b) showed that focused stimulation intervention was delivered effectively as part of a parent-implemented program for children born with cleft palate aged 14 to 36 months. The mothers of 10 children with cleft palate were trained to deliver the intervention. Speech and language measures were compared with 10 children without cleft palate and their mothers, who formed a no intervention control group. Mothers in the intervention group used more language facilitation strategies following intervention than mothers in the control group. Although children in the intervention group did show a significant reduction in their glottal plosive usage, the absence of glottal plosives in the non-cleft control group limits the extent to which this can be attributed to intervention. Similarly, although children showed significant increases in true consonant inventory, percentage consonants correct and vocabulary size, these were also seen in the non-cleft control group.

In contrast to Scherer (1999) and Scherer et al. (2008b), Hardin-Jones and Chapman (2008) did not find a significant benefit of early intervention in their study, but a number of factors limit the validity and comparability of this finding. The study involved 40 children, 30 with repaired cleft palate and 10 without cleft palate. Speech measures were taken at 17 and 27 months of age. Ten of the children with repaired cleft palate had received therapy, 10 had been referred but did not receive it (no therapy group), and 20 were not referred. It was a retrospective study so the authors were unable to provide explicit descriptions of the intervention that children received. Intervention was delivered by six clinicians with different levels of experience and no criteria to measure their competency in the intervention. The authors did not know the amount of parent training that was provided or the extent to which treatment goals and strategies had been

incorporated into the child's daily activities. Children in the therapy group demonstrated similar phonetic and lexical profiles to children in the no therapy group at 17 months of age. Although children who received therapy produced higher levels of correct consonant production at 27 months of age, this was not (except for glides) statistically significant. The authors conclude that it is unclear if vocabulary-based approaches can produce better or comparable outcomes than speech-based approaches for children with severely restricted phonetic development.

Enhanced milieu teaching with phonological emphasis (EMT+PE; Scherer & Kaiser, 2010) is an adaptation of the EMT approach described by Scherer (1999). Phonological emphasis involves the therapist modelling words that contain target sounds and providing 'phonological recasts' in response to the child's unintelligible or partially unintelligible utterances. A phonological recast is the therapist's repetition of the child's utterance using the correct phonological form; it gives feedback to the child to modify his/her production. Kaiser et al. (2017) investigated the effects of EMT+PE provided by speech and language therapists (SLTs) to toddlers born with cleft palate aged 15 to 36 months of age. The study had several strong design features. Children were randomly allocated to one of two treatment conditions: eight children received EMT+PE and 11 children received business as usual (BAU). Enhanced milieu teaching with phonological emphasis was delivered by trained SLTs who were required to demonstrate competency prior to taking part. A systematic procedure was put in place to measure fidelity throughout the intervention period and high levels were reported. In the BAU group, six children received community-based speech and language therapy intervention, which may or may not have involved EMT+PE, and five children did not receive any therapy. In the EMT+PE group, the authors describe how intervention was differentiated so that children had individualised language and speech sound targets "consistent with their emerging abilities" (p809). The authors do not provide further details of the targets for individual children so it is unclear precisely what this means. It may mean that children's targets were sounds they were producing accurately on some occasions but not others. If so, this would suggest they had accurate motor programs for the target sounds but they were not being implemented all of the time. This could be because of inaccurate phonological representations or because motor programs have not been fully updated.

The study found that children in the EMT+PE group showed significant gains in some, but not all, of the speech and language measures. For example, they made greater gains in percentage of consonants correct (PCC; measured using Profiles of Early Phonological Skills (PEEPS); Stoel-Gammon & Williams, 2013) and expressive vocabulary (measured using the parent-completed MacArthur-Bates Communication Development Inventory; Fenson et al., 2007), but not in intelligibility. The word lists in the PEEPS assessment are based on developmental norms of

vocabulary and phonological acquisition, and we know that children born with cleft palate have altered speech and lexical development. It is therefore possible that the PCC measure does not fully reflect phonetic or phonological changes made by children in both groups. Furthermore, the authors acknowledge that measures may have been over or underestimates because not all test items were elicited at both pre- and post-test assessments. Given the young age of children and developing nature of speech production systems in young children, percentage of consonants correct-adjusted (PCC-A; Shriberg, Austin, Lewis, McSweeney & Wilson, 1997), which does not count common distortions of sounds (e.g. dentalisation) as errors, may have been a more suitable measure. Willadsen et al. (2018) used PCC-A to report articulation outcomes in three-year-old Danish children with unilateral cleft lip and palate. Percentage of consonants correct-revised (PCC-R; Shriberg et al., 1997) is an alternative measure that additionally does not count uncommon distortions of sounds (e.g. palatalisation) as errors. Such a measure might more effectively capture changes in articulatory placement that are deemed favourable in a developing child who may have structural restrictions on articulatory precision, e.g. /s/ → [X]_{PRE-INTERVENTION} and [sʲ]_{POST-INTERVENTION}.

In their discussion, Kaiser et al. (2017) highlight the multicomponent nature of EMT+PE. Their study was unable to determine the impact of individual components and they encountered several challenges because of its multicomponent design, e.g. selecting appropriate speech and language targets that could be embedded within naturalistic interactions and when following the child's lead. They also acknowledged that low rates of verbal communication in such young children present a challenge for implementing the EMT+PE approach. They suggest a sequential approach may be more suitable, which would begin with the EMT component to increase rates of expression followed by the addition of PE strategies to promote accuracy of production.

These existing descriptions of naturalistic interventions for toddlers with cleft palate suggest they are designed to stimulate changes in speech output from a top-down direction. Repetitive models of functional real words will tap a child's existing lexical knowledge. If the child imitates a word with inaccurate phonological form, the therapist's recast will inform the child that there is a mismatch. Kaiser et al.'s (2017) findings suggest EMT+PE was sufficient for triggering change in stored lexical representations, but we do not know enough about the nature of impairment in the children's speech processing systems nor the specific nature of changes that occurred.

Multisensory input modelling (MSIM; Harding & Bryan, 2000) and MISIM with output (MSIM+O; Calladine & Vance, 2019) are examples of alternative approaches for treating CPSSD in young children. They share some of the same principles as naturalistic approaches, but are grounded in

a different theoretical perspective, which informs a distinctly different approach to target selection and stimulus design.

2.2.2 Multisensory input therapy

Multisensory input modelling (MSIM; Harding, 2001; Harding & Bryan, 2000) was designed to be used with young children with speech output difficulties for whom output-based approaches are inappropriate. The nature of the clinician–child relationship is modelled on Pamplona and colleagues’ non-directive and non-corrective naturalistic approach (Pamplona, Ysunza & Jiménez-Murat, 2001). It is more clinician-led than some naturalistic approaches but adheres to Pamplona and Ysunza’s (2000) recommendation that young children will respond more positively when they are not required to perform and there is no risk of failure. The clinician engages the child in activities within which innovative behaviour is repetitively modelled and the child is invited to fulfil a role without expectation of speech output practice.

Multisensory input modelling is based on Stackhouse and Wells’ (1997) speech processing model. It aims to create a context within which new motor programming for specific articulatory gestures evolves gradually. Since it is likely that inaccurate motor programs are rooted at the level of motor execution rather than phonological representation, input therapy stimulates output processing for sounds as articulatory gestures rather than as units of lexical knowledge. For example, [f:] as a voiceless sound made with a sustained oral airstream and gentle contact between the upper front teeth and lower lip, rather than /f/, the phoneme at the beginning of /fi:f/. The rationale for repetitive exposure to models of speech stimuli is provided by mirror neurone theory, which has shown that the neural pathways stimulated by observation of a specific behaviour are the same pathways stimulated by subsequent performance of the same behaviour (Oberman, Pineda & Ramachandran, 2007). Frequent observation of appropriate and achievable articulatory gestures can therefore be assumed to develop the neural pathways in preparation for the time when output processing for that gesture, or sequence of gestures, is initiated. Infant imitation theory also suggests that infants and young children will naturally imitate novel adult models of unusual sounds and words (Meltzoff, 1988). The MSIM approach seeks to promote natural, evolutionary learning and behavioural change similar to the process of typical development. The process of change from a restricted consonant repertoire in babble and early words to the use of mature phonological and linguistic systems seems to occur with minimal conscious awareness (Vihman & Harding-Bell, 2019). Therapy to change speech while patterns are still evolving has two aims: to destabilise atypical patterns while they might still be susceptible to change; and to influence continuing phonological development. Children with a

history of fluctuating conductive hearing loss may benefit from a multisensory approach to updating phonological representations. In addition to supplementing auditory perception, innovative visual and tactile stimuli may increase the salience of new sounds, increasing the likelihood of both neural mapping and spontaneous imitation.

Dodd's (1979) research into the role of vision in speech perception confirms that supplementing acoustic information in speech tasks with visual, tactile and, later, orthographic referents is a well-grounded precaution with all children. In MSIM, stimuli are modelled with attention to visibility and audibility. The therapist sits opposite the child, holds toys or objects up to his/her face and models sounds at close proximity so the acoustic features do not degrade with distance. Tactile feedback of airflow by producing sounds close to the palm or back of the hand or the cheek can provide additional sensory information (Peterson-Falzone, Trost-Cardamone, Karnell & Hardin-Jones, 2006; Russell & Albery, 2005). This may be performed in different ways as it is not a manualised technique. Multiple repetitions ensure that the visual and acoustic features of the modelled sounds are sufficiently salient to ensure storage of accurate phonetic information.

The non-directive, non-corrective and facilitative style of MSIM is underpinned by pragmatic approaches, such as social learning theory, naturalistic approaches, communicative facilitation techniques, and focused stimulation (Andersen-Wood & Smith, 2001; Bandura, 1977; Harding-Bell, personal communication). Activities and roles are designed to ensure the child does not experience any sense of failure. The vocabulary used is encouraging, gently spoken and affirming and therapy sessions are designed to be collaborative. A primary goal of MSIM is to create an environment where young children develop confidence to initiate communicative exchanges. Whether or not consonant production is normalised with therapy, it has been suggested that confident communicative participation may be an equally valuable therapy effect (Thomas-Stonell, Oddson, Robertson & Rosenbaum, 2009; Thomas-Stonell, Washington, Oddson, Robertson & Rosenbaum, 2013).

2.2.3 Multisensory input modelling with output

Calladine and Vance (2019) describe an extension to the original MSIM approach. Their description draws on underlying theory, empirical evidence from the study by Calladine (2009) and two subsequent clinical case studies, and the clinical experience accumulated by Dr Anne Harding-Bell and the lead author. This extension of MSIM focuses on application of the approach with young children from around 18-24 months of age.

Multisensory input modelling with output (MSIM+O; Calladine, 2009, 2010; Treslove, 2014) is a

development of MSIM (Harding & Bryan, 2000). Calladine (2009) hypothesised that when a child spontaneously imitates a real word with a mismatched production, lexical speech processing has occurred. Acceptance of the mismatched imitation could potentially be counterproductive if it reinforces existing inaccurate motor programs. Therefore, MSIM+O incorporates explicit verbal feedback for both matched and mismatched productions. This includes the use of specific articulatory and phonetic descriptions and frequent praise and encouragement. A specific approach to stimuli selection is also adopted. Invented words are used as an alternative to sound sequences since they may be more engaging and linguistically evocative for young children. It is hypothesised that these novel but meaningful sound sequences have potential to generate neural pathways through input processing whilst avoiding mapping to pathways that are linked to an existing phonological representation or motor program. The use of invented words to stimulate specific aspects of processing has been reported in other speech sound intervention approaches, such as Parents and Children Together (PACT; Bowen & Cupples, 1999) and complexity approaches (see Baker & Williams, 2010).

Calladine (2009) evaluated the effect of MSIM+O on speech output in four two-year-old children with repaired cleft palate using a non-experimental pre-post treatment case study design. The intervention period was three months. Two children received six fortnightly sessions with video-therapy (non-standard model of service delivery) and two children received three monthly sessions without video-therapy (standard model). All children had familiar and unfamiliar real word targets to support comparison of the speech processing involved: updating existing motor programs (for familiar words) versus generating new motor programs (for unfamiliar words). Untreated words were used to probe generalisation. For all children, the intervention included a specific procedure for responding to child productions of target stimuli based on Gardner (1994). This involved explicit articulatory or phonetic feedback in response to matched and mismatched productions. In responding to a mismatched production, the therapist highlighted the difference between her and the child's production by referring to her own articulation, e.g. 'I do it with my lips.'

The Phonological Assessment of Child Speech (PACS Toys; Grunwell & Harding, 1995) was used to categorise the nature of target sound realisations and measure changes in speech output. These were analysed using Stackhouse and Wells' (1997) psycholinguistic theory of speech processing. All children made matched and mismatched productions for target words during therapy and modified mismatched productions closer to target in response to therapist feedback. The analysis found that all four children showed favourable changes in their speech output and the nature of these changes reflected their specific targets: three of the four children acquired 10 new sounds

(range 2-5) that were absent from their phonetic inventories or marginal pre-intervention; the fourth child did not acquire new sounds but ceased to use four atypical sounds. All four children produced more accurate target sound realisations in both treated and untreated words following intervention. The two children who received the non-standard model of service delivery, which provided a higher dosage of intervention, made more progress on treated words. However, the heterogeneity of participants and study design prohibited a valid comparison of these two models of delivery in this study. Analysis in relation to word knowledge provided some evidence that invented word stimuli facilitate bottom-up processing to create new motor programs and avoid mapping to existing lexical knowledge.

Calladine's (2009) findings provide empirical evidence that two-year-old children can be engaged to make and modify target speech productions within the context of MSIM+O and in response to explicit feedback. These findings have contributed to existing descriptions of the approach (Treslove, 2014; Calladine & Vance, 2019). However, the study does not provide strong evidence for the impact of MSIM+O due to its methodological limitations. The study raised questions about the optimum nature of target selection and stimulus design and how engagement in therapy and output elicitation is facilitated by the nature of therapist-child interaction.

2.2.3.1 Aims and goals of therapy and therapy techniques

The principal aim of MSIM+O is to influence change in the child's output processing by helping the child establish an inventory of accurate motor programs linked to accurate phonological representations. In keeping with the original MSIM approach, MSIM+O aims to stimulate change from a bottom-up direction by creating new off-line motor programs and gradually integrating them into the child's speech output. Activities and stimuli are designed to give the child opportunity to observe, and eventually practice, the new accurate motor programs. A gradual process of varying the input models in small incremental steps leads the child in their awareness, and eventually their output processing, so that motor programs are gradually introduced into their phonological system. 'New' and 'old' motor programs and representations may exist in equipose at certain stages. Subsequent activities and stimuli are designed to help the child select the new accurate motor program rather than the original by strengthening links between new motor programs and their corresponding phonological representations that will over-ride existing connections. These stages can be summarised as follows:

1. Stimulate the development of new accurate motor programs for target sounds using single sound stimuli
2. Help the child to stabilise new accurate motor programs at a single sound level

3. Stimulate the development of phonological representations and motor programs for sound sequences and new words containing target sounds using invented word stimuli
4. Strengthen links between motor programs for these new words and the corresponding phonological representations
5. Establish links between new motor programs and stored phonological representations for known real words containing the target sounds using real word stimuli
6. Destabilise links between phonological representations and old inaccurate motor programs

Calladine and Vance (2019) describe therapy activities as play-based and input-based. The therapist does not present the child with tasks; rather she leads the child towards desired behaviours through engaging activities. The responsibility for successfully fulfilling a 'task' lies with the therapist. The tasks include activities, which consist of the following components:

1. Repetitive modelling of speech sound stimuli
2. Modelling provided at close proximity to the child for quality of acoustic signal
3. Therapist sits opposite the child (except when filming) for clarity of visual signal
4. Multisensory stimulation providing auditory, visual and tactile information
5. Specific verbal description and labels to draw the child's attention to articulatory and phonetic features of speech stimuli
6. Provision of clear semantic associations for speech stimuli
7. Explicit verbal and nonverbal feedback given to the child for spontaneous productions
8. Encouragement for the child to be an active participator

Specific verbal description and labels highlight particular features of the modelled sound using vocabulary that is meaningful to the child, e.g. 'with my lips' for bilabial sounds, 'soft' or 'quiet' for voiceless sounds, 'I made the bubble sound' for the [p] stimulus. The therapist uses 'I' and 'my' to draw the child's attention to important features of the sound s/he models. Along with multisensory stimulation, these descriptions promote the development of phonetic information for use in the development of new motor programs.

Responding to child productions

The procedure for responding to child productions is based on the work by Gardner (1994; 2006). When the child makes a stimulus production, the therapist provides explicit feedback to reinforce a matched production (accurate motor program) or encourage reflection and repair of a mismatched production (inaccurate motor program). To confirm a matched production, feedback refers specifically to features of the child's output that match the modelled sound, e.g., 'you used

your tongue' or 'you did the drum sound'. To encourage repair of a mismatched production, feedback refers to the mismatched feature(s). In keeping with the non-corrective style of MSIM+O, feedback focuses on the therapist's modelled stimulus rather than the child's production, e.g. 'I used my lips', 'I did the ball sound'. Cued articulation gestures (Passy, 1993) or other nonverbal gestures, e.g. pointing to the lips, provide extra visual support.

Video-therapy

Video-therapy is the process of filming therapy activities for parents/carers to watch with their child at home in-between therapist-led sessions. Harding and Bryan (2000) describe it as a tool for increasing the frequency of input stimulation, giving the child consistent and accurate models of speech sound stimuli, and providing opportunity for the child to participate in input processing, rehearsal and experimentation at their most attentive times. In the study by Calladine (2009), video-therapy was used in two of the four children's intervention. The therapist provided a new therapy video after every session. She asked parents to watch the video with their child once or twice a day when the child was attentive. Parents completed a video diary to log the time their child spent watching the therapy videos, who they watched with (which parent and anyone else who was present, e.g. sibling), and what observations the parent made, e.g. did the child copy any of the productions. Analysis of the diaries found that both children typically watched their video on 10 separate occasions between two consecutive therapy sessions. One of the children, Max, typically watched one activity (about 10 minutes long) at a time whereas the other, Sameeah, watched the full video (about 25 minutes long). Widdowson (2015) described a similar procedure for home implementation of therapy videos in her study of MSIVT involving seven children aged 2;2 to 3;9 with emerging signs of speech disorder in the absence of cleft palate. In this study, the therapist asked parents to watch the video with their child for at least five minutes every other day. Quantitative data from video diaries was not reported.

Both studies report qualitative findings from the video diaries completed by parents. They show that parents found the videos helpful by serving as a reminder of how to model sounds and complete the activities. They also show that the videos supported other family members who did not routinely attend sessions to be involved in the therapy (Calladine, 2009; Widdowson, 2015). These studies provide empirical evidence of how video-therapy is implemented outside of MSIVT sessions and the role the therapy videos can play. However, they are the only known studies that have reported on this aspect of MSIVT delivery, indicating the limited availability of evidence in this area.

The two case studies presented by Calladine and Vance (2019) provide empirical evidence of how

video-therapy is implemented within MSIVT sessions. Their videos demonstrate the therapist (the lead author) using the video camera as a tool in activities, asking the child to show pictures to the 'telly'. The therapist also appears to use the camera as a recipient of models when the child is engaged in another activity, which appears to lessen the need for direct requests for the child's attention. This account, however, is descriptive not analytical and draws on clinical data only. There is a need to examine the nature of video-therapy within and outside of therapy sessions so that this component of therapy can be better understood and defined.

2.2.3.2 Stimulus design and selection

In MSIM+O, stimulus refers to the sound modelled by the clinician for a given target phoneme (sound in the child's native language). There is principled guidance in the cleft palate speech literature on the selection of therapy targets (Harding & Grunwell, 1998; Harding-Bell & Howard, 2011). It is underpinned by linguistic theory and takes into account the phonetics and physiology of consonant production and the structural and functional characteristics of cleft palate and hearing loss. The developmental trajectory of consonant acquisition should be considered but is usually second to physiological considerations in children with cleft palate. A selected stimulus is modelled in such a way as to maximise the likelihood of it being perceived by the child as novel. This is so it does not immediately stimulate a stored motor program for the corresponding target; for example, an interdental [t^{h}] or linguolabial [t^{h}] model for an alveolar target helps to avoid the stimulus being recognised as the phoneme /t/ in phoneme recognition. It is recommended that articulation is relaxed and not exaggerated or effortful. However, it may be necessary to introduce an adaptive placement to achieve novelty or to adapt around physical constraints, which can be seen as a short-term facilitator. It is important the child may have success in matching the stimulus if they produce it themselves. Effortful articulation may encourage or perpetuate compensatory articulation (Russell & Harding, 2001) so the stimulus should be modelled with minimal effort using gentle placement and soft, stretched airflow. Verbal labels for the stimuli apply meaning and concreteness, e.g. [f^{h}] = 'leaky balloon sound' (see Harding & Bryan, 2000, and Speake & Harding-Bell, 2019, for recommendations). An important principle of MSIM+O is that the child's stimulus productions are interpreted using Stackhouse and Wells' (1997) speech processing model. If output suggests an inaccurate motor program was produced, it would be appropriate to modify the stimulus and/or the label.

The therapist initially models the stimulus as a single sound to make phonetic features salient. Calladine and Vance (2019) recommend progressing to sound sequences with meaning, also known as invented words, e.g. names for characters or food. Simple consonant and vowel (CV),

VC and VCV structures are recommended depending on the selected target word position(s). Modelling the stimulus in real words (assumed to be known to the child) follows. Ideally, real-word modelling is delayed until there is successful spontaneous output at the invented word level. Overlap between stages in the sequence highlights the common stimulus at both levels; for example, when moving from single sound to invented words, characters with invented names can be introduced into a single sound modelling activity, and when moving from invented word to real words, real-word stimuli can be introduced in an activity with the named characters.

2.3 Conclusions

This chapter has drawn on the existing literature to provide an explanation of the phonetic and phonological nature of cleft palate-related speech sound disorder (CPSSD) and the effect of cleft palate and hearing loss on the developing speech processing system. Early diagnosis of CPSSD provides opportunity for early intervention. Currently, the two main approaches to early intervention are naturalistic and speech-based. A review of the literature on naturalistic approaches, namely EMT+PE, has found that evidence of efficacy is emerging. However, there are challenges with using this type of approach to target speech development in young children at an early stage of language development because of their low levels of verbal output (Kaiser et al., 2017). Speech-based approaches, which target speech processing directly, are typically not suitable for children under three years of age. However, the MSIM+/-O approach has been designed with features to support its suitability for very young children (Calladine & Vance, 2019; Harding & Bryan, 2000). It is input-based and uses a non-directive and non-corrective style of interaction to create opportunities for output rather than engaging the child in explicit output practice. Therapy videos can be made and provided to families to watch in-between therapy sessions; a process known as video-therapy (VT).

This review of the literature has found that current descriptions of MSIM+/-VT are rooted in coherent theory. They suggest the approach is suited to the specific nature of CPSSD and the early nature of intervention for children born with cleft palate. However, as shown, very few experimental studies have been carried out to investigate MSIM+/-VT and this is a significant limitation of the current evidence base supporting its use in clinical practice. There is likely to be variation in how it is delivered from one SLT to another and key components of the approach are not fully understood. This must be addressed to support its use in clinical practice and so that the intervention can be adequately defined in order to be evaluated (Craig et al., 2008; MRC, 2000). This is an important research priority in the UK. Questions about the optimum age at which to start intervention and optimum method and model of delivery are in the top 12 research

priorities identified by individuals, families and professionals affected by cleft lip and palate (James Lind Alliance, 2012).

The findings of the literature review support an investigation into the nature of MSIVT as it is implemented in the NHS as an early intervention to treat CPSSD in young children, in order to identify and describe characteristic features of the approach and make hypotheses about its key components. A pilot study was undertaken to trial methods for analysing therapy sessions and to inform the specific aims and research questions for the main study.

Chapter 3: Pilot Study

This chapter presents the methods and results of the pilot study and explains how the findings informed Phases 1, 2 and 3. The study aimed to examine the author's own clinical practice of multisensory input video-therapy (MSIVT) in order to formulate ideas for how to analyse the therapy and develop skills in managing and organising data extracted from NHS clinical records.

3.1 Methods

3.1.1 Design

The study used retrospective data generated from clinical records at the author's place of work, a specialist speech and language therapy (SLT) service hosted by a regional cleft lip and palate centre in the NHS. One SLT and two children participated, forming two therapist-child dyads. Data consisted of video recordings and written information on 30 MSIVT sessions delivered as part of routine clinical care between January and December 2012. The study trialled qualitative and quantitative methods for analysing the therapy data.

3.1.2 Procedure

3.1.2.1 Ethical approval and consent

The study fell within the remit of service evaluation, therefore national NHS ethics approval was not required; see Appendix A. However, due to the patient identifiable nature of video data and a requirement for off-site analysis, informed consent for the specific purposes of this service evaluation study was obtained. The study was also approved by the ethical review process in the Department of Human Communication Sciences, University of Sheffield. Parents/carers were informed of the study verbally and in writing and they signed a consent form giving their approval for their child's records to be used. The Participant Information Sheet and Consent Form are provided in Appendices B and C. Two children met inclusion criteria and were recruited; they are referred to in this chapter as James and Holly. The author is the SLT participant; she is referred to in this chapter as Sam.

3.1.2.2 Data collection

Data were collected by the author from the clinical records of children's MSIVT sessions. All 30 sessions took place in children's homes and were delivered by the author. The majority of sessions were filmed using the camcorder Sony HDR-UX19, which records sound using an

external microphone. The therapist stopped and started filming during the sessions at the beginnings and ends of activities, such that one session featured a collection of videos corresponding to individual activities. At the end of every session, the therapist made a ‘therapy video’ for home practice using the camcorder’s integral mini DVD player. This was produced from the recordings made in that session. During data collection, all video recordings generated in the MSIVT sessions were identified and copied to an external hard-drive secured with advanced encryption hardware. They were organised into two folders, one for each child participant. Each participant’s videos were organised in session and date order. Videos were named X.Y, where X depicts the session number and Y depicts the video number, corresponding to an individual activity within that session. Videos were reviewed to assess their suitability for inclusion in the study. A video was excluded if the recorded activity was not an intervention activity, (e.g. if it was assessment or advice giving, or if a therapy task was abandoned).

3.1.3 Materials

The study used data generated from the clinical records of the two child participants, James and Holly. Data consisted of video recordings of MSIVT sessions and accompanying case notes. Sessions took place between January and December 2012 and the children were 1;8 and 1;11 at the onset of therapy and both aged 2;7 at completion. Data were available on every session. The study collected 135 video recordings from 30 MSIVT sessions and 99 of these were included; see Table 3.1 for a summary of the data.

Table 3.1: Summary of the data

	Sam-James	Sam-Holly	Totals
Number of MSIVT sessions	14	16	30
Month and year of first and last session	March to December 2012	January to December 2012	January to December 2012
Age of child at first and last session	1;11 to 2;7	1;8 to 2;7	1;8 to 2;7
Number of videos collected	63	72	135
Number of videos excluded	18	18	36
Number of videos in the study	45	54	99
Mean duration of videos (range)	6:20 (1:59-15:51)	6:37 (2:50-12:16)	6:26 (1:59-15:51)

Note. Duration given in the format M:S, where M=minutes and S=seconds.

3.1.4 Participants

3.1.4.1 Therapist participant

The SLT participant is the author, Sam. At the time of providing the therapy, Sam had been working as a specialist SLT in a regional cleft lip and palate centre for eight years. She had received training on multisensory input therapy with/without video-therapy at local, national and postgraduate levels. She also had a special interest in this particular therapy approach and had studied it as part of her MSc in the three years prior to this study.

3.1.4.2 Child participants

Children were eligible for inclusion if they met the following criteria:

- Monolingual English
- Born with a cleft palate
- Underwent primary palate surgery in line with regional protocols
- Had a specialist 18 month SLT assessment
- Had no diagnosed or suspected developmental delay
- Had a diagnosis of cleft palate-related speech sound disorder (CPSSD)
- Received speech and language therapy between 18 months and three years of age targeting speech, and the therapy approach was MSIVT
- Received therapy from the author in her clinical role in the period after the Calladine (2009) study
- Received a complete episode of care, defined as at least six consecutive sessions

Participant information was collected from the children's NHS clinical records at the regional cleft lip and palate centre and, where appropriate, from local hospitals involved in their care. See Table 3.2 for a summary of the information.

Table 3.2: Characteristics of child participants

	James	Holly
Cleft diagnosis	Unilateral cleft lip and palate	Cleft palate and Robin Sequence
Age at palate repair	0;10	1;0
Healing	Complicated by soft palate fistula	Uncomplicated (no fistula)
Age at 18-month assessment	1;7	1;6
Receptive language	Age-appropriate	Age-appropriate
Expressive language	Mildly delayed	Mildly delayed
Resonance and airflow	Unable to rate	Normal
Phonetic inventory	(m n ?)	m p b k g w l
Cleft speech characteristics	(Absent pressure consonants; glottal articulation)	Backing to velar
Age at first therapy session	1;11	1;8
Age at last therapy session	2;7	2;7
Number of therapy sessions	14	16
Middle ear history	OME from 0;8-2;10	OME at 0;10; no OME at 3;2
Hearing levels	Mild bilateral loss at 0;8-0;10; moderate bilateral loss at 2;0-3;10	Mild bilateral loss at 0;10; normal levels at 3;2
Hearing management	Bilateral hearing aids at 1;11	Surveillance

Note. Ages given in format Y;M, where Y=years and M=complete months. Healing information documented at post-operative surgical review at ages 1;0 (James) and 1;4 (Holly). Speech and language information (receptive language, expressive language, resonance and airflow, phonetic inventory, cleft speech characteristics) documented at 18-month assessment at ages 1;7 (James) and 1;6 (Holly). James' resonance and airflow was 'unable to rate' because he did not vocalise during the assessment. Sounds have been transcribed using IPA (2018) and ExtIPA (Ball, Howard & Miller, 2018). OME=otitis media with effusion. Information in round brackets denotes parental report.

Both children were seen for their 18-month SLT assessment by the author in her role as specialist SLT. The regional protocol involved informal assessment during play, observation and parent report. Speech development was assessed from vocalisations elicited during play. These were transcribed live using the International Phonetic Alphabet (IPA; 2018) and Extended IPA (Ball et al., 2018). Transcribed productions were analysed phonetically and phonologically with reference to the cleft speech literature (Harding & Grunwell, 1996) and the Great Ormond Street Speech Assessment (GOS.SP.ASS.; Sell, Harding-Bell & Grunwell, 1994; 1999) framework. This included ratings of resonance and airflow. The author's postgraduate qualifications and specialist clinical experience made her qualified to undertake this transcription.

James

James had a diagnosis of unilateral cleft lip and palate and had palate surgery at 10 months of age. He experienced some wound healing complications resulting in a soft palate fistula. This was monitored during therapy for its potential impact on James' speech development. James was 19 months of age at the time of his 18-month assessment and he did not vocalise during the initial appointment. His parents described a limited inventory of sounds consisting of [m n ?]. James had a diagnosis of CPSSD characterised by absent pressure consonants suspected to be related to underlying impairments in motor execution and motor programming (Calladine & Vance, 2019). He received 14 MSIVT sessions between the ages of 1;11 and 2;7. At least one parent was present in every session.

James was diagnosed with OME and mild bilateral conductive hearing loss at 0;8, and these persisted or had reoccurred at subsequent testing at 1;6, 1;7 and 1;10. A further four recordings of OME were made at 2;0, 2;5, 2;10 and 3;10. From the age of 2;0 up to 3;10, he was diagnosed with moderate bilateral conductive hearing loss; this covers the full length of his MSIVT episode of care. James was given bilateral hearing aids at 1;11, one week before he started therapy, and wore these throughout the therapy period.

Holly

Holly had a diagnosis of cleft palate and Robin Sequence and had palate surgery at 12 months of age. Holly was 18 months at the time of her 18-month assessment. Her phonetic inventory consisted of the following sounds: [m p b k g w l]. Phonological analysis of the utterances heard at assessment revealed the cleft speech pattern of backing to velar affecting phonemes /t d/ and the typical developmental process of voicing. Holly's resonance was rated as normal and her [p b] realisations were evaluated as having adequate intra-oral pressure. It was hypothesised that Holly had adequate velopharyngeal function for speech production. She had a diagnosis of CPSSD characterised by backing to velar, suspected to be related to an underlying impairment in motor programming. She had 16 MSIVT sessions between the ages of 1;8 and 2;7. Louise's grandmother, who was her legal guardian, was present in every session.

At 0;10, eight months before starting MSIVT, Holly was diagnosed with OME and suspected mild bilateral conductive hearing loss. Holly's next follow-up was at 3;2, five months after completing MSIVT. At 3;2, she did not have OME and her hearing levels were within normal limits.

3.1.5 Data analysis

The therapy data were analysed in three ways:

1. Descriptive analysis of the MSIVT episodes of care, including speech sounds targeted and stimuli used in MSIVT activities
2. Descriptive analysis of the behaviours used by the therapist during MSIVT activities and the ways the child responds
3. Quantitative sequential analysis of therapist behaviours and child responses

3.1.5.1 Descriptive analysis of MSIVT episodes of care

Analysis began with the production of metadata using the therapy data collected on the two therapist-child dyads. This provided a means of presenting the data in such a way as to facilitate within-session and session-by-session analysis. The process involved watching video recordings in consecutive order, reviewing written information about sessions and gradually constructing the metadata with relevant information.

Nature of target selection and stimulus design

A speech sound target is a phoneme of the child's native linguistic system that is targeted in therapy. Speech sound targets were identified from the written session information and video observation. These were analysed to identify which phonemes were targeted in each activity across the episodes of care. A speech sound stimulus is a phone(s) produced by the SLT as a stimulus for the target phoneme. Stimuli orientate to phoneme targets. Harding and Bryan (2000) and Calladine (2009) describe the use of novel sounds with adaptive articulation, e.g. interdental stimuli for target alveolar phonemes. Stimuli were identified from video observation and transcribed phonetically using the International Phonetic Alphabet (2018) and Extended IPA (ExtIPA; Ball et al., 2018). Where more than one stimulus was used for the same target phoneme, all of the different productions were transcribed.

As well as the number and type of speech sound targets and stimuli, data were analysed to identify the linguistic level, i.e. the verbal context in which the stimulus was produced and the target position it orientated to. This is important because of the implications that different levels have for the child's speech processing. Linguistic level was identified from written information and video observation. It was documented according to the following descriptions in the existing literature: single sound; sound sequence; invented word; and real word; see definitions below (Calladine, 2010; Calladine & Vance, 2019; Gierut, Morrisette & Ziemer, 2010; Harding & Bryan,

2000; Treslove, 2014; Williams et al., 2010). At every level, the stimulus can occur as part of a short phrase or sentence.

- **Single sound:** A phone or phoneme produced in isolation
- **Sound sequence:** A consonant (C) and vowel (V) combination, where the stimulus may occur in any position, e.g. CV, VC, VCV
- **Invented word:** A sound sequence with meaning assigned, e.g. the name for a character, or a word that does not exist in the child's language system
- **Real word:** A word that exists in the child's language system and is understood to be known to the child; the stimulus may occur in any position, e.g. initial (WI), medial (WM) or final (WF)

Child stimulus productions

The videos were analysed to determine the number of activities in which the child participants made productions and the nature of their productions, i.e. which sounds and at what linguistic levels.

3.1.5.2 Descriptive analysis of therapist behaviours and child responses

Sampling the data

A subset of data was used to trial a method for identifying the types of behaviours the therapist used during therapy and the ways children responded. Five videos (5% of the dataset) were randomly sampled using the random number generator at <https://www.random.org>. The names of the videos selected are: 14.2; 10.7; 12.2; and 4.1 from Sam-James and 1.2 from Sam-Holly. The first two minutes of every video were extracted to create a sample of 10 minutes of video of MSIVT activities.

Nature of therapist behaviours

A therapist behaviour is a verbal or nonverbal behaviour by the therapist that relates to the speech sound stimulus, is directed toward the child directly or indirectly via the camera, and is interpreted as an attempt to provide the child with input stimulation in accordance with the MSIVT approach. In her study of phonology therapy, McCartney (1989) used the term 'speech teaching strategies' to refer to the behaviours therapists used during therapy. They include 'model', which is where the therapist produces the target sound (for the child to imitate), and 'metalinguistic utterance', an utterance that describes target articulation features of speech sounds (to support the child to revise their production).

The 10 minutes of video were analysed to identify all of the therapist behaviours that occurred. A list of qualitative descriptions was produced and analysed using a method similar to thematic analysis (Braun & Clarke, 2006) to identify categories of behaviours that display similar characteristics. Descriptive codes were generated and assigned to the categories.

Nature of child responses

A child response is a verbal or nonverbal behaviour by the child that relates to the speech sound stimulus and occurs in response to a therapist behaviour. The 10 minutes of video were analysed to identify all of the child responses that occurred. A list of qualitative descriptions was produced and analysed using a method similar to thematic analysis to identify categories of responses that display similar characteristics. Descriptive codes were generated and assigned to the categories.

3.1.5.3 Quantitative analysis of therapist behaviours and child responses

Sampling the data

A subset of data was used to trial a method of quantitative sequential analysis for examining therapist behaviours and child responses. To simplify the data, only videos featuring speech sound stimuli at single sound level were included. One sample was produced for each therapist-child dyad using the following steps:

1. One target phoneme was randomly selected (using <https://www.random.org>)
2. The metadata were examined to identify all videos featuring that target
3. A 30% sample of the videos identified in step 2 was randomly selected (using <https://www.random.org>)

Sampling selected the /b/ phoneme and videos 6.2, 10.6, 10.7, and 12.2 from the Sam-James dyad and the /t/ phoneme and videos 1.2, 5.2, 5.3, and 6.1 from the Sam-Holly dyad.

Method of sequential analysis

A quantitative method was piloted on the data to inform method development for the main study. Of particular interest was exploring ways that therapist behaviours and child responses could be analysed and the potential insights this might provide into what MSIVT looks like as a dynamic therapy approach. Very little research has examined verbal and nonverbal behaviours in speech and language therapy, so the study drew on research from outside of the field. Of particular influence were the study by Wiseman and Rice (1989) and the paper by Elliott (2010). An important methodological feature of studies that use sequential analysis is a clear definition of turn boundary and sequence. In the present study, a therapist turn could consist of one or

more coded behaviours and the end of the turn was the point at which either a) the child responds or b) there is a time lag of more than one second before the therapist displays another behaviour, as this would mark the start of a new turn. A child turn could also consist of one or more coded responses and the end of the turn was the point at which either a) the therapist displays a new behaviour or b) there is a lag of more than one second before the child displays another response. A sequence comprises one complete therapist turn and one complete child turn. Therapist turns that did not elicit a response were excluded from the analysis.

The video samples were viewed three times. Sequences were identified and the start and finish times of turns were noted. Therapist turns were coded to indicate the presence or absence of two categories of behaviour that emerged from the analysis, *Active involvement* and *Specific description*. These behaviours were chosen because of their high frequency of occurrence in the therapy data. Child turns were coded to indicate whether or not they included a *Production* of the stimulus. This category of child response emerged from the analysis and was chosen because the therapy environment in which child stimulus productions occur during MSIVT is of particular interest. Two-by-two tables were created with these two different categories of therapist behaviour and child production response, as shown below:

- *Active involvement* (Yes/No) and *Production* (Yes/No)
- *Specific description* (Yes/No) and *Production* (Yes/No)

The tables were then populated with frequency of occurrence data. To see if a relationship exists between the two variables, statistical analysis was undertaken using the Pearson's chi-square test and IBM SPSS Statistics 21 software. This test compares the frequencies observed in the different categories to the frequencies that would be expected by chance (Field, 2009).

3.2 Results

The findings have been generated from the analysis of 99 MSIVT activities. The mean duration of activities was six minutes and 26 seconds, with a range of 1:59 to 15:51.

3.2.1 Descriptive analysis of MSIVT episodes of care

3.2.1.1 Nature of target selection and stimulus design

Table 3.3 summarises the findings from within-dyad analysis. It shows the phonemes targeted during the two episodes of care and the number of activities in which they were targeted. It also shows the speech sound stimuli produced by the therapist for the different phonemes and the linguistic levels at which they were produced. For example, in the episode of care featuring Sam-

James, phoneme /b/ was targeted with stimulus [b̥] at single sound, invented word and real word levels.

Table 3.3: Speech sound targets, stimuli and linguistic levels

Sam-James							
Total number of sessions = 14; Total number of activities = 45							
Target phoneme	Stimulus	Number (and %) of sessions	Number (and %) of activities	Linguistic levels			
				Single sound	Sound sequence	Invented word	Real word
/p/	p̥	11 (79)	20 (44)	Y	Y	Y	Y
/b/	b̥	10 (71)	21 (47)	Y	NA	Y	Y
/t/	t̪	4 (29)	5 (11)	Y	NA	NA	NA
/s/	ʃ̪:	2 (14)	2 (4)	Y	NA	NA	NA
Sam-Holly							
Total number of sessions = 16; Total number of activities = 54							
/t/	t̪	15 (94)	42 (78)	Y	Y	Y	NA
/d/	d̪	9 (56)	13 (24)	Y	Y	Y	NA
/n/	n̪	1 (6)	1 (2)	Y	NA	NA	NA
/s/	ʃ̪:	3 (19)	8 (15)	Y	NA	Y	NA

Note. Sounds have been transcribed using IPA (2018) and ExtIPA (Ball et al., 2018). Y=yes. NA=not applicable because this level was not targeted in therapy.

James' MSIVT episode of care targeted four phonemes, two bilabial plosives /p b/ and two alveolar phonemes, one plosive /t/ and one fricative /s/. They were targeted with four corresponding stimuli. Therapist Sam used modified articulation in her stimulus productions: weakened articulation for the bilabial plosives; interdental stimuli for both alveolar phonemes; and lengthened articulation for the fricative. The stimuli [p̥ b̥], which share the same place of articulation, featured most frequently, in 79% (11/14) and 71% (10/14) of sessions and 44% (20/45) and 47% (21/45) of activities, respectively. They also featured at the widest range of levels: single sound; invented word; and real word; and [p̥] featured additionally at sound sequence level. In contrast, [t̪ ʃ̪:] featured in fewer sessions (29% and 14%, respectively) and fewer activities (11% and 4%, respectively) and at single sound level only. Closer analysis of the metadata revealed that some of James' activities featured multiple speech sound stimuli, e.g. Session 11-Activity 2 featured [p̥ b̥] both at invented word level, and some activities featured stimuli at more than one linguistic level, e.g. Session 1-Activity 3 featured [p̥] at single sound and sound sequence levels.

Holly's episode of care also targeted four phonemes /t d n s/, which share the same alveolar place of articulation. Sam's stimulus productions for the phonemes /t d/ featured a combination of interdental and linguolabial articulations. She also used interdental articulation for phonemes

/n s/. The stimuli [t̥ t̥] featured prominently, in all but one of the 16 sessions and in three quarters (78%; 42/54) of all activities, followed by the voiced stimuli [d̥ d̥], which featured in more than half (56%; 9/16) of sessions and a quarter (24%; 13/54) of all activities. These stimuli were produced at single sound, sound sequence and invented word levels. No stimuli featured at real word level.

3.2.1.2 Child stimulus productions

Both children made stimulus productions during therapy in response to the stimuli presented in activities. Table 3.4 summarises the number of sessions and activities that featured productions and the levels at which the children made productions for their individual target stimuli.

Table 3.4: Summary of child stimulus productions

Sam-James						
Total number of sessions = 14; Total number of activities = 45						
Number (and %) of sessions	Number (and %) of activities	Target stimuli	Linguistic levels			
			Single sound	Sound sequence	Invented word	Real word
13 (93)	23 (51)	p̥	Y	N	Y	Y
		b̥	Y	NA	Y	Y
		t̥	Y	NA	NA	NA
		ʒ̥:	Y	NA	NA	NA
Sam-Holly						
Total number of sessions = 16; Total number of activities = 54						
16 (100)	49 (91)	t̥ t̥	Y	Y	Y	NA
		d̥ d̥	Y	Y	Y	NA
		n̥	Y	NA	NA	NA
		ʒ̥:	Y	NA	Y	NA

Note. Sounds have been transcribed using IPA (2018) and ExtIPA (Ball et al., 2018). Y=yes. N=no. NA=not applicable because this level was not targeted in therapy.

As shown, James made productions in nearly all (93%; 13/14) of his sessions and in about half of all activities. He made productions for all of his targets and at nearly all of the levels targeted. The only target stimulus he did not produce was [p̥] at sound sequence level. Holly made productions in all (100; 16/16) of her sessions and in nearly all (91%; 49/54) of activities. She made stimulus productions for all of her targets at all the levels targeted.

3.2.2 Descriptive analysis of therapist behaviours and child responses

3.2.2.1 Nature and frequency distribution of therapist behaviours

The analysis produced 113 qualitative descriptions of therapist behaviours in the 10 minutes of video. Further analysis led to the development of eight themed categories. Table 3.5 presents the eight categories of behaviour and shows the number of times they occurred in the five sampled videos.

Table 3.5: Therapist behaviour categories: Codes, descriptions and frequency of occurrence

Category code	Category description	Video					Totals
		14.2	10.7	12.2	1.2	4.1	
Model	Therapist produces the target stimulus	7	2	5	10	5	29
Label	Therapist verbally labels the stimulus	1	4	8	4	6	23
Specific description	Therapist describes the specific articulatory or phonetic features of the stimulus	4	2	2	12	1	21
Representation	Therapist uses a picture to represent the stimulus	7	1	2	8	6	24
Active involvement	Therapist verbally, or nonverbally through gesture, asks or tells the child to do something related to the stimulus	13	4	4	8	6	35
Gesture	Therapist makes a gesture relating to the stimulus	5	2	6	18	0	31
Praise	Therapist gives the child verbal praise about a stimulus production	3	0	0	1	1	5
Tactile	Therapist gives the child tactile stimulation from stimulus production	0	0	0	0	2	2
Total occurrences		40	15	27	61	27	170

Note. Video names given in the format X.Y where X=Session number and Y=activity number. The numbers in the video columns indicate how many times the behaviour occurred in the video. 0=the behaviour did not occur in the video.

The analysis found that Sam used a combination of verbal and nonverbal behaviours during the MSIVT activities. For example, behaviours in the categories *Model* and *Label* are verbal, and *Representation* and *Gesture* are nonverbal. Behaviours in the category *Active involvement* were either verbal or nonverbal. Some behaviours involved a direct reference to the child, e.g. *Praise*, whereas others referred to the stimulus or its representation, e.g. *Label*. There was some variation within categories. For instance, *Specific description* sometimes referred to Sam's own stimulus production, e.g. "with my lips" (Sam-James; Video 12.2), and at other times referred to the child's production, e.g. "but you used your lips" (Sam-James; Video 14.2). One of the distinguishing features was the nature of the pronoun in Sam's utterances. When describing her own stimulus model she showed a preference for a pronoun that referred to her, e.g. 'my'. Yet, when responding to a child stimulus production, she did not show a strong preference for a pronoun referring to the child, such as 'your'; she would also respond to child productions with a pronoun referring to her instead of the child.

Sam did not display all categories of behaviour in every video. Those displayed most frequently were *Active involvement* (21%; 35/170), *Gesture* (18%; 31/170) and *Model* (17%; 29/170). Those displayed least frequently were *Tactile* (1%; 2/170) and *Praise* (3%; 5/170). *Model*, *Label*, *Specific description* and *Active involvement* behaviours featured in all five videos. These behaviours involved Sam producing, describing or using a picture to represent the speech sound stimulus. *Tactile* behaviours are similar in nature, though they only featured in one of the videos, 4.1.

3.2.2.2 Nature and frequency distribution of child responses

The analysis produced 63 qualitative descriptions of child responses in the 10 minutes of video. Further analysis led to the development of four themed categories. Table 3.6 presents the four categories of response and shows the number of times they occurred in the five sampled videos.

Table 3.5: Child response categories: Codes, description and frequency of occurrence

Code	Description	Video					Totals
		14.2	10.7	12.2	1.2	4.1	
<i>Production</i>	Child makes a stimulus production	2	NA	NA	12	1	15
<i>Manipulation</i>	Child does something with the picture representation of the stimulus	18	3	6	1	6	34
<i>Gesture</i>	Child makes a gesture relating to the stimulus	NA	NA	NA	3	2	5
<i>Label</i>	Child produces a verbal label relating to the stimulus	NA	NA	NA	1	NA	1
		20	3	6	17	9	55

Note. Video names given in the format X.Y where X=session number and Y=activity number. The numbers in the video columns indicate how many times the response occurred in the video. NA=not applicable because the response did not occur.

The analysis found the children made verbal and nonverbal response during the MSIVT activities. For example, *Production* and *Label* are verbal and *Manipulation* and *Gesture* are nonverbal. The most frequently exhibited response was *Manipulation*, which involved the child doing something with the picture used to represent the sound stimulus, e.g. holds the picture facing the camera; it made up 34/55 (62%) of all the responses observed. The second most frequently exhibited response was *Production*, which involved the child making a stimulus production; this represented 15/55 (27%) of all responses observed. Responses belonging to the categories *Gesture* and *Label* occurred infrequently.

3.2.3 Quantitative analysis of therapist behaviours and child responses

3.2.3.1 Sam-James

The analysis identified 48 complete behaviour-response sequences in the sampled videos. The frequency data is shown in Tables 3.6 and 3.7 and the results of Pearson chi-square analysis are presented in Table 3.8.

Table 3.6: Frequency data for *Active involvement* and *Production* (Sam-James)

<i>Active involvement</i>	<i>Production</i>		Totals
	Yes	No	
Yes	4	24	28
No	6	14	20
Totals	10	38	48

Note. Yes=behaviour/response present. No=behaviour/response absent.

Table 3.7 Frequency data for *Specific description* and *Production* (Sam-James)

<i>Specific description</i>	<i>Production</i>		Totals
	Yes	No	
Yes	8	9	17
No	2	29	31
Totals	10	38	48

Note. Yes=behaviour/response present. No=behaviour/response absent.

Table 3.8 Chi-square analysis (Sam-James)

Relation	Value	Df	Asymptotic significance (2-sided)
<i>Active involvement X Production</i>	1.747	1	0.186
<i>Specific description X Production</i>	10.977	1	0.001*

Note. * indicates statistical significance ($p < 0.05$). df=degrees of freedom.

Table 3.6 shows that sequences in which Sam used *Active involvement* did not typically involve a *Production* response by James. Sequences in which Sam did not use *Active involvement* were also more likely not to involve a *Production* response. However, Table 3.8 shows that analysis of these variables using Pearson's chi-square test did not find a relationship between them. That is, the presence/absence of *Production* in sequences that did/did not include *Active involvement* was no more or less frequent than would be expected by chance alone. The analysis did, however, find a statistically significant relationship between the variables *Specific description* and *Production*. When Sam did not use *Specific description*, James was statistically significantly unlikely to make a *Production*; see Table 3.8.

3.2.3.2 Sam-Holly

The analysis identified 65 complete behaviour-response sequences in the sampled videos. The frequency data is shown in Tables 3.9 and 3.10 and the results of Pearson chi-square analysis are presented in Table 3.11.

Table 3.9: Frequency data for *Active involvement* and *Production* (Sam-Holly)

<i>Active involvement</i>	<i>Production</i>		Totals
	Yes	No	
Yes	14	14	28
No	18	19	37
Totals	32	33	65

Note. Yes=behaviour/response present. No=behaviour/response absent.

Table 3.10 Frequency data for *Specific description* and *Production* (Sam-Holly)

<i>Specific description</i>	<i>Production</i>		Totals
	Yes	No	
Yes	15	6	21
No	15	29	44
Totals	30	35	65

Note. Yes=behaviour/response present. No=behaviour/response absent.

Table 3.11 Chi-square analysis (Sam-Holly)

Relation	Value	Df	Asymptotic significance (2-sided)
<i>Active involvement</i> X <i>Production</i>	0.12	1	0.914
<i>Specific description</i> X <i>Production</i>	7.974	1	0.005*

Note. * indicates statistical significance ($p < 0.05$). df=degrees of freedom.

Table 3.9 shows that sequences in which Sam used *Active involvement* sometimes did and sometimes did not include a *Production* response from Holly. Table 3.11 shows that analysis using Pearson's chi-square test found no relationship between these variables. The analysis did, however, find a statistically significant relationship between the variables *Specific description* and *Production*, indicating their occurrence was more than would be expected by chance alone. When Sam did not use *Specific description* Holly was statistically significantly unlikely to make a *Production*; when Sam did use this behaviour Holly was likely to make a *Production*.

3.2.3.3 Methodological considerations

The therapist behaviours *Active involvement* and *Specific description* were analysed irrespective of whether they occurred as single behaviours in a therapist turn or as a composition of multiple behaviours. Inspection of the turn timings revealed that some turns were longer in duration than others. Where the behaviour occurred as a composition, its position in the turn was not factored into the analysis. This contextual information might be

important. The combination and/or turn position might inform the child’s response more than the isolated behaviours under study. Table 3.12 provides some examples of therapist turns that included *Active involvement* to illustrate some of these contextual issues.

Table 3.12: Examples of turns featuring *Active involvement*

Video name	Turn sequence	Duration of turn (in seconds)	Therapist behaviours (in order of appearance)
6.1	1	3	<i>Representation</i> <i>Label</i> <i>Active involvement</i>
6.1	2	16	<i>Active involvement</i> <i>Label</i> <i>Model</i> <i>Specific description</i>
12.2	1	3	<i>Active involvement</i> <i>Label</i> <i>Representation</i> <i>Gesture</i> <i>Model</i> <i>Active involvement</i>
12.2	7	2	<i>Label</i> <i>Representation</i> <i>Active involvement</i>

As shown in these examples, the behaviour of interest, *Active involvement*, appeared in different positions and alongside up to four other behaviours. Composite turns, featuring more than one behaviour, were more typical of the data than turns featuring single behaviours. Although this analysis highlights limitations of the method, revealing the composite nature of Sam’s turns was an important finding.

3.3 Discussion

This pilot study aimed to examine the author’s own clinical practice of multisensory input video-therapy (MSIVT) in order to consider how to analyse the recorded therapy data. It examined data on 30 therapy sessions delivered as part of two episodes of care. The sessions featured therapist-child dyads Sam-James and Sam-Holly. The study trialled a combination of

methods to analyse and describe what happened during therapy. It looked specifically at speech sound targets and stimuli, therapist behaviours and child responses. The findings informed the analytical methods used in the main study.

3.3.1 Nature of target selection and stimulus design

Both children had multiple phoneme targets. These featured some common class relationships. For example, all four of Holly's targets were alveolar phonemes. This reflects the specific nature of her speech diagnosis, cleft palate-related speech sound disorder (CPSSD) characterised by the backing to velar process affecting alveolar phonemes. Targeting all phonemes affected by a phonological process is consistent with the phonological and cleft speech literature (Harding-Bell & Howard, 2011). James started in therapy with a less specific diagnosis; CPSSD characterised by absent pressure consonants. The variety of voiced, voiceless, bilabial and alveolar targets in his therapy might reflect the absence of specific substitution processes. Calladine and Vance (2019) present a more detailed and psycholinguistic analysis of James' (and Holly's) speech profiles and response to therapy, which includes rationale for his target selection. Although she targeted multiple phonemes, in both dyads Sam showed a preference for activities orientating to just two of the phonemes, /p b/ in Sam-James and /t d/ in Sam-Holly.

Each of James' phonemes were targeted with one sound stimulus and all four stimuli featured adaptive articulations. This included weakened pressure, which is a recognised principle of cleft speech intervention, especially in the presence of suspected or confirmed velopharyngeal insufficiency (Harding & Bryan, 2000; Russell & Albery, 2005). Harding and Bryan (2000) and Calladine and Vance (2019) describe lengthened fricatives and the use of interdental and linguolabial placements to make sounds more acoustically and visually salient, as well as to maximise the likelihood of them being processed as novel. Sam used these specific articulations in both dyads. An interesting finding from Sam-Holly was that Sam used a combination of interdental and linguolabial stimuli for each of the phonemes /t d/. The use of multiple stimuli for the same phoneme has not previously been described in the literature in relation to the MSIVT approach.

In both dyads, Sam targeted all phonemes with stimuli at single sound level. The most frequently targeted stimuli occurred at a variety of levels, e.g. James' /p/ stimulus featured at single sound, sound sequence, invented word and real word levels, and Holly's /t d/ stimuli featured at single sound, sound sequence and invented word levels. The use of multiple levels

is consistent with existing descriptions of the MSIVT approach (Calladine, 2009; Calladine & Vance, 2019; Harding & Bryan, 2000; Treslove, 2014) and makes it different to other approaches for young children with CPSSD, such as EMT+PE (Kaiser et al. 2017). Naturalistic approaches, like EMT+PE, advocate the use of meaningful real word stimuli rather than single sounds. In MSIVT, the use of single sounds relates to the underlying psycholinguistic basis of the approach and reflects the goal of therapy to support the development of new and accurate internal templates from a bottom-up direction.

3.3.2 Child stimulus productions

Both children demonstrated high levels of stimulus production during therapy and this included all of their stimuli at single sound level. Holly made productions in every (100%; 16/16) session, nearly every (91%; 49/54) activity, for every one of her four targets and at every one of the three levels targeted. The finding that child stimulus productions can occur during therapy using an input-based approach reinforces the previous findings by Calladine (2009).

3.3.3 Profiling therapist behaviours and child responses in MSIVT

The analysis generated eight categories of therapist behaviour, which include a range of verbal and nonverbal behaviours. Some behaviours are described in the existing literature on MSIVT, whereas others are not. *Model*, which consists of a stimulus production, featured highly. This is consistent with existing descriptions of the MSIVT approach (Calladine & Vance, 2019; Harding & Bryan, 2000; Treslove, 2014). It offers the child auditory and visual information about speech sound stimuli, which has potential to stimulate input and output processing (Stackhouse & Wells, 1997). *Active involvement* and *Gesture* also featured highly. *Gesture* involves the therapist making a gesture relating to the stimulus, e.g. points to tongue, moves finger around lips. The original description of MSIVT by Harding and Bryan (2000) does not refer to this specific behaviour, though it does refer to highlighting visual features of target stimuli, and *Gesture* is a way of providing visual information. The use of finger gestures, such as cued articulation (Passy, 1993) is described in the speech intervention literature. The prevalence of *Active involvement* behaviours, which involve the therapist asking or telling the child to do something relating to the stimulus, was surprising. Neither Harding and Bryan (2000) nor Calladine (2009) describe this type of active instruction. In line with the principles of the MSIVT approach, none of the *Active involvement* behaviours involved a request for stimulus production. Mostly, they involved asking the child to do something with the stimulus picture, e.g. show it to the telly.

The analysis found *Tactile* behaviours were used infrequently. This was a surprising finding given the principle of the approach is to provide multisensory stimulation (Harding & Bryan, 2000). However, this may be related to the nature of target stimuli in the sampled videos. *Tactile* was only observed in one video, which featured James. The stimulus was the voiceless plosive [t̥]. In three of the remaining four videos, the stimulus was the voiced plosive [b]. Voiced plosives do not provide the same opportunities for tactile stimulation due to their lack of aspiration and friction. However, *Tactile* behaviours were similarly not seen in one of Holly's videos, which also targeted [t̥]. This suggests possible variation relating to specific characteristics of the child. The participant histories show that James had hearing loss whereas Holly did not, therefore it may have been used as a strategy for James to help compensate for his loss of auditory perception.

An important finding of the analysis was Sam's display of *Specific description* behaviours, which she used to refer to her own stimulus production as well as production by the child. This suggests they served different functions. In the author's previous study, Calladine (2009), she used such verbal descriptions in response to child stimulus productions as a way of giving feedback to the child, not alongside her own productions.

The analysis generated four categories of child response, which included verbal and nonverbal behaviours. The most frequently exhibited responses were *Manipulation* and *Production*. The finding of *Production* responses in the absence of therapist requests for production reinforced Calladine's (2009) findings and aligns with the non-directive style of MSIVT (Calladine & Vance, 2019; Harding & Bryan, 2000).

3.3.4 Associations between behaviours and responses

Quantitative analysis of behaviour-response associations yielded similar findings for both dyads. Statistical analysis found no significant association between Sam's use of *Active involvement* behaviours and *Production* responses. In other words, *Active involvement* did not appear to elicit or inhibit stimulus production. Actively involving children in therapy is a recommended principle of output-based therapy approaches (Weiss, 2004), but it has not previously been reported in relation to the MSIVT approach. In contrast, Sam's use of *Specific description* behaviours did show an association with *Production* responses. In both dyads, sequences that did not display this behaviour were significantly unlikely to contain a *Production* response. In sequences where Sam did display this behaviour, Holly's response was more likely to include a *Production* than not. Identifying therapist behaviours that may or may

not elicit a *Production* response was not an objective of this study, rather it was the process of trialling different methods for analysing the data. The analysis revealed that Sam typically displayed a composition of behaviours in her turns, and some behaviours appeared to serve more than one function.

3.4 Conclusions and implications for the main study

The pilot study has shown that descriptive analytic methods can be used successfully to identify and describe features of MSIVT delivery, such as what speech sounds are targeted and what stimuli are used. This is extended in the main study to also track the levels at which stimuli were targeted throughout the course of therapy. This is important for understanding how the delivery of MSIVT reflects the underlying theoretical basis of the approach and how therapists are implementing this in practice. The pilot study trialled qualitative and quantitative methods for profiling and examining therapist behaviours and child responses during MSIVT as a way of developing understanding about the dynamic nature of MSIVT and how it is described. This is informative in that it has revealed the types of verbal and nonverbal behaviours that the author, as therapist, used in these MSIVT sessions and their alignment with existing descriptions of the approach. It has also revealed the verbal and nonverbal nature of responses made by the two children. The findings provide further empirical evidence that two-year-old children engage in stimulus production during MSIVT, thus extending the findings from Calladine (2009).

However, two key findings from the analysis of therapist behaviours and child responses led to a different focus in the main study. Firstly, the therapist often displayed a multitude of behaviours in one turn and, similarly to the metalinguistic utterances in McCartney's (1989) study, some of these behaviours served more than one function. Secondly, whilst both children made stimulus productions, some of their nonverbal responses, such as displaying a picture to the camera, appeared to be an important part of what was happening in the session. These findings support a different focus of analysis in the main study. They suggest that examining therapist and child interaction rather than therapist and child behaviours will reveal more about the dynamic nature of the MSIVT approach and its key components.

Chapter 4: A Literature Review of Clinical and Non-Clinical Interaction

This chapter is organised in two sections. Section 1 reviews the literature on clinical interaction in speech and language therapy (SLT) and Section 2 reviews some of the non-clinical literature.

4.1 Section 1: Clinical interaction in SLT

4.1.1 Introduction

Of particular interest in relation to the interactional features of MSIVT are: the ways in which the therapist establishes a child's attention; how the therapist creates an environment that exposes the child to high levels of multisensory sound input and opportunities to engage in production in a non-directive manner; and how the child responds to this in the local context. To inform this, the literature review was extended to include the structure, pattern and nature of therapist-child interaction during speech and language therapy.

Multisensory input therapy is an early intervention for children with speech output difficulties who are at an age where traditional output approaches are inappropriate. It is therapist-led but the authors describe an interaction style that is non-directive and nurturing, which is typically associated with child-led intervention. The goal of therapy is help the child reach their potential in developing optimally intelligible speech. On this basis, this review focuses primarily on the literature on therapist-child interaction during therapist-led intervention for children with speech difficulties, but also draws on the literature on child-led intervention. The review is expanded in places to deal with limitations in the speech intervention literature.

4.1.2 Therapist-child interaction in SLT: An overview

Central to the literature on therapist-child interaction in speech and language therapy are studies of discourse; what the therapist and the child say and do during the therapy session. These typically involve large sets of discourse data from relatively small numbers of participants, i.e. small n studies. Only one of the studies identified in this review involved children under three years of age (Norris & Hoffman, 1990) and all the children in this study were pre-verbal. Neither the children's communication profiles nor the intervention provided resemble the participants and intervention in the present study. All of the studies reviewed examined discourse data cross-sectionally, either from single or multiple sessions; none of the

studies explored therapy interaction from a longitudinal perspective.

The earliest studies of speech and language therapy discourse involving children emerged in the late 1970s and studied intervention in the broad sense. For example, Prutting, Bagshaw, Goldstein, Juskowitz and Umen (1978) examined interaction in language intervention, Ripich, Hambrecht, Panagos and Prelock (1984) compared discourse patterns in language and articulation intervention, and Letts (1985) examined sessions involving children (and adults) with a range of speech, language and communication needs. Many of these early studies used discourse analysis (DA) as the analytic method, whereby therapy discourse was examined for the occurrence of certain types of speech acts or taxonomies. Such analyses produced descriptive accounts of the characteristics of therapy interaction. Since then, studies have narrowed their focus to specific types of intervention and more micro levels of interaction. Gardner (1994; 1997; 1998) studied interaction during phonology therapy and specifically the ways that therapists engender imitation and repair in children with phonological disorder. She used the micro-analytic approach of conversation analysis (CA), examining discourse at the level of the turn and its consequences, not in relation to any pre-identified speech acts. In doing so, Gardner (1994; 1997) demonstrated CA's ability to uncover subtle phonetic and prosodic features that would have otherwise gone undetected. Even more importantly, such analyses provided explanatory accounts of how certain features of interactional turns explicate specific actions and contribute to the therapeutic process. Hulterstam and Nettelbladt (2002) used a similar approach, initiative response analysis, to compare therapist-child interaction during two different types of speech intervention, Metaphon (Howell & Dean, 1994) and traditional therapy, for children with phonological impairment. Most recently, Ronkainen (2011) used CA to study interaction during auditory-verbal therapy involving children with cochlear-implants at a pre-linguistic stage of development. None of the reviewed studies examined interaction involving children with cleft palate-related speech sound disorder (CPSSD).

The following themes emerge from the existing literature on interaction during speech and language therapy: the organisational structure of a therapy session and the activities or tasks that make up the main business of therapy; therapist and child roles; the pattern and nature of interaction; and the nature and design of therapist turns and consequences they have for the child. In order to illustrate the accumulation of current knowledge, I will use a chronological perspective to explicate these themes.

4.1.3 Structure of the therapy session

The literature describes therapist-led sessions as highly structured (Panagos, Bobkoff, Kovarsky & Prelock, 1988; Panagos, Bobkoff & Scott, 1986). Panagos et al. (1986) drew on findings from a range of studies to provide a descriptive account of what happens during a speech and language therapy session ('lesson'). They do not provide details of the therapists' experience, age of the children, nature of their communication impairment, or the therapy approach used. However, they describe the picture of discourse that they accumulated as "remarkably constant" (p224) and this has not been overtly challenged by the more recent studies. A session typically consists of three parts: the opening phase; the remedial phase; and the closing phase. They involve two or more 'tasks', which have a similar three-part structure: opening; remedial; closing. It is in the remedial phase where the main business of therapy takes place. It is characterised by a series of 'remedial sequences', a specific type of interactional sequence in which the therapist tries to elicit a target response from the child, and is regulated by nonverbal support (Panagos et al., 1986; 1988).

In a more recent study, Tykkyläinen (2009) examined interaction during task setting. She compared interaction between six therapists and five-year-old children with specific language impairment and six mothers with their five-year-old typically developing children. Tykkyläinen's (2009) description of a three phase structure, with the central 'task phase' comprising a chain of 'task sequences', resembles the description by Panagos et al. (1986) and, like Panagos et al. (1986; 1988), characterises therapy as a highly structured context. Therapy sessions evolve around a series of activities and it is within the activities, through a series of therapist-initiated sequences, where the therapy takes place.

4.1.4 Therapist and child roles and the therapeutic alliance

The roles of the therapist and the child during therapy reflect the type of intervention approach and the goals of therapy. In therapist-led intervention, the goal is to facilitate change in one or more of the child's linguistic forms (Fey, 1986; Kovarsky & Duchan, 1997). The therapist takes the lead in organising the session, designing the activities, setting the targets, and choosing the materials and stimuli (Letts, 1985; Panagos et al., 1986). She creates learning opportunities for the child through instruction (often direct) and evaluation (Letts, 1985; Ripich et al., 1984). The child takes on a respondent role following the therapist's lead and complying with the therapist's requests (Ripich et al., 1984). These roles have been illuminated by studies that have combined discourse analysis and quantitative procedures, e.g. Prutting et al. (1978)

and Ripich et al. (1984). Such studies examine extracts of discourse to identify the occurrence of specific types of speech act and use quantitative procedures to count and compare the relative frequencies exhibited by the therapist and the child.

Prutting et al. (1978) analysed audiotapes of language intervention sessions involving eight therapist-child pairs. The children were aged 5 to 8;6 with language impairment and intelligible speech. Four of the eight therapists had certificates in clinical speech competencies. They used discourse analysis to identify verbal utterances produced by therapist and child and coded them under the notion of the illocutionary act, that is, according to their linguistic properties representing the speaker's intention, not how the utterance was treated by the listener. Utterances were coded according to three broad types of communicative act: requests; responses; and statements. Of the total acts produced by therapists, 52% were requests and most (43%) of these were requests for verbal information. In contrast, only 9% of the acts produced by children were requests. Children predominantly made response-type acts, which represented 75% of the total acts they produced. Prutting et al.'s (1978) findings characterise the therapist as the primary initiator of interaction and the child as the respondent.

Ripich et al. (1984) used a very similar method of analysis when they compared interaction during articulation and language intervention in six therapist-child pairs. Therapists had a mean of 4.10 years' experience. The children, who had a combination of articulation and language difficulties, had a mean age of 7.67 years. Each pair provided two audiotape recordings, one during each type of intervention. Ripich et al. (1984) used the same three broad categories as Prutting et al. (1978) to code the discourse data: request; response; and statement. A statement in which a judgement was made of a previous response was identified as an 'evaluation'. The authors found only minor differences between the two types of intervention sessions. In both the articulation and language contexts, therapists interacted primarily using instruction type utterances, making frequent requests, and they made almost all (95%) of the evaluations. The children talked less than the therapists did and their role within the interaction was primarily as respondent. Ripich et al. (1984) liken the therapist and child roles to that of teacher and pupil in the classroom. These findings reinforce those from Prutting et al. (1978) and extend description of the therapist's role to include making evaluations of responses as well as requesting them.

In another early study Letts (1985) distinguished between different aspects of the therapist's role in relation to 'organising acts', used to get an activity up and running and ensure it does

not break down, and 'on-going acts', which make up the main essence of the activities. On-going acts consist primarily of eliciting acts, in which the therapist tries to elicit a response from the child. Similar to Prutting et al. (1978) and Ripich et al. (1984), Letts (1985) also identifies the therapist's role in giving feedback and information within the category of on-going acts. Letts' (1985) data consisted of audiotapes of sessions involving children (and adults) with a range of speech, language and communication needs, although she does not provide details of the therapy that took place. Nonetheless, her characterisation of the roles of therapist and child during therapy are consistent with the previously cited studies.

In their studies of therapist-child discourse during stuttering therapy, Leahy (2004) and Leahy and Walsh (2010) characterise the therapist's role as facilitator of both the 'art' and 'science' of therapy. The art of therapy represents the rapport that develops between the therapist and the child; a mutual agreement to work together to resolve problems. Leahy and Walsh (2010) identify techniques that a therapist may use to establish rapport and demonstrate that rapport building is an on-going process, not something that happens in discrete moments. They use an extract of discourse from a study by Van Riper (1977, as cited in Leahy & Walsh, 2010) to illustrate how humour can be used as a vehicle for rapport building. This is in contrast to Panagos et al. (1986) who, from reviewing studies of articulation and language intervention, state that "little laughter and smiling" (p214) takes place during therapy. The science of therapy represents the process of identifying and finding solutions to problems, and happens simultaneously. Leahy (2004) analysed audio recordings of an experienced therapist working with a 13-year-old child on a fluency technique. One feature of interaction revealed by the analysis was the therapist's use of the pronoun 'we'. Leahy suggested this signified inclusivity and warmth; a collaborative effort by the therapist and the child to work together to achieve the goals of therapy.

4.1.5 Structure and pattern of therapist-child interaction

The majority of studies of interaction during speech and language therapy relate to therapist-led intervention. They describe therapy as a highly structured exchange between the therapist and the child. Quantitative and qualitative studies have shown that roles are established and maintained through an asymmetrical pattern of interaction where the therapist takes up more of the talking space than the child (Letts, 1985; Prutting et al. 1978; Ripich et al. 1984). Structure is regulated and maintained by the use of 'boundary markers', such as 'now' and 'OK',

which initiate new sequences of interaction, and 'attention-getters', such as 'ready' (Letts, 1985), which get the child's attention and simultaneously assert the therapist's role.

Hulterstam and Nettelbladt (2002) compared patterns of interaction in two different types of speech intervention, Metaphon (MT; Howell & Dean, 1994) and traditional therapy (TT). The authors' description of traditional therapy, which is limited, suggests an articulation (rather than phonological) intervention in which the clinician supports the child to achieve accurate productions using imitation and other operant learning techniques. They analysed 10 video recordings obtained from seven clinicians with a range of four to 25 years' experience and 10 children with phonological impairment. Recordings comprised five MT sessions and five TT sessions. They analysed discourse data using a method called initiative response analysis, which has similarities with CA (Linell, Gustavsson & Juvonen, 1988). In contrast to the procedure used by Prutting et al. (1978) and Letts (1985), Hulterstam and Nettelbladt (2002) coded turns based on the recipient's reaction (rather than the speaker's intention), which fits the notion of the perlocutionary act. Their analysis found that in both types of intervention interaction was asymmetrical. However, there was less asymmetry in the MT sessions compared with the TT sessions, operationalised by the degree of linkage between participants' turns, though this did not reach statistical significance. The authors discuss how their findings align with the underlying principles of the two approaches. MT sessions place emphasis on ensuring the child is an active participant. The focus is on supporting successful communication and repair. A metalanguage is used to describe and discuss speech sounds, which Hulterstam and Nettelbladt (2002) suggest encourages active participation. In TT sessions, the focus is on getting the child to achieve correct production so the therapist uses techniques more consistent with operant learning techniques, eliciting and reinforcing child productions.

Early studies of discourse during speech and language intervention described the structure of interaction as a systematic recurrence of a three-part sequence: therapist request or initiation; child response; and therapist evaluation of the child's response, commonly abbreviated as RRE or IRE. It resembles the IRE sequence that is characteristic of teacher-pupil discourse in the classroom (Sinclair & Coulthard, 1975). Prutting et al. (1978) and Ripich et al. (1984) used discourse analysis to code the types of requests that occurred in their therapy data, e.g. Prutting et al. (1978) coded nine different types of request, which included request for action; attention; completion; and imitation. Eight of the nine requests were verbal and the examples they provide suggest they were explicit. The ninth was a nonverbal request; they gave an

example of showing the child a picture of a boy riding. The target response was a verbal description of what the boy was doing. In the absence of an accompanying verbal utterance, this resembles an implicit request. Prutting et al. (1978) used audiotape data so it seems reasonable to suggest their analysis would not have revealed a thorough picture of the nature of therapist requests. They also coded on the basis of the illocutionary act, thus not taking into account how the child responded.

Letts (1985) describes a more data-driven approach that did not use pre-determined categories. One category that emerged from her analysis is a type of stimulus “which involved neither direction nor questioning” yet “a response was plainly expected of the child” (p326). Letts (1985) categorises this as a type of on-going act within the subcategory ‘to elicit specific responses’ and refers to it as a ‘signal’. She provides the following examples: models (to elicit an imitation); prompts (to elicit a completion); and (nonverbal) points to a picture (to elicit a name). Although she does not describe them as such, signals appear to be a type of implicit request. None of these early studies examined the nature of therapist requests at the level of directness or explicitness.

Prutting et al. (1978), Ripich et al. (1984) and Letts (1985) describe the evaluation part of the sequence structure, but with limited scope. Ripich et al. (1984) expanded on Prutting et al.’s (1978) description by adding evaluation as a sub-type of statement. Letts (1985) acknowledges giving feedback as a subcategory of on-going act used by the therapist to create opportunity for the child to learn and change. She states that therapists generally indicate approval with “straightforward comments such as good, right” (p327) and tend to show more approval than disapproval, but does not describe the nature of feedback. Letts (1985) does not present a sequential analysis of therapy discourse and so does not deal with the consequences of elicitation and evaluations. Panagos et al. (1986) differentiate between ‘simple’ sequences, in which the therapist gives a positive evaluation to a prompt and adequate response by the child, and ‘complex’ sequences, in which the therapist reformulates or repeats in the evaluation slot in response to an erroneous or non-response by the child, typically until an adequate response is reached.

4.1.6 Nature of therapist turns

Early studies of therapy interaction described therapist turns as requests or acts and focused predominantly on verbal utterances, perhaps limited by their use of audio rather than video data. However, a handful of studies did pay specific attention to nonverbal behaviours and

made some early attempts to describe prosodic features. Panagos et al. (1986) introduced the notion of a 'clinical register' to describe the therapist's manner of speaking in therapy, which they describe as "businesslike... the clinician's tone is firmer, her voice is louder" (p214). In addition to distinctive verbal features, they describe characteristic nonverbal features, such as pausing, the use of gesture, and postural changes, namely methodical forward-backing body leaning, forwards on initiations and backwards on evaluations, following a child response. Prutting et al. (1978) and Letts (1985) described nonverbal requests, but their studies did not examine the nature of nonverbal action. Panagos et al. (1988) examined nonverbal behaviours in videotaped therapy sessions conducted by therapists (sample size unknown). They were interested in what therapists used these behaviours for and how they were coordinated. Similarly to Bobkoff and Panagos' (1986) findings on the use of pointing, Panagos et al. (1988) found nonverbal behaviours were used to fulfil a variety of functions, e.g. to get the child's attention, to invite production, and to evaluate a response. They were coordinated with verbal directives alongside prosodic alteration, e.g. accompanying a request: child-directed eye gaze; point to the object; head tilt; rising intonation; and amplified voice. Nonverbal behaviours also appeared in clusters to reinforce a message and add intensity, and were utilised in increasing number in pursuits when the child did not comply.

4.1.6.1 Speech acts in phonology therapy

McCartney (1989) examined the 'speech teaching strategies' that therapists used in phonology therapy. She analysed audiotape recordings of therapy sessions featuring 10 therapist-child pairs. Pairs consisted of eight therapists who had a range of two to 29 years of experience; one of the therapists featured in three pairs. The 10 children ranged in age from 4 to 8;11 and had articulation or phonological difficulties. McCartney (1989) looked specifically at 'speech-teaching exchanges'; sequences in which the child made an attempt ('try') at the target word. In a similar manner to Prutting et al. (1978) and Ripich et al. (1984), analysis was at the level of speech act, not the turn, and the method used pre-determined categories, thus resembling discourse analysis. McCartney (1989) examined the speech teaching strategies therapists used to initiate a try, provide feedback on a try, and initiate a re-try. She also coded utterances that gave the child metalinguistic information about their production or the target form and those that directed the child's attention to the task or the therapist. McCartney (1989) coded therapist initiations according to five categories: no model; model; augmented model; error copy; and meta. She explains a 'no model' initiation typically elicited production by pointing to a picture. However, there is presumably a whole host of possible utterance types that would

fall within this category if the only defining feature is that it does not contain a model; verbal, nonverbal or both. McCartney (1989) acknowledges the use of prosodic modification in the initiation categories both in an 'augmented model', whereby prosody is used to make the target sound or word more salient, and in an 'error copy', whereby interrogative prosody is used to signal to the child that a re-try is required. However, restriction at the speech act level prevents distinction between descriptive features and functional consequences. It also seems that these categories combine both implicit and explicit requests for tries and re-tries. To identify the ways therapists elicit attention and participation from young children during multisensory input video-therapy (MSIVT), an input-based approach, it seems important to make a distinction between implicit and explicit initiations. McCartney (1989) found that modelling featured highly in the phonology therapy sessions she studied; it accounted for 48% of therapist initiations. Augmented models occurred infrequently and mostly in re-initiations after an inaccurate try.

Some of the limitations of McCartney's (1989) study are relevant considerations for the present study. McCartney (1989) did not code therapist initiations that did not elicit a try, and she looked specifically at 'try' responses. Findings from the pilot study show that, during MSIVT, therapist stimulus models do not necessarily elicit tries by the child. In addition, child tries do not necessarily follow stimulus models by the therapist. The pilot study showed children may participate in lots of other ways that might still be meaningful to the process of therapy, and such behaviour may be nonverbal.

These early studies provide descriptive accounts of the verbal and nonverbal features of therapist requests or initiations, but they do not explain their significance to the process of therapy. Subsequent studies using conversation analysis (CA) have taken us away from the RRE structure. Analysing interaction at the level of therapist and child turns has moved the focus to the design of therapist turns, the actions they fulfil, and the consequences they have for the child. As these studies have typically used video rather than audio data, they have also paid greater attention to nonverbal aspects of therapy interaction. Gardner's (1994; 1997; 1998; 2007) research using CA not only addressed this in relation to phonology therapy; it also revealed there was more to discover about the extent and contribution of prosodic modification, at least in intervention targeting speech change.

4.1.6.2 Turn design, action and consequence

Gardner (1994) analysed interaction between speech and language therapists (SLTs) and eight

children with phonological disorder aged 3;11 to 5;9 during phonology therapy sessions. She also carried out a comparison between therapist-child and mother-child interaction (see also Gardner, 2007). Her research had two main foci of analysis: the nature of therapist models of target words and types of response they engendered; and the nature of therapist and mother repair work, i.e. how they responded to the child's inaccurate productions.

Gardner (1994) found that therapists showed a dispreference for overt requests for target word production, suggesting the nature of therapist initiations is not always explicit. Therapists used other verbal and nonverbal markers to mark out words that they wanted the child to produce. This is an interesting finding and it has implications for the present study because of the nature of differences between MSIVT (for young children with CPSSD) and the type of phonology therapy that Gardner (1994) examined. It raises the consideration that therapists may make implicit requests for output during MSIVT using similar practices to the therapists in Gardner's (1994) study.

Like McCartney (1989), Gardner (1994) observed a high level of modelling. It was regarded as a direct method of eliciting an initial try from the child as well as a means of other-correction to elicit a re-try following an erroneous production. Compared to the types of direct requests for production described in the early studies of therapy discourse, a model seems to be an implicit way of requesting production from the child. Gardner's (1994) detailed turn-by-turn analysis revealed that not all therapist target word productions engendered an imitation. Distinctive prosodic and nonverbal features in therapist turns determined the action they fulfilled and the consequences they had for the child. With specific attention to target word production in initiation turns, Gardner (1994) found that target words set up as models for imitation were characterised by a configuration of features including a definite break before the word, phonetic distortion, and syntactic alteration, e.g. omission of the definite article. Such features served to tell the child something of interest was coming. Sustained eye gaze accompanied by a pause was also a strong indicator to the child that a response was required. In contrast, a 'redoing' of the child's production, serving as an acceptance and closing statement, was typically characterised by a mimicking of the child's prosodic pattern, shift in gaze and sometimes accompanying verbal praise. These findings are particularly relevant for the present study. In MSIVT, subtle prosodic and nonverbal features in therapist turns may distinguish the responses they elicit from the child, for example, imitation of the stimulus versus attention to the stimulus.

Interestingly, Gardner (1994) found that the prosodic subtleties in therapist turns meant the child did not always decipher what the therapist intended. This contrasts with Tarplee's (1996) findings from her longitudinal study of everyday talk between a parent and a child aged 1;7 to 2;3, much younger than the children in Gardner's (1994) study, and more similar in age to the children in the present study. Tarplee (1996) found that when the parent paused before repeating a given word, the child engaged in phonetic repair, i.e. the parent's utterance served as a model or re-elicitation, whereas a direct repeat with no pause did not lead to repair, i.e. it served as an affirmation. Other studies of everyday talk in young children, e.g. Gallagher (1977), have shown that 18 month to two-year-old children engage in phonetic repair, but it tails off quickly as their language develops.

Another design feature of therapist turns revealed by Gardner's (1994) analysis is explicit phonetic description and evaluation in repair sequences, i.e. where the child made an inaccurate production and the therapist encouraged the child to have another try. Gardner (1994; 2007) found that the therapists, in contrast to mothers, became increasingly more explicit in their pursuit for a repair, suggesting this is a specific characteristic of therapy interaction. Although McCartney (1989) had previously identified that these types of utterances were a feature of phonology therapy talk, analysis at the level of speech act was not capable of providing insight into their effects on the continuing trajectory of talk. Gardner's (1994) analysis revealed that such explicit phonetic comments led to phonetic repair whereas non-explicit comments had a tendency to engender semantic repair. Gardner (1994) used quantitative analysis to show that, in therapist-child repair sequences, the child reached target production in fewer turn 'bouts' than in mother-child sequences. Like Leahy (2004) and Leahy and Walsh (2010) in their studies of stuttering therapy, Gardner's (1994) findings explicate the process of phonology therapy as a collaborative accomplishment between the therapist and child. A dispreference for overt, explicit elicitation requires the child to decode the therapist's implicit requests. When the child's response is off target, the therapist uses her theoretical knowledge to steer the child towards accurate production: therapist and child work together to achieve changes in the child's speech.

Gardner (1994) suggests this design feature of therapist repair initiations reflects the therapists' theoretical knowledge about phonological development and the goals and aims of therapy. These conclusions are reinforced by the findings of Hulterstam and Nettelbladt (2002) in their comparison of interaction patterns in two types of speech intervention. They found Metaphon sessions were characterised by lower levels of clinician questions and requests for

direct action (soliciting initiatives) and a preference for strategies that elicited a comprehension response, whereas traditional therapy sessions were characterised by higher levels of soliciting initiatives and production-eliciting strategies. These findings suggest the types of requests and strategies therapists use during speech intervention relate to the specific principles of the approach.

4.1.6.3 Multimodality

More recently, Tykkyläinen (2009) and Ronkainen (2011) use the term 'multimodal' to capture the design of therapist turns during language therapy and auditory-verbal therapy, respectively. Tykkyläinen (2009) used CA to examine interaction during task setting in two different contexts: six pairs of mothers with their typically developing children; and six pairs of SLTs (with 6-20 years of professional experience) with children who had specific language impairment (SLI). Typically developing children ranged in age from 4;10 to 5;1; children with SLI ranged in age from 4;11 to 6;0. In both contexts, adults led the interaction. Tykkyläinen (2009) found that both the mothers and therapists utilised verbal, nonverbal and prosodic resources in their turns, e.g. named items, pointed to pictures, changed their body posture, placed emphasis on key words, and exhibited changes in pitch, i.e. their turns were multimodal in nature. Previous accounts of discourse reflect multimodality in descriptions of a clinical register and the nonverbal components that therapists utilise alongside verbal components (e.g. Panagos et al., 1986; Panagos et al., 1988). Tykkyläinen (2009) found that whilst both therapists and mothers used multimodal turns for common purposes, e.g. to attract the child's attention and highlight important information, differences existed in how they used them. Tykkyläinen (2009) identified two specific practices in the therapists' use of multimodality that she did not see in the parents. Therapists segmented turns before introducing the critical information of the task and they added on extra elements once the task setting was complete; both practices were used systematically. Tykkyläinen (2009) discusses how such specialised practices serve to secure the child's attention and facilitate their understanding and learning, and reflect the therapeutic goals underlying this context of interaction. These findings support the notion of a clinical register characterised by a special myriad of verbal and nonverbal features and prosodic modification.

Ronkainen's (2011) research reinforces Tykkyläinen's (2009) findings of multimodality. She also extends the enquiry by looking specifically at how it is used in different types of turn that fulfil different actions and have different consequences for the child. Ronkainen (2011) used CA to

examine interactions between three children with congenital deafness and an experienced SLT. All the children had a cochlear implant. At the time of the study, they ranged in age from 3;0 to 4;1 and were all at a pre-linguistic stage of development. The children received auditory-verbal therapy, an early intervention approach used by SLTs in Finland to enhance a child's listening and imitation skills and increase their use of oral interaction (Eastbrooks, 2006, as cited in Ronkainen, 2011). Ronkainen's (2011) data consisted of 55 minutes of video recordings from therapy sessions of dollhouse play in which the therapist used auditory-verbal techniques. She looked specifically at sequences ($n=30$) where the therapist introduced a new object. Ronkainen (2011) describes two types of sequence that occurred in her data: sequences of mutual orientation; and attention-seeking sequences. The former begins from a place of mutual orientation, i.e. the child's attention is directed at the therapist so the therapist can introduce the object. These feature one pair of turns: the therapist introduces the new object and the child responds, e.g. vocalises. In attention-seeking sequences, the child is not orientated to the therapist at the beginning of the sequence so the therapist seeks to establish joint attention before introducing the object. These sequences feature two pairs of turns: the therapist seeks the child's attention and the child responds, e.g. looks at the therapist (referred to as a 'pre-sequence'); and the therapist introduces the new object. Ronkainen (2011) is characterising the therapist's turns as actions, differentiating between seeking attention and introducing an object.

Similar to Tykkyläinen (2009), one of Ronkainen's (2011) key findings was that both types of turn were multimodal, i.e. they featured a myriad of verbal and nonverbal features. The therapist modified her verbal utterances with prosodic features, such as songful intonation, rising pitch, prolongation of vowels, and changes in tempo and voice quality. In contrast to Panagos et al.'s (1986) description of the clinical register, Ronkainen (2011) says, "one important feature in therapy... is playfulness and laughter" (p264). She shows how, during auditory-verbal therapy, multimodality and intense prosodic modification work together to establish and maintain the child's attention and interest, and encourages them to listen to and imitate speech. In the extracts Ronkainen (2011) presents, attention-seeking turns do not feature explicit verbal requests for the child's attention; instead, the therapist appears to utilise a number of different verbal and nonverbal resources that create a type of implicit request, e.g. saying the child's name, hiding a toy between her hands, producing interjections like 'ooh' with exaggerated prosody (Extract 4, p257). This is also seen in object introducing turns, whose features appear to exhibit implicit rather than explicit requests for attention and

vocalisation, e.g. producing an utterance with melodic prosody alongside a synchronous gesture while leaning closer to the child (Extract 1, p23). There are similarities here with Gardner's (1994) finding of implicitness in the nature of therapist models and redos during phonology therapy.

Ronkainen's (2011) study is relevant to the present study because the therapy approaches, auditory-verbal therapy and MSIVT, share some common therapeutic goals, namely to elicit attention and imitation. It raises the consideration that interaction during MSIVT may exhibit some of the same features of multimodality, prosodic modification and implicit directive work by the therapist.

4.1.7 Child-led interventions to support early language development: Nature, structure and patterns of interaction

A relatively small number of studies have examined the interactional features of child-led intervention, also referred to as child-centred, child-initiated, and naturalistic intervention (Cole & Dale, 1986; Kovarsky & Duchan, 1997; Norris & Hoffman, 1990). Child-led intervention has a different theoretical basis to adult-led intervention; it considers the form of language and social use of language as indivisible and places greater value on the use of language than on the accuracy of specific linguistic forms (Fey, 1986). This gives rise to different interactional features. Child-led approaches emerged, in part, as a response to concern that the highly structured and asymmetrical nature of adult-led intervention is so far from everyday interaction that the child's learning may not generalise to their everyday life (Cole & Dale, 1986; Kovarsky & Duchan, 1997). In child-led intervention, the therapy context closely resembles everyday situations; it is less structured, characterised by free play activities, and the child takes the lead in the activities and with the toys that take their interest (Cole & Dale, 1986; Norris & Hoffman, 1990). The therapist attributes meaning to the child's behaviours so that, over time, they recognise that their behaviour has an effect on their social environment. The goal of therapy is meaningful communication, not accuracy of the child's forms, so the therapist does not explicitly evaluate the child's behaviours or initiate repair of inaccurate forms; evaluation is implicit and any repair work is orientated to clarifying the child's message. The three-part RRE sequence described in studies of therapist-led intervention is not a typical feature of child-led intervention (Kovarsky & Duchan, 1997).

Studies largely fall into one of two categories: comparison of the effects of therapist-led versus child-led intervention, to investigate whether one style of interaction produces better

outcomes (Cole & Dale, 1986); and conceptualisation of the dimensions of therapist-led versus child-led interventions, as a way of distinguishing between their contrasting characteristics (Kovarsky & Duchan, 1997).

Cole and Dale (1986) did not find a difference in the effects of direct and interactive language interventions on language levels in 44 children aged 38 to 69 months with language delay. They used a randomised treatment design: 22 children received direct intervention, which involved techniques that are common to therapist-led intervention, e.g. elicited imitation and structured reinforcement; and 22 children received interactive intervention, which had characteristics of child-led intervention. Intervention fidelity was measured using a method resembling discourse analysis, which featured a speech act taxonomy similar to that described by Prutting et al. (1978). They found that both interventions resulted in change, shown by statistically significant pre-treatment post-treatment differences on eight of nine language measures.

Norris and Hoffman's (1990) study asked a similar question but used a different study design. They compared the effects of therapist-led and child-led intervention on child behaviours in five children aged 2;6 to 2;10. The children were at a pre-linguistic stage of language development and had concomitant cognitive delay. They were videotaped during a single therapy session conducted by an experienced therapist; each child received 25 minutes of intervention in each of the two conditions. The effects of intervention were measured by examining frequency of occurrence of a range of communicative behaviours. All five children exhibited higher levels of communicative behaviours in the child-led condition, demonstrating that direct attempts to elicit and reinforce vocal behaviours did not increase their occurrence relative to spontaneous productions. These findings do not support those reported by Cole and Dale (1986). However, unlike Cole and Dale (1986), the authors did not include a measure of intervention fidelity, so one might question the validity of the two intervention conditions. This seems plausible given that the treatment design relied on the therapist switching from one style to the other within the space of a single session.

Kovarsky and Duchan (1997) approached this topic from a different perspective. They propose a framework of five dimensions to differentiate between the interactional characteristics of adult-centred and child-centred language interventions: intervention event (activity structure); therapist's agenda; interactional lead; evaluation; and repairs. Fey (1986) describes approaches that have features of both therapist-led and child-led interventions as 'hybrid'.

Kovarsky and Duchan (1997) used turn-by-turn analysis to examine interaction during an intervention session to show how these dimensions manifest. The session was conducted by an SLT and a student SLT and the child was 4;7 with limited verbal expressive language and autistic tendencies. The approach was described as child-centred and the goal was to increase the level of child requests. The authors analysed a single videotaped session lasting 20 minutes. The analysis revealed an assortment of verbal and non-verbal features from both styles of intervention. They describe a highly controlled routine, e.g. withholding toys while inviting the child to make a request, and an asymmetrical pattern of interaction, underpinned by the therapist's agenda (adult-centred). Activities and interaction orientated to the therapist's goal (adult-centred) but opportunities were provided for the child to make his own choices (child-centred). Interactional exchanges did not conform to the RRE sequence structure and the level of repair work was low (child-centred). The level of structure within the session was explained as a way of dealing with the low level of child initiations. These findings suggest therapists may draw on different features in therapy to deal with specific circumstances. In this study, it was the nature of the child's communication difficulties. It seems the child's age, stage and rate of expressive language development might influence the interactional features of a particular approach or session. In their study of enhanced milieu teaching with phonological emphasis, a naturalistic intervention, Kaiser et al. (2017) highlighted that the low rate of productive language in their 15-36 month old participants posed a challenge to the delivery of the intervention.

4.1.8 Making therapy practice explicit: Lessons from studying aphasia therapy sessions

There is a large volume of studies on aphasia therapy that demonstrates a pursuit to make the processes by which therapy takes place more explicit, and a recognition of the potential of CA as a method for achieving this (Horton, 2006; Horton & Byng, 2000; Simmons-Mackie, Damico, & Damico, 1999). Horton (2006) examined therapy sessions that had been video- and audio-recorded. Participants were 14 therapists and 13 persons with aphasia, making 14 dyads (one of the persons with aphasia paired with two different therapists). Horton produces a broad descriptive framework for the structure of therapy sessions, which includes general domains such as 'opening up the business', 'doing therapy tasks' and 'the closing down period'. He then draws on CA to produce a detailed analytical account of what happens within individual domains in terms of the interaction between therapy participants. Horton's (2006) analysis

illustrates the complex nature of therapy. He describes “the participants in this process do not just go through therapy routines, or have conversations, or give feedback – they are doing all of these things all the time in a sequential process that ebbs and flows through a whole session...” (p557). This volume of research supports a deeper level of understanding of the interactive work that takes place during therapy. It supports a notion that making therapy explicit will enhance communication and teaching about therapy and improve the quality and effectiveness of therapy. There is no known research of this kind that has yet been carried out on input-based therapy.

4.2 Section 2: Non-clinical interaction: Insights from CA research

In this section I review some studies whose methods and findings are relevant to the present study. Before I begin it is necessary to introduce and define some key terms and concepts. *Response mobilisation* is the process by which one individual elicits a response from another. Stivers and Rossano (2010) suggest that speakers use a combination of resources to mobilise a response from the recipient. Resources include: the social action which the speaker produces; the sequential position in which it is delivered; and non-verbal and prosodic features of the speaker’s turn (*turn-design features*). They make a distinction between canonical and non-canonical types of social action. *Canonical* action types make it relevant for the recipient to respond, for example, offers (e.g. ‘want some tea?’) and requests for information (e.g. ‘what time is it?’). If the recipient does not respond, this would generally be regarded as problematic. *Non-canonical* action types make a response less relevant, holding the recipient less accountable for a response. Examples include assessments (e.g. that’s great’) and noticings (e.g. ‘it’s getting late’). One of the turn-design features that Stivers and Rossano (2010) discuss is *recipient-tilted epistemic asymmetry*. This describes turns about affairs within the recipient’s (not the speaker’s) domain, e.g. their plans, past experiences, likes, etc. As such, these utterances more routinely attract a response from the recipient.

4.2.1 Response mobilisation

Stivers and Rossano’s (2010) paper on interaction raises some relevant considerations for the present study. They investigated response mobilisation in 50 hours of videotaped conversation in English and Italian using methods utilising CA. They identified four types of (canonical) actions performed by speakers that routinely and reliably received a response: invitations;

requests for action; requests for information; and offers. They then examined turns in which the actions were performed in order to identify their design features, i.e. how they were delivered. They found four particular design features occurred across the turns: interrogative prosody; interrogative lexico-morphosyntax; recipient-directed gaze; and recipient-tilted epistemic asymmetry (i.e. within the recipient's domain not the speaker's). However, there was not one feature that was present in all cases, suggesting a combination of features (not one single feature) mobilises a response. Absent responses to these types of actions were typically treated as problematic; evidenced by a pursuit by the speaker. Stivers and Rossano (2010) also examined turns that represented non-canonical actions, i.e. actions that are less frequently cited as first-pair parts, e.g. announcements, noticings and assessments. They found that absent responses to these actions were less commonly treated as problematic. However, they identified that when one or more of the aforementioned turn design features were present they more reliably engendered a response. Their other key finding was that the same four turn-design features that mobilise response are also utilised in pursuits following an absent response.

Stivers and Rossano (2010) discuss how the design of a speaker's turn has implications for the level of pressure on the listener to respond. In the absence of response-mobilising features, there is minimal pressure on the listener, but this increases as the turn design becomes more response mobilising. They consider how certain situations may have a dispreference for highly response-mobilising turn designs in order to create opportunity for volunteered responses. Their extracts show how initial turns that are minimally response-mobilising can become more response-mobilising in pursuit sequences.

These findings have relevance for the present study because of the high level of implicit action we might see in therapist turns and the need to create a non-pressurising environment for the child. In existing descriptions of MSIVT, the therapist is described as aiming to create opportunities for production responses but not explicitly request them; does this mean s/he displays a preference for non-canonical action types? In phonology therapy, Gardner (1994) made a distinction between models (for imitation) and redos (for clarification), which were distinguished by their turn design features. There are similarities between the features Gardner (1994) identified in models and the response-mobilising features Stivers and Rossano (2010) identified in some non-canonical actions, e.g. child or recipient-directed gaze, distinctive prosody. In Gardner's (1994) data, absent responses (child imitations) to therapist models were treated as problematic. In light of Stivers and Rossano (2010), a model may be

regarded as a non-canonical action with response-mobilising features. In a similar vein, there are similarities between the multimodal object introducing turns that Ronkainen (2011) describes and non-canonical announcements with response-mobilising features.

Stivers and Rossano (2010) suggest that the response engendering consequences of a speaker's action are only understood with consideration of both linguistic and sociological perspectives. They propose a scalar model of response relevance that delineates the response-mobilising degree of a speaker's turn as relating to the action it implements and how it is delivered. It recognises that some actions, e.g. (canonical) request for action, are more response-mobilising than others, e.g. (non-canonical) announcement, and some turn designs, e.g. one with interrogative prosody and recipient-directed gaze, are more response-mobilising than others, e.g. interrogative prosody only. The model also acknowledges that actions in sequence initial position are highly response-mobilising (Schegloff & Sacks, 1973).

4.2.2 Clustered actions and response relevance

Reed, Reed and Haddon (2013) used CA to examine interaction between professional masters, accompanists and singers in vocal master classes; sample size not provided. The accompanists and singers were mostly students. Vocal master classes are characterised by some of the same patterns of interaction seen in therapist-led speech and language therapy sessions, e.g. asymmetry between the master and singer/accompanist, and constraints on when and how the students can contribute (Reed et al., 2013). Masters primarily talk and students (singers/accompanists) primarily perform. As in therapist-led therapy sessions, the interaction is geared towards goals and both the master and the student work together to achieve them.

The authors use the term directives to refer to a recurrent type of action that occurred in their data. They make a distinction between directives, which tell the recipient to do something, and requests, which recognise that recipients can choose whether or not to respond. Interestingly, and unlike the types of invitation actions identified by Gardner (1994) in phonology therapy and Ronkainen (2011) in auditory-verbal therapy, Reed et al. (2013) explain that verbal and nonverbal components and spatial orientation are utilised in directives and some are entirely nonverbal. They delineate directives by two specific criteria: the selection of a recipient; and the specification of a change in recipient behaviour. Using Stivers and Rossano's (2010) theory, these two criteria make directives maximally response-mobilising.

Reed et al. (2013) identified that the decision dilemma for students in vocal master classes was

not whether or not to respond to a directive, but when to respond. They found that directives frequently appeared within clusters in a single turn, where masters would deliver several clusters in succession without giving the students opportunity to respond. The students therefore had to negotiate whether to respond immediately on completion of a particular directive (now) or await further directives (a not now response).

They identified two types of directives in their data, distinguished by their spatial orientation and the type of response they received: 'local' and 'restart-relevant' directives. The former engendered an immediate response by the students, typically a repeat or a specific verbal or nonverbal behaviour. The latter engendered a response at a later stage, typically a restart of the performance rather than a repeat of a particular section. The authors found it was the restart-relevant directives that typically featured in clusters with no opportunity for students to respond until after the final directive was delivered. As such, sequences containing clustered restart-relevant directives did not conform to the IRE structure that is characteristic of teacher-learner interaction. This was in contrast to sequences containing local directives, which did include a response slot to be taken up by the students and therefore did resemble the IRE structure.

A final observation from these two studies (Reed et al., 2013; Stivers & Rossano, 2010), relating to the nature of responses, may have relevance for the present study. Stivers and Rossano (2010) discuss response-mobilisation in relation to the mobilisation of talk, primarily verbal responses. In the Reed et al. (2013) study, responses were more commonly action, in their case musical performance: playing the piano (student accompanist); and singing (student singer). Both studies illustrate the necessity of fine-grained turn-by-turn analysis in capturing design features of the initiation action and the nature of response. These are important considerations for the present study given the input-based nature of MSIVT and the types of verbal and nonverbal responses therapists might elicit.

4.2.3 Longitudinal studies of interaction

Since therapy is a process that takes place over time, it is appropriate to examine MSIVT longitudinally to find out how interaction changes over time as the therapist and child become more familiar with therapy and each other. This is important because the aim of therapy is to facilitate change in the child, so an understanding of how the interactional context for this change evolves over time will provide insight into how this change is accomplished. In Pekarek Doehler, Wagner and González-Martínez (2018), authors draw on conversation analysis (CA) to

study and document change across time in a variety of settings and practices. One field of interest has been the interactional practices of very young children. Flipi (2018) studied the development of the response token 'yes' in two children, Cassie and Rosie, aged 10-24 months. She used data on interactions that took place at home during everyday activities: Cassie with her father and Rosie with her mother. Flipi (2018) analysed the way the children exhibited 'yes' in response to summons and yes/no questions. Her analysis revealed a developmental trajectory from nonverbally designed turns at 10 months, featuring speaker-directed gaze and body orientation, to a more distinctive head nod at around 15 months and, by age two, verbally formulated 'yes' responses. Flipi (2018) explains how, in addition to learning how to design their own turn, a young child must learn how to make sense of another speaker's turns. She illustrates this with an extract of Cassie, in which her rejection of her father's responses and pursuit for something more acceptable explicates her previous interactional experiences. This study by Flipi (2018) is an example of studying how the same action (affirming, agreeing, acknowledging) is accomplished differently at different moments in time.

Forrester (2008) used a longitudinal design to study the emergence and developing profile of self-repair practices in his daughter, Ella. His data consisted of video recordings of Ella from when she was one year old up to 3;5 when she was interacting during meal times. Forrester's (2008) analysis identified the earliest instances of self-repair when Ella was 15 months old (65 weeks). The analysis revealed differences in the profile of Ella's self-repair practice at different ages, for instance, in the nature of the immediate antecedent and in the form and design features of Ella's repair. Up to around two years of age, self-repair was usually preceded by a nonresponse from a participant, whereas beyond age two there was a predominance of spontaneous self-repair. Ella employed a variety of resources, e.g. altered sounds, gestures and gaze, which became more complex over time. In the early years, the analysis showed that Ella's self-repairs were typically repeats or reformulations to gain attention or make requests, whereas by age two she was using repair for a wider range of functions, e.g. to ask questions.

These studies by Flipi (2018) and Forrester (2008) show the potential of longitudinal analysis to reveal differences in the way the same action is accomplished at different moments in time, as well as differences in the developing profile of action. They involved children of a similar age to the children in the present study. However, where Flipi (2018) and Forrester (2008) focused on changes in the child's practices over time, the present study is particularly interested in how the therapist's interactional practices change and the consequences this has for the child.

4.2.4 Frames and footing

The review of the existing literature on MSIVT (Chapter 2) found some references to the video-therapy component that suggests Goffman's (1981) concepts of 'frames' and 'footing' have relevance for the present study. Harding & Bryan (2000) describe the provision of therapy videos as a means of optimising the level of input stimulation that the child receives and creating opportunities for self-rehearsal and practice. The purpose of making the videos is therefore to provide them to the family to watch in-between therapist-led sessions. Calladine and Vance's (2019) findings of the therapist in their data using the camera within the therapy session itself raises the idea that the therapist in an MSIVT session interacts with the child in two different frames of events: the 'live' frame of therapy activities with the child in the here and now; and the 'future' frame of therapy with the child at a later moment in time, not now; (there are similarities here with Reed et al., 2013). Goffman (1981) describes movement or switching between frames as a change in footing, which "implies a change in the alignment we take up to ourselves and the others present as expressed in the way we manage the production or reception of an utterance" (p128). Goffman (1981) describes it as a persistent feature of naturally-occurring talk and uses examples from broadcasting (where a broadcaster has both a 'live' and broadcast audience) and teaching to illustrate its characteristic nature in constitutional talk. However, rather than merely switching from one alignment to another, Goffman (1981) goes on to add "...we are not so much terminating the prior alignment as holding it in abeyance with the understanding that it will almost immediately be reengaged" (p155). This results in layers of embedded interaction and multiple participation frameworks through which the speaker guides delivery. The potential role of the video camera in facilitating the delivery of MSIVT suggests the therapist's interaction with the camera is an important consideration for the present study.

4.3 Conclusions

This chapter has reviewed the literature on clinical interaction in speech and language therapy, primarily involving children. Studies that have drawn on discourse analysis methods describe therapist-led intervention as highly structured with an asymmetrical pattern of interaction. The session is underpinned by the therapist's agenda to stimulate change in one or more aspects of the child's development. Child-led or naturalistic intervention is less structured, with the therapist responding to the topics and objects that capture the child's interest and attention. Kaiser et al. (2017) describe challenges in using naturalistic intervention to target

speech change in young children who are at an early stage of language development. Studies using conversation analysis (CA) extend these descriptive accounts to capture the nature of therapist-child interaction during therapy. The fine-grained and inductive nature of CA explicates the ways therapists use multimodal features to interact with a child during therapy and the consequences their turns have for the child in the immediate context and for the continuing trajectory of talk. Studies like those by Gardner (1994) and Ronkainen (2011) show the potential of CA to provide both practical and theoretical insights into the practices therapists use to create a facilitative therapeutic environment for stimulating learning and development.

The present study is interested in the MSIVT approach. Whilst there are similarities between the long-term goal of MSIVT and some of the approaches discussed in this literature review, namely phonology therapy and auditory-verbal therapy, their underlying theoretical bases and the short-term aims of therapy are different. As an input-based approach primarily for children at an early stage of language development, MSIVT is designed to promote changes in speech output by stimulating speech input processing, such that a non-pressurising, non-directive and non-corrective environment is created. The literature review suggests a CA examination of MSIVT will be an effective way of identifying the interactional features and dynamic nature of the approach. Of particular interest are: the types of actions that therapist turns project and how these compare to those seen in other speech interventions; the consequences these have for the child; and how these align with the aims of MSIVT. The studies of non-clinical interaction give relevant insights into response mobilisation, clustered action, change over time and frames of talk, which also inform the study.

4.4 Aims and rationale for the study

The overall aim of the present study was to examine the nature of MSIVT as it is implemented within the NHS for young children with CPSSD. A review of the existing literature on MSIVT and early intervention for children with CPSSD (Chapter 2) has found that MSIM+/-VT is underpinned by a coherent theoretical basis and there is anecdotal preference for this approach in the UK. Current descriptions draw on theory and empirical evidence from clinical case studies and the non-experimental case study by Calladine (2009). They make recommendations for the model of delivery and describe a theoretically-sound approach to target selection and stimulus design. With regards to dynamic aspects of therapy, they suggest MSIVT has some features of output-based approaches, such as the procedure for responding

to child target productions, and some features of naturalistic approaches, such as the style of interaction. However, the absence of a manual means there is likely to be significant variation in how MSIVT is delivered from one therapist to another, and the empirical basis of MSIVT is weak. This must be addressed to support clinical practice and inform future research to develop and evaluate the approach.

The pilot study (Chapter 3) trialled methods for analysing and describing the nature of MSIVT. The findings support a specific aim of the main study to examine the delivery of MSIVT as an episode of care in order to extend current knowledge about characteristic features of service delivery and identify the sources of any variation in practice. The pilot study examined the dynamic nature of MSIVT by profiling therapist behaviours and child responses. The findings support an analytical focus in the main study on the ways therapists interact with the child during MSIVT rather than a descriptive focus on the behaviours they display.

The literature review presented in the current chapter (Chapter 4) has shown that CA studies of therapy talk have revealed characteristic features of therapy interaction and provide analytical insights into how therapeutic goals are achieved in interactional terms. However, very few studies have involved young children or interaction during an input-based approach. The review suggests a CA examination of MSIVT will reveal how it is enacted as an interactional process and extend knowledge of how very young children are engaged in speech intervention.

To give justice to the complex nature of speech and language therapy intervention, the present study focuses on what happens within therapist-led MSIVT sessions. It is a progression from previous studies because of the methods of analysis it employed, but this focus is based on the author's view that studies of the home-delivery component of MSIVT will be more effectively studied and understood if we first give adequate attention to the session itself.

4.4.1 Research questions

The study addresses the following research questions:

1. How do therapists deliver MSIVT as an episode of care within the NHS: what is the structure of an episode of care; what speech sounds do therapists target; and what types of activities, materials and speech sound stimuli do they use?
2. What are the interactional features of MSIVT, and in what ways do therapists establish a child's attention and stimulate their awareness and production of speech sounds?

3. How does therapist-child interaction change over the course of an episode of care; (comparing first and last sessions)?

Chapter 5: Methods

This chapter describes the methods used to identify, collect, organise and analyse data in Phases 1, 2 and 3.

Phase 1 used all data to investigate delivery of MSIVT as an episode of care within the NHS. It examines the structure of MSIVT episodes of care and the nature of activities, speech sound targets and stimuli, and compares these features across five therapist-child dyads (**Chapter 6**).

Phase 2 used a subset of data to analyse interaction during MSIVT. It used data from three of the therapist-child dyads; recordings from the first, middle and last sessions in the therapy episodes, plus additional data from the pilot study (three recordings from one of the two episodes). It closely examines the nature of therapist-child interaction in order to identify the specific practices SLTs use to support children with CPSSD during MSIVT sessions (**Chapter 7**).

Phase 3 looked in more detail at one of the therapist-child dyads, using recordings from the first and last sessions in the episode. It examines the nature of therapist-child interaction from a longitudinal perspective over the course of an episode of care (**Chapter 8**).

5.1 Study design

This was a qualitative and longitudinal study. It used data collected from pre-existing clinical records. A descriptive approach was used to investigate patterns of delivery of MSIVT and conversation analysis was used to investigate the nature and pattern of interaction between the therapist and the child during therapy. Five therapist-child dyads, comprising three SLTs and five children, at two specialist SLT services in the NHS, participated. Therapists provided video recordings and case notes of MSIVT sessions that were delivered as part of routine clinical care, not for research purposes. The number of recordings, corresponding to sessions, ranged from five to seven per therapist-child dyad, totalling 29 recordings and 573 minutes of video. Sessions were delivered between the months of June 2013 and December 2016 and episodes of care lasted between five and 11 months. All but two of the MSIVT sessions took place prior to the recruitment of SLT participants, so it is highly unlikely that the research influenced the delivery of therapy. The nature of the dataset allowed within-session and session-by-session analysis.

5.2 Procedures

5.2.1 Ethical approval

The study raised a number of ethical, legal and management issues due to the nature and age of child participants and the type of data collected. An application for NHS ethical approval was made to the West of Scotland Research Ethics Service. The application and Research Protocol are provided in Appendices D and E, respectively. Approval was granted on 2nd September 2015 following proportionate review; see Appendix F. The study was sponsored by Leeds Teaching Hospitals NHS Trust (LTHT), who granted Research and Development (R&D) approval on 21st September 2015; see Appendix G.

Ethical approval addressed concerns about the inclusion of participant identifiable data (PID), which includes patient information held by the NHS. All of the procedures listed here and described above: participant identification; participant recruitment; informed consent; data collection; data handling; and data storage, were developed in accordance with the International Conference on Harmonisation Good Clinical Practice guidelines (ICH GCP; 1996) and in collaboration with the Research & Innovation Department at LTHT. The ethical application also addressed potential legal issues that arose due to working with confidential data collected by the NHS. NHS treatment records are legal and confidential documents. The procedures for data collection, handling and storage were developed in accordance with the Data Protection Act 1998 and in collaboration with the Information Governance Department at LTHT.

5.2.2 Identification of eligible sites

There are 12 NHS nationally-commissioned regional cleft lip and palate networks in the UK (Cleft Lip and Palate Association; CLAPA, 2016). Six of these networks operate as twin-site centres, making 18 clinical units (The Cleft Registry and Audit Network; CRANE, 2015). With the exception of one twin-site centre, which has one regional lead specialist SLT overseeing both units, all other twin-site centres have a lead speech and language therapist at each unit. There are therefore 17 lead SLTs in the UK.

Children are typically under the shared care of the regional specialist SLT service within the cleft lip and palate unit and a local SLT service in the community (NHS England, 2013). A recent survey conducted in the UK showed there is extreme variability in the nature and resources of

regional and local provision across the country (Britton, Calladine, Extence, Phippen & Pinkstone, 2017).

To identify networks where there were potential SLT and child participants, it was necessary to identify those whose routine therapy provision included MSIVT. The author contacted all 17 lead specialist SLTs and requested information about therapy provision in place in the networks they oversee; see Appendix H. Twelve of the 17 lead SLTs responded; of these:

- 12/12 said therapy is provided to children between 18 months and three years
- 11/12 said therapy consists of multisensory input therapy
- 6/11 said therapy consists of multisensory input modelling with video-therapy, i.e. MSIVT
- 5/6 said they store copies of the MSIVT videos
- 1/6 said they do not store videos but might be able to get copies from families

One of the five units storing videos, unit A, is where the author takes up the role of lead SLT. This unit was excluded because the author considered it unethical to approach therapists for whom she has managerial responsibility in case they felt obliged to volunteer. This left four units: G, I, L and Q. Units L and Q were outside of England and so different NHS ethical procedures would apply. They were therefore excluded. Once ethical approval had been obtained, the author contacted the lead SLT at the two remaining units: G and I plus the lead SLT at unit N, where videos may be retrievable from families, with an invitation to take part in the study. All three lead SLTs responded to say they and their clinical teams were interested in taking part.

5.2.3 Research and Development (R&D) approvals

The author contacted the R&D departments at the three NHS organisations in order to seek R&D approval. One of these units, unit N, was not recruited because the organisation only takes part in studies on the National Institute for Health Research (NIHR) Portfolio and the current study did not meet criteria. The two other units, G and I, were recruited, and are hereon referred to as Site 1 and Site 2, respectively. NHS R&D approvals were granted by Site 1 on 27th April 2016 and Site 2 on 6th June 2016; see Appendices I and J, respectively. On receiving R&D approval, applications were made to both sites for a Letter of Access to give the author permission to access and retrieve data; these was obtained from Site 1 on 6th June 2016 and site 2 on 8th June 2016; see Appendices K and L, respectively.

5.2.4 Identification and recruitment

Potential SLT participants were identified and approached first. They were identified initially by the lead specialist SLTs at the recruited sites and then by themselves following a self-eligibility check. The author contacted the lead SLTs and provided written information about the study. The Therapist Invitation Letter and Therapist Participant Information Sheet are provided in Appendices M and N, respectively. The author requested that lead SLTs disseminate the information to SLTs in their respective specialist teams who they believed might meet inclusion criteria for the study. This method ensured the identities of potential participants were not disclosed to the author. The author requested that SLTs contact her if they wished to take part and believed they were eligible. Eligibility was confirmed in a subsequent telephone contact between the author and the SLT. Three SLTs were identified as meeting inclusion criteria for the study, one at Site 1 and two at Site 2.

To identify potential child participants, a review was undertaken of patient healthcare records at the recruited units. For ethical reasons, and to ensure the identities of potential participants were not disclosed to the author, this was carried out by clinicians in the direct healthcare teams who had routine access to this information. The participating SLTs therefore identified children who might participate in the study. The author requested that the SLTs contact her to discuss the children they believed met criteria so that eligibility could be verified. No patient identifiable information was shared in this exchange. Five children were identified as meeting inclusion criteria, one at Site 1 and three at Site 2.

Parents/carers of eligible children were invited to give consent to take part in the study on behalf of their children. Parents/carers of children identified as meeting the inclusion criteria were contacted by telephone by the participating SLT who provided their child's therapy. The SLT made a verbal invitation and supported this in writing with the Parent/Carer Invitation Letter and Parent/Carer Participant Information Sheet that were provided by the author; see Appendices O and P, respectively.

5.2.5 Informed consent

On receiving the information sheet, all participants were given time to review the written information they had received about the study and the opportunity to ask questions. The author completed the process of informed consent for SLT participants. To take part in the study, SLTs were required to sign the Therapist Consent Form; see Appendix Q. A second

consent form was provided requesting permission to use video data in presentations and teaching; see Therapist Consent Form: Use of Videos in Presentations and Teaching in Appendix R. This was an optional level of consent; SLTs could still take part in the study if they did not wish to give their permission for video data to be used in this way. The author arranged follow-up telephone discussions with the SLTs two weeks after they had received the written information. Verbal consent was obtained from all three SLTs and the consent forms were completed and returned to the author. One of the three SLTs signed the additional consent form, permitting the researcher to use her therapy videos in presentations and teaching.

The participating SLTs completed the process of informed consent for child participants. Consent was obtained from parents/carers on children’s behalf. To be involved in the study, parents/carers were required to sign the Parent/Carer Consent Form; see Appendix S. As with the SLT participants, an optional level of consent was invited to allow use of video data in presentations and teaching; see Parent/Carer Consent Form: Use of Videos in Presentations and Teaching in Appendix T. Parents of all five children signed both consent forms. For video data to be used in presentations and teaching, consent needed to have been obtained by both primary and secondary participants involved.

The study recruited five therapist-child participant dyads comprising three SLTs and five children. See Table 5.1 for a summary of this information and the identification names assigned to participants. One therapist, Helen, formed three dyads with three different children.

Table 5.1: Participant identification names

Site	Therapist	Child	Dyad
1	Sarah	Louise	Sarah-Louise
2	Laura	Ellie	Laura-Ellie
2	Helen	Thomas	Helen-Thomas
2	Helen	Hannah	Helen-Hannah
2	Helen	Naomi	Helen-Naomi

5.2.6 Data collection

The study used data derived from pre-existing NHS clinical records of MSIVT sessions that took place as part of the children’s routine clinical care. The clinical records consisted of written case notes and therapy videos. All therapy sessions took place in a hospital clinic setting and, for each dyad, all sessions took place in the same room. At Site 1, therapy sessions were filmed

using two different camcorders: Sony HDR-XR160E and Sony DRC-SR15E. Sound was recorded using the camera's inbuilt microphone; no external microphone was used. During filming, the camera was connected to a television and external Digital Versatile Disc (DVD) recorder. The SLT and child could therefore see themselves on the television screen as the session played out. The therapy DVD was provided to the family at the end of every session to take home with them. At Site 2, therapy sessions were filmed using a Canon XA10 camcorder. An external microphone (Rode NT3) was used in all sessions. During filming, the camera's small inbuilt screen was flipped around. The SLT and child could therefore see themselves on the camera screen as the session played out. At this site, the therapy DVD was produced after the session and posted to the families approximately one week later.

The author collected the data during site visits. These took place on 3rd November 2016 at Site 1 and 21st November 2016 and 2nd February 2017 at Site 2. Site-specific procedures were developed for accessing and extracting video and case note data in order to comply with local information technology (IT) and governance (IG) procedures. These are described below:

Site 1

The participating SLT (Sarah) identified and located the video data on the Trust network and saved it temporarily to a memory stick provided by the site's IT department. The data were then transferred to the author's two external devices:

- DataLocker Enterprise™ encrypted external hard-drive
- Ergo Engage 123 encrypted laptop

The author identified and located the speech and language therapy case note data corresponding to the child participant's (Louise) filmed therapy sessions on the electronic patient records system, SystemOne. Anonymised data were extracted into a Microsoft Word document and then transferred to the author's external devices via NHS mail.

Site 2

The two participating SLTs (Laura and Helen) identified the video data on the Trust network. A media technician in the Trust extracted the videos and saved them to encrypted and password-protected DVDs using the program Windows 7-Zip. These were un-locked using the 7-Zip program when the author returned to the host site and then transferred them to the external hard-drive device.

Speech and language therapy case note data corresponding to the filmed therapy sessions of the four child participants (Ellie, Thomas, Hannah and Naomi) was spread across two systems: EpicCare, an electronic patient records system, and paper records. The participating SLTs identified the data on Epic and the author identified the data in the paper records. All data were anonymised and extracted and then transferred to the author's external devices. No participant identifiable data (PID) was removed from the sites in hard copy format.

5.3 Materials

The study used data derived from pre-existing NHS clinical records. The rationale for using pre-existing clinical records was that the study wanted to investigate real life delivery of MSIVT that was free from any influence of a research study. The primary data for the study were the video recordings of MSIVT sessions conducted by the SLT participants as part of the child participants' routine clinical care. The SLT case notes corresponding to the MSIVT sessions formed supplementary data. For each therapy session, a 'complete dataset' consisted of the therapy video and the corresponding written case notes.

5.3.1 Summary of the data

The study collected 29 complete therapy data sets from five therapist-child dyads at two NHS sites. Therapy sessions took place between June 2013 and December 2016. Therapists Sarah and Helen were recruited in July 2016 and October 2016, respectively, which meant the last sessions in Sarah-Louise and Helen-Naomi's episodes took place after recruitment. The children ranged in age from 1;6 at the start of therapy to 2;11 at the end of therapy. The total number of sessions that took place was 42. Therapy videos were made in 32/42 (76%) sessions and were available for 31 sessions. Two videos from the Helen-Naomi dyad did not meet criteria for inclusion because therapy was provided by a different SLT. Case notes were available for all sessions. See Table 5.2 for a summary of the data.

Table 5.2: The complete dataset

Site	Dyad	Number of sessions	Month and year of first and last session	Age of child at first and last session	Number of complete datasets	Total duration of video data
1	Sarah-Louise	9	January 2016 to October 2016	2;3 to 2;11	7	168:03
2	Laura-Ellie	9	May 2014 to February 2015	2;0 to 2;9	6	95:36
2	Helen-Thomas	6	January 2014 to May 2014	1;6 to 1;11	5	59:47
2	Helen-Hannah	9	June 2013 to May 2014	1;6 to 2;7	6	90:18
2	Helen-Naomi	9	June 2016 to December 2016	2;3 to 2;10	5	159:26
Total		42	June 2013 to December 2016	1;6 to 2;11	29	573:10

Note. Video duration shown in minutes and seconds. Ages shown in years and months.

5.4 Participants

5.4.1 Child participants

5.4.1.1 Inclusion and exclusion criteria

Children were eligible for inclusion if they met the following criteria:

- Born with a cleft palate \pm cleft lip
- Under the care of an NHS Regional Cleft Lip and Palate Service
- Diagnosis of cleft palate-related speech characteristics
- Had a specialist SLT assessment before 27 months of age¹
- Received therapy for CPSSD between the ages of 18 months and three years
- The SLT who delivered the therapy meets the secondary participant inclusion criteria
- Therapy consisted of multisensory input modelling and video-therapy, i.e. MSIVT
- Therapy took place in the UK and since 2010²
- A minimum of five therapy records were available (from five separate sessions)³

¹ There is a national process standard of care in place in the UK, which states all children born with a cleft palate will be seen for a speech and language assessment by a specialist SLT before 27 months of age (Lead SLT Forum, 2016). This assessment is typically undertaken at 18-24 months of age.

² 2010 is the year after the first national dissemination of the Calladine (2009) study.

- The parent/carer of the child who received therapy consents to take part in the study

The following exclusion criteria applied to the child participants and data:

- The following information about the child is not available:
 - Cleft type
 - Surgical and wound healing history
 - Presence/absence of palatal fistula
 - Speech diagnosis
- The following information about the SLT is not available:
 - Previous training and experience
 - Current role
- Written case note entries for the therapy sessions are not available
- Video records of the therapy sessions are not available
- Video records are of poor audio and/or visual quality

5.4.1.2 Characteristics of child participants

Information about child participants was obtained by reviewing children’s medical notes and speech and language therapy case notes. The author extracted anonymised information and transferred it to the Child Data Collection Form shown in Appendix U. The children consisted of four girls and one boy. Four of the children were monolingual English speaking and one child, Ellie, was bilingual, speaking both English and Portuguese. The characteristics of child participants are summarised in Table 5.3 and described in detail below.

Louise

Louise had a history of Robin Sequence associated cleft palate. She underwent surgery at nine months of age (0;9) and this healed well with no palatal fistula. In line with the regional protocol at Site 1, Louise was seen for her 18-24 month assessment at age 1;11. According to the reports of the assessing SLT, Louise had age-appropriate receptive and expressive language skills, age-appropriate listening and attention levels, and age-appropriate play skills. In relation to her speech development, Louise had well-balanced resonance, which is a strong indicator of

³ Analysis of the two episodes of care in the pilot study found that it took at least five therapy sessions for the therapist to progress target phonemes and stimuli through several different linguistic levels. This informed a decision in the main study to apply an inclusion criterion of a minimum of five therapy records.

adequate velopharyngeal function (VPD). Her phonetic inventory consisted of [m b w g ŋ x h ʔ]. Louise was producing velar plosives for alveolar phoneme targets, which resulted in a diagnosis of backing to velar. Louise was 2;3 when she started therapy and received nine sessions up to the age of 2;11. She had a history of adequate middle ear function and hearing levels.

Ellie

Ellie had a repaired cleft of the soft palate and underwent surgery at six months of age (0;6). She did not have a palatal fistula. In line with the regional protocol at Site 2, Ellie was seen for her 18-24 month assessment at age 1;7. Her speech and language skills were developing appropriately for her age and she presented as sociable and communicative. Ellie's phonetic inventory consisted of [p b t d f v s z k g w ɹ j] and she demonstrated age-appropriate phonological development. At a subsequent review at age 2;0, Ellie was producing turbulent alveolar and velar active nasal fricatives [ɲ̥ ɲ̥̃] for fricative and affricate phoneme targets. She did not present with any hypernasal resonance but abnormal nasal airflow (nasal turbulence) was present accompanying her speech; this suggests a marginal degree of VPD. Ellie started therapy at age 2;0 and had nine therapy sessions up to the age of 2;9. Prior to starting therapy, she had been diagnosed with bilateral middle ear effusion following Tympanometry testing and this persisted throughout the therapy period. At 2;5, the findings from Visual Reinforcement Audiometry and McCormick Toy Test resulted in a diagnosis of moderate conductive hearing loss. Ellie was referred to Ear, Nose and Throat at age 2;8 and although bilateral middle ear effusion persisted, her hearing levels had normalised by 2;10.

Thomas

Thomas had a repaired bilateral cleft lip and palate, and underwent surgery at a delayed age of 15 months (1;3) after illness delayed his prior lip surgery. There were no healing complications. In line with the regional protocol, Thomas was seen for his 18-24 month assessment at age 1;6. He had age-appropriate comprehension and, expressively, was at a variegated babbling stage of development, though some early words were emerging. This would suggest a mild delay in this area of development. Thomas' phonetic inventory consisted of [m n] in babble strings and emerging words. No oral pressure consonants were heard; this is a cleft speech characteristic that is strongly associated with VPD. Thomas started therapy at 1;6 and received six sessions up to age 1;11. At 1;3, three months before starting therapy, Thomas was diagnosed with bilateral middle ear effusion, and testing with Visual Reinforcement Audiometry found this was

mildly affecting his hearing. When his hearing was tested again at age 2;1, two months after the therapy period, he had normal to near normal hearing levels.

Hannah

Hannah had a repaired cleft of the soft palate. She underwent surgery at seven months (0;7) of age; this healed well with no palatal fistula. Hannah was seen for her 18-24 month assessment at age 1;6. Her receptive and expressive language skills were developing appropriately and she presented as very communicative and interactive. Hannah's phonetic inventory consisted of [m n l w ɪ j] and possibly a weak [d]. Her speech characteristics were indicative of possible VPD. Hannah started therapy at 1;6 and received nine sessions up to age 2;7. She underwent grommet insertion at 1;7, one month after she started therapy following a history of bilateral middle ear effusion, diagnosis of mild to moderate conductive hearing loss, and suspected tinnitus. At age 1;9, two months following grommet insertion, Visual Reinforcement Audiometry showed normal hearing levels, and again at 2;7, at the end of the therapy period.

Naomi

Naomi had repaired unilateral cleft lip and palate. She underwent surgery at six months (0;6) of age; her palate healed well with no palatal fistula. Naomi was seen for her 18-24month assessment at age 1;8. She was a sociable and communicative little girl with good attention and interaction skills. Her receptive and expressive language skills were developing appropriately. Her phonetic inventory consisted of [m n l j] and vowels were nasalised. These observations are strongly indicative of VPD. Naomi was 2;3 when she started therapy. It would appear from the case notes that her initial speech presentation was suspected to be related to her hearing history. Therapy was recommended following a review at 2;2 when earlier characteristics had not resolved. Naomi presented with persistent and consistent middle ear effusion affecting one or both ears. She was referred to Ear, Nose and Throat at 1;0 and later diagnosed with a mild to moderate conductive hearing loss. She had bilateral myringotomy and grommet insertion at age 1;2. Despite treatment, effusion and hearing loss persisted. At 2;2, one month prior to starting therapy, Naomi was fitted with bilateral hearing aids, which she wore up to and beyond 3;1, spanning the full therapy period.

Table 5.3: Characteristics of child participants

	Louise	Ellie	Thomas	Hannah	Naomi
Cleft diagnosis	Hard and soft palate (Robin Sequence)	Soft palate	Bilateral cleft lip and palate	Soft palate	Unilateral cleft lip and palate
Age at palate repair	0;9	0;6	1;3	0;7	0;6
Fistula present?	No	No	No	No	No
Age at 18-24m assessment	1;11	1;7 and 2;0	1;6	1;6	1;8
Receptive language	Age-appropriate	Age-appropriate	Age-appropriate	Age-appropriate	Age-appropriate
Expressive language	Age-appropriate	Age-appropriate	Mildly delayed	Age-appropriate	Age-appropriate
Resonance and airflow	Normal	No hypernasal resonance; nasal turbulence	Not reported; absent oral pressure consonants	Not reported; limited oral pressure consonants	Nasal tone on vowels; absent oral pressure consonants
Phonetic inventory	[m b w g ŋ x h ʔ]	1;7: [p b t d f v s z k g w ɹ j] and at 2;0: [ŋ ʃ]	[m n]	[m n l w ɹ j] Possibly [d]	[m n l j]
Cleft speech characteristics	Backing to velar	Turbulent active nasal fricatives	Absent oral pressure consonants	Limited strong oral pressure consonants	Absent oral pressure consonants
Age at first therapy session	2;3	2;0	1;6	1;6	2;3
Age at last therapy session	2;11	2;9	1;11	2;7	2;10
Number of therapy sessions	9	9	6	9	9
Middle ear history	Normal at 1;8	Persistent effusion at 0;10, 1;3, 1;8, 2;5, 2;8, 2;10	Bilateral effusion at 1;3	Bilateral effusion at 0;10 and 1;1; recurrent infections; tinnitus	Bilateral effusion at 0;6, 1;0, 1;1, 1;4; unilateral at 0;8, 0;9
Hearing levels	Normal at 0;10 and 1;8	Moderate bilateral loss at 2;5, 2;8; normal at 2;10	Mild loss at 1;3; normal at 2;1	Mild to moderate loss at 1;1	Mild loss at 0;9; mild to moderate loss at 1;0, 1;4; satisfactory at 1;7; moderate loss at 2;2
Hearing management	Surveillance	Surveillance	Surveillance	Grommets at 1;7	Grommets at 1;2; bilateral hearing aids at 2;2

5.4.2 Speech and language therapist (SLT) participants

5.4.2.1 Inclusion and exclusion criteria

Speech and language therapists were eligible for inclusion if they met the following criteria:

- Works as a speech and language therapist in the National Health Service (NHS)
- Provided therapy to a child with CPSSD aged between 18 months and three years
- Therapy took place in the UK
- Therapy took place from 2010 onwards
- Therapy consisted of MSIVT
- A minimum of five therapy records were available (from five separate sessions)
- The therapist who provided the therapy consents to take part in the study

Being a specialist in cleft palate and related disorders was not an inclusion criterion because in the UK therapy is not always provided by a specialist SLT.

The following exclusion criteria applied to the SLT participants:

- The following information about the SLT is not available:
 - Previous training and experience
 - Current role

5.4.2.2 Characteristics of SLT participants

To obtain information about SLTs' roles, training they had received on multisensory input modelling with/without video-therapy and their experience of using the approach, all three SLT participants completed the Therapist Data Collection Form shown in Appendix V.

Sarah

Sarah had been qualified between five and 15 years and was working as a specialist SLT in a regional cleft lip and palate centre. She had received training on the multisensory input modelling approach at local/regional and postgraduate levels, and had discussed the approach in clinical supervision. She had experience of multisensory input modelling with more than 10 children with cleft palate-related speech sound disorder (CPSSD) aged 18 months to three years and between one and five children over the age of three. In addition, Sarah had experience with between five and 10 children with other types of speech characteristics aged 18 months to three years and between one and five children over three.

Laura

Laura had been qualified between five and 15 years and was working as a specialist SLT at a regional cleft lip and palate centre. She had received training on multisensory input modelling at local/regional and postgraduate levels. She had experience of the approach with between one and five children under 18 months with CPSSD, more than 10 children aged 18 months to three years, and more than 10 children over three. In relation to other types of speech characteristics, Laura had experience of using the approach with between one and five children in each of the following age groups: under 18 months; 18 months to three; over three.

Helen

Helen had been qualified between five and 15 years and was working as a specialist SLT at a regional cleft lip and palate centre. She had received training on multisensory input modelling at local/regional and postgraduate levels. She had experience of using the approach with between one and five children with CPSSD under 18 months of age, more than 10 children aged 18 months to three years and between five and ten children over three. Helen did not have experience of using the approach with children with other types of speech characteristics.

5.5 Data analysis

Different methods of analysis were employed in each phase of the study. Phase 1 used descriptive methods to describe the delivery of MSIVT in the NHS; the results of the analysis are presented in Chapter 6. Phase 2 drew on conversation analysis to examine therapist-child interaction during MSIVT; the results of the analysis are presented in Chapter 7. Phase 3 combined conversation analysis and a quantitative method to examine therapist-child interaction over the course of a therapy episode; the results are presented in Chapter 8.

5.5.1 Phase 1: Delivery of MSIVT in the NHS

This phase of the study used the complete dataset, which consisted of 573 minutes and 10 seconds of video recordings and corresponding case note records from five therapist-child dyads at two specialist SLT services in the NHS.

5.5.1.1 *Descriptive analysis using metadata*

A modified version of the procedure developed in the pilot study was used to catalogue the data and prepare it for analysis. In the pilot study, videos corresponded to activities; in the

main study, videos corresponded to sessions. Analysis began with the production of metadata from the therapy datasets representing each of the five MSIVT episodes of care. This process was completed by watching the therapy videos in consecutive order and reviewing the SLT case note records. There was no limit on the number of times a video could be watched. The metadata were analysed within-session and session-by-session for each therapist-child dyad to examine the delivery characteristics of each episode of care. The metadata were then analysed episode-by-episode to examine variation in delivery across the five episodes.

Structure of MSIVT episodes of care

The episodes of care were analysed to identify the number, duration and frequency of MSIVT sessions, where sessions took place and who was present. The duration of a session was operationalised by the length of the corresponding therapy video. The MSIVT sessions were analysed to identify the number and duration of MSIVT activities. The duration of an activity was operationalised by the length of time between the closure of one activity and the onset of the following activity. Onset was identified by the onset of verbal talk by the SLT in which she introduces a new activity. Closure was identified by the onset of verbal talk by the SLT in which she closes an activity. Any time in-between was excluded, therefore the duration of a session does not always equate to the summed duration of its composite activities.

Nature of MSIVT activities

The nature of activities within the sessions was analysed by synthesising the SLT's written description of the activity and watching the therapy video. Informed by existing accounts of the MSIVT approach (Calladine, 2009; Harding & Bryan, 2000) and the pilot study, this involved making a distinction between toy-based and paper-based activities, and identifying the type or name of toy (if used) and whether or not pictures were used.

Nature of target selection and stimulus design

The nature of speech sound targets and stimuli within activities were analysed to determine the target selection and stimulus design characteristics of the episodes of care. A speech sound target is a phoneme of the child's native linguistic system that is targeted in therapy. Speech sound targets were identified from SLT written descriptions and author video observations. These were analysed to identify which, and how many, phonemes were targeted in each session and activity, and examine patterns within and across MSIVT episodes of care. A speech sound stimulus is a phone(s) produced by the SLT as a stimulus for the target phoneme. Stimuli

orientate to phoneme targets. Harding and Bryan (2000) and Calladine (2009) describe the use of novel sounds with adaptive articulation, e.g. interdental stimuli for target alveolar phonemes. In the pilot study, the SLT at times produced different phones for the same target phoneme. This reflects the principle of the approach to design stimuli that will stimulate non-lexical speech processing, i.e. generation of new motor programs rather than retrieval of existing motor programs for known phonemes (Calladine & Vance, 2019; Harding & Bryan, 2000). For these reasons, it was important to make a distinction between speech sound targets and stimuli. Speech sound stimuli were identified from video observation and transcribed phonetically using the International Phonetic Alphabet (IPA; 2018) and ExtIPA (Ball et al., 2018). Where more than one stimulus was used for the same target phoneme, all of the different productions were transcribed. In English, the voiceless plosives /p t k/ are either aspirated or unaspirated depending on their phonological context; they are aspirated at the beginning of a word and when they are in front of a stressed syllable. For the purposes of this study, both aspirated and unaspirated versions were transcribed as [p t k] to prevent over-counting.

As well as the number and type of speech sound targets and stimuli, data were analysed to identify the linguistic level, i.e. the verbal context in which the stimulus was produced and the target position it orientated to. This is important because of the implications that different levels have for the child's speech processing. Linguistic level was identified from SLT written descriptions and author video observations. It was documented according to the following descriptions in the existing literature and in the pilot study: single sound; sound sequence; invented word; and real word; see definitions below (Calladine, 2010; Calladine & Vance, 2019; Gierut et al., 2010; Harding & Bryan, 2000; Harding & Grunwell, 1998; Williams et al., 2010). At every level, the stimulus can occur as part of a short phrase or sentence.

- **Single sound:** A phone or phoneme produced in isolation
- **Sound sequence:** A consonant (C) and vowel (V) combination, where the stimulus may occur in any position, e.g. CV, VC, VCV
- **Invented word:** A sound sequence with meaning assigned, e.g. the name for a character, or a word that does not exist in the child's language system
- **Real word:** A word that exists in the child's language system and is understood to be known to the child; the stimulus may occur in any position, e.g. initial (WI), medial (WM) or final (WF)

Child stimulus productions

To determine if child stimulus productions were a characteristic feature of the five MSIVT episodes of care, therapy videos were analysed to identify the presence or absence of stimulus productions and document this on the metadata. The metadata were analysed within and across sessions to identify how many sessions and activities featured productions and when they emerged in each episode of care. Within each episode of care, the number of activities featuring productions provided a measure of the child's rate of production. The number and accuracy of productions were not recorded.

5.5.2 Phase 2: Interactional features of MSIVT

This phase of the study used a subset of data to allow detailed analysis of therapist-child interaction within the MSIVT activities. It consisted of 120 minutes of video recordings from four therapist-child dyads at three specialist SLT services in the NHS.

5.5.2.1 Sampling the data for analysis

Since the study was interested in identifying the ways SLTs support children with CPSSD during MSIVT, it was important to use data from as many SLTs as possible. One of the SLTs, Helen, featured in three of the episodes. To ensure comparable data across the three SLTs, only one of Helen's episodes was included. The Helen-Naomi episode was purposively selected on the basis that it was more similar in delivery to the episodes featuring Sarah-Louise and Laura-Ellie. To extend the data to include a fourth therapist, one of the episodes from the pilot study was also included. A review of child participant and episode delivery characteristics of the pilot study and Phase 1 data did not identify a stronger justification for one child over the other, therefore the episode Sam-James was randomly selected using the online generator <https://www.random.org>.

Phase 1 of the study revealed that speech sound targets and stimuli changed over the course of a therapy episode of care. It was important that any potential differences in therapist-child interaction over the course of an episode were captured in Phase 2 of the analysis so that the interactional features identified reflect the approach broadly and not specifically at a particular stage in therapy. (This is the focus of Phase 3). Therefore, purposive sampling was used to select the first, middle and last videos of each of the four MSIVT episodes, 12 videos in total, for analysis in Phase 2. In order to allow detailed examination of the video data and ensure the datasets across episodes were comparable, a further stage of sampling was conducted.

Informed by the sampling procedures in other studies of interaction in speech and language therapy, such as Wilkinson, Gower, Beeke and Maxim (2007), 10-minute excerpts of consecutive therapy talk within MSIVT activities were identified in each of the 12 videos, beginning at the time 01:00. Since the study was interested specifically in the interaction within an MSIVT activity, verbal and nonverbal talk and action relating to setting up and closing down an activity were excluded in order to produce 30 minutes of continuous therapy talk per dyad (Panagos et al., 1986). See Table 5.4 for the names of videos selected from each dyad and the ages of the children in these sessions.

Table 5.4: The video data for Phase 2 analysis

Video information		Sarah-Louise	Laura-Ellie	Helen-Naomi	Sam-James
FIRST	Name	V1	V1	V1	V1
	Age of child	2;3	2;0	2;4	1;11
MIDDLE	Name	V4*	V4*	V3	V8*
	Age of child	2;6	2;5	2;6	2;4
LAST	Name	V7	V6	V5	V14
	Age of child	2;10	2;9	2;9	2;7

Note. Ages shown in years and months. V=video. Videos that are starred (*) were randomly selected from the two middle videos in the episode using the online random generator <https://www.random.org>.

5.5.2.2 Conversation analysis

Interaction can be analysed using different methodologies, as shown by the review of the literature in Chapter 4. The studies by Gardner (1994) and Ronkainen (2011) used the well-established method of conversation analysis (CA), which has its origins in ethnomethodology (Heritage, 1984). Their research demonstrates the potential of CA to uncover interactional features that would not have been visible by other methods. Gardner (1994) found that subtle differences in the way therapists produced target words during phonology therapy differentiated between words set up as a model for imitation and words set up as a redoing to close a sequence. These included subtle differences in prosody and use of syntax as well as nonverbal features, namely eye gaze. Ronkainen's (2011) application of CA to the study of auditory-verbal therapy revealed the multimodal nature of the SLT's interaction with the child when establishing their attention and introducing new objects to elicit vocalisation.

The present study was interested in examining interaction during MSIVT in order to identify the practices SLTs use to establish a child's attention and participation in order to expose them to speech sound stimuli and create opportunities for production. The findings of the literature review on MSIVT and interaction (Chapters 2 and 4) suggest a potentially significant dimension of the interaction in MSIVT is between the therapist and the camera. Application of Goffman's (1981) theories of 'frames' and 'footing' to Calladine and Vance's (2019) findings has raised the consideration that therapists use the camera to facilitate different frames of interaction. It was important in studying the nature of this unusual method of therapy to try and capture how therapists manage this potentially complex manner of interaction. This required a highly sophisticated method of analysis. The studies by Gardner (1994) and Ronkainen (2011) suggested the detailed nature of CA and its ability to capture the complexities as well as the subtleties of talk made it the most suitable methodological choice for studying the interactional features of MSIVT. Conversation analysis is a qualitative data-driven approach for studying social interaction between people (Goodwin & Heritage, 1990). Herein lay a problem if some of the interaction that takes place during MSIVT is with the camera, an inanimate object. However, since existing descriptions of MSIVT suggest the camera facilitates a form of interaction with the child (through the camera), albeit at a future moment in time, and in all other respects CA was the most suitable choice, it was adopted for the study.

Central to CA is the notion that interaction is the accomplishment of its participants, whose shared understanding or knowledge allow them to interpret the speech and actions of others and what and when to speak or act themselves; this is referred to as 'intersubjectivity' (Sidnell, 2010). A fundamental assumption of CA is that utterances are understood primarily according to their sequential placement, so the basic unit of analysis is a sequence consisting of a minimum of two turns, a first-pair part and a second-pair part (Sidnell, 2010). Unlike in discourse analysis, the study of talk using CA is not led by pre-existing categories or theories and anything in the data is potentially of interest. Accordingly, no verbal or nonverbal talk should be excluded before it has been analysed, and categories or patterns only emerge from meticulous analysis of all of the data (Hutchby & Wooffitt, 2008). This is particularly important for studying interaction during MSIVT because it has not been studied before. In CA, the data itself provides the evidence. Therefore, the action of a speaker's turn is understood by its consequence, i.e. how the listener responds in their turn and, if a response is not forthcoming or appropriate, by the speaker's pursuit (Sidnell, 2010; Stivers & Rossano, 2010).

Conversation analysis focuses on specific aspects of conversation and the practices within them, such as turn-taking, action and understanding, and repair. Much of Gardner's (1994; 1997; 1998; 2007) research focused on repair. She examined the ways SLTs (in comparison to mothers) responded to children's erroneous productions to initiate a re-try and help the child progress towards self-initiated repair. Ronkainen's (2011) study was concerned with action and understanding. Close examination of sequences in which the SLT introduced a new object differentiated between turns that sought the child's attention and those that sought vocalisation.

Detailed transcripts of the therapy talk data were produced for each of the 12 10-minute excerpts using a method derived from Jefferson's (2004) system of conventions; see Appendix W. Given the nature of the conversational context, speech sound stimuli were transcribed phonetically using the IPA (2018) and ExtIPA (Ball et al., 2018). Volume changes and hesitations were not transcribed, but length of pauses was, where necessary, as were overlaps in talk. Additional annotations were used to capture nonverbal details, such as eye gaze, gesture and manipulation of toys and pictures. Transcripts were produced through repeated viewings of the video data and continuous refinement. The process of producing the transcripts formed a central part of the analysis. The video data and written transcripts were examined to identify the different types of action SLT turns fulfilled and the specific practices SLTs used to establish the child's attention to, and create opportunities for production of, speech sound stimuli. In line with CA principles, the action of a therapist's turn was identified by analysing its design (verbal, nonverbal and prosodic features) and the nature of child response it elicited. However, since a listener can choose whether or not to respond to a speaker (Drew, 1981) it was also necessary to look for evidence of the action in the SLT's evaluation (acceptance or rejection) of the child's response.

Whilst the primary focus of the analysis was the nature of initiation in SLT first-pair part turns, it also included analysis of the child's response. This was done not just to understand the nature of action being accomplished, but also to explore how the child responds to therapy in the local context of interaction. It also gave insight into the structural pattern of interaction. An additional focus, to gain preliminary insight into SLT repair practices during MSIVT, was the design features of SLTs' third turn evaluations when the child made a stimulus production.

5.5.3 Phase 3: MSIVT interactions over an episode of care

This phase of the study used one episode selected by purposive sampling to analyse therapist-

child interaction from a longitudinal perspective over the course of an episode of care. It consisted of two 10-minute excerpts of video, one each from the first and last video-therapy sessions.

5.5.3.1 Sampling the data for analysis

The author excluded her own data (Sam-James) from this phase of the study. Purposive sampling selected the Sarah-Louise dyad because the analysis in Phase 2 showed this episode of care most closely resembled existing descriptions of the MSIVT approach. The aim of Phase 3 was to examine changes in interaction practices over time, so the first and last videos of the episode were purposively selected for comparative analysis. To allow a detailed level of analysis a 10-minute excerpt was extracted from each video, thereby producing two 10-minute excerpts of continuous therapy talk for analysis. To ensure the excerpts were as comparable as possible they were selected from the same point within each video, beginning at 01:00, as in Phase 2. Verbal and nonverbal talk relating to setting up and closing down an activity was excluded.

5.5.3.2 A mixed methods approach combining CA and quantitative analysis

Longitudinal studies of social interaction using CA is a relatively new area of research that offers new insights but is not without methodological challenges (Wagner, Pekarek Doehler & González-Martínez, 2018). One approach is to study the process of emergence or change of a particular type of practice in accomplishing a particular type of action over a period of time (Wagner et al., 2018). The availability of two videos, one of the therapist and child (age 2;3) at the start of therapy, and the other six months later (child aged 2;10) as the episode of care draws to a close, provided the unique opportunity to compare two distinct moments in time.

The video data and detailed written transcripts of the first and last videos in Sarah-Louise's episode were examined side by side and all of the SLT initiated stimulus-related sequences of therapy talk occurring in the two 10-minute excerpts were isolated. Each sequence was examined to identify its turn-constructive units and the nature of action and response that unfolded within it. This phase of the analysis used a different method to Phase 2. The SLT's first-pair part turns were examined in relation to the pre-determined types of action that had been excavated from detailed analysis of the data in Phase 2. An innovative application of quantitative analysis was used to compare and contrast the interactional profiles of the first and last videos in relation to frequency and distribution of different types of action.

Conversation analysis was then used to examine specific actions that occurred in both videos in order to identify any differences in the way they were designed and delivered as well as the nature of child responses they elicited. Sequences involving child stimulus productions were analysed in relation to psycholinguistic theories of speech processing (Stackhouse & Wells, 1997).

5.6 Summary

This chapter has provided detailed explanations of the methods used to collect, organise and analyse data and the rationale behind them. The study recruited five therapist-child dyads comprising three SLTs and five children. The children received MSIVT as part of their routine clinical care in two specialist SLT services in the NHS. The children were aged 1;6 to 2;3 when therapy commenced and 1;11 to 2;11 when it concluded. Data were generated from children's pre-existing clinical records and included 573 minutes of video recordings. Mixed methods were used to analyse the data, which included the well-established and systematic approach of conversation analysis. The following three chapters present the results: Chapter 6 presents the results of Phase 1; Chapter 7 presents the results of Phase 2; and Chapter 8 presents the results of Phase 3.

Chapter 6: Delivery of MSIVT in the NHS (Phase 1)

6.1 Introduction

This chapter presents the results of Phase 1. This phase of the study aimed to examine the delivery of multisensory input video-therapy (MSIVT) as an episode of care in order to describe the nature of sessions, activities, speech sounds targets, speech sound stimuli and materials that characterise MSIVT.

The data were analysed within-session and session-by-session within and across the episodes. The findings of the analysis are presented below, firstly in relation to the features of the five MSIVT episodes of care, and secondly, in relation to variations in delivery.

6.2 Features of MSIVT episodes of care

6.2.1 Structure of MSIVT episodes of care

Table 6.1 gives details about the number, duration and frequency of sessions within the five episodes of care. Mean frequency (in row two) relates to the occurrence in time of consecutive sessions, e.g. in the Helen-Thomas episode of care, sessions took place, on average, every 4.2 weeks.

Table 6.1: Summary of MSIVT episodes of care

	Sarah-Louise	Laura-Ellie	Helen-Thomas	Helen-Hannah	Helen-Naomi	Totals
No. of sessions	9	9	6	9	9	42
Mean frequency of sessions (range)	5.1 (3-9.3)	5.2 (2-10.3)	4.2 (2.3-5.1)	5 (1.5-8.6)	4 (1.4-6.4)	4.2 (1.4-10.3)
Age of child at first/last session	2;3/2;11	2;0/2;9	1;6/1;11	1;6/2;7	2;3/2;10	1;6/2;11
Length of the episode (in months)	8	9	5	11	7	NA
No. of MSIVT sessions	8	6	5	6	7	32
No. of videos available	7	6	5	6	5	29
Mean duration of MSIVT sessions (range)	24:00 (18:29-29:46)	15:56 (10:23-18:52)	11:57 (08:52-15:42)	15:03 (09:12-20:41)	31:53 (25:39-37:00)	19:46 (08:52-37:00)
People present	Mum +/- sibling	Mum or dad	Mum	Mum	Mum	NA
No. of MSIVT activities	35	23	17	22	25	122
Modal no. of activities per session (range)	5 (4-6)	3 (3-5)	3 (2-5)	3 (3-5)	5 (4-6)	5 (2-6)
Mean duration of MSIVT activities (range)	04:38 (01:19-09:12)	04:24 (01:37-06:38)	03:09 (01:05-06:06)	04:18 (0:47-07:22)	06:09 (03:26-12:20)	04:28 (0:47-12:20)

Note. No.=number. MSIVT=multisensory input video-therapy. There is a discrepancy between the number of sessions (row one) and the number of MSIVT sessions (row four) because SLTs did not make a therapy video in every session. There is also a discrepancy between the number of MSIVT sessions (row four) and the number of videos available (row five) because of factors such as the video did not meet criteria for inclusion. Video duration is in minutes and seconds. Ages are in years and months. NA=not applicable. +/- =with/without.

Episodes of care most commonly consisted of nine therapy sessions (Sarah-Louise; Laura-Ellie; Helen-Hannah; Helen-Naomi). There was little variation in the number of sessions per episode. The shortest episode (Helen-Thomas) consisted of six sessions; the other four episodes all consisted of nine sessions. Sessions took place, on average, approximately monthly. The children ranged in age from 1;6 (Thomas and Hannah) to 2;3 (Louise and Naomi) when therapy commenced and 1;11 (Thomas) to 2;11 (Louise) when it concluded. The length of the episode ranged from five months (Helen-Thomas) to 11 months (Helen-Naomi). In every episode, at least one session did not include video-therapy and was therefore not an MSIVT session. The majority (32/42; 76%), however, were MSIVT sessions, ranging from 67% (Laura-Ellie and Helen-Hannah) to 89% (Sarah-Louise). Of the 29 complete datasets available, the average duration of an MSIVT session was a little under 20 minutes, though this ranged from 11:57 (Helen-Thomas) to 31:53 (Helen-Naomi). The shortest MSIVT session was 08:52 (Helen-Thomas) and the longest was 37:00 (Helen-Naomi). Sarah-Louise was the only episode to feature a sibling in some of the sessions. One parent was present in all of the 29 sessions.

In the 29 complete datasets available, 124 activities took place across the five therapy episodes. Two of these activities (one in Helen-Thomas and one in Helen-Hannah) were abandoned by the SLT and were therefore excluded from the analysis, leaving 122 activities. An MSIVT session commonly consisted of five therapy activities, ranging from two (Helen-Thomas) to six (Sarah-Louise and Helen-Naomi). The average duration of an MSIVT activity was approximately four and half minutes, ranging from 03:09 (Helen-Thomas) to 6:09 (Helen-Naomi). The shortest MSIVT activity was 0:47 (Helen-Hannah) and the longest was 12:20 (Helen-Naomi). The findings in relation to the nature of activities are presented in the following section.

6.2.2 Nature of MSIVT activities

Therapists described the nature of activities in different ways. Most commonly, in terms of the toy, game or pictures used with or without reference to what the activity required of the child, for example:

- "...blowing activity with the straw..." (Sarah+Louise-S1-V1)
- "Started video-therapy...activities: i) cotton reels..." (Laura+Ellie-S1-V1)
- "Posting of food cards plus 'bye bye'" (Helen+Thomas-S3-V2)
- "Tower building /t/ in isolation, plus /t/ in tower, your turn and tall" (Helen+Hannah-S7-V5)

- “Hiding game with WF /p/ pictures” (Helen+Naomi-S3-V1)
- “Listening activity for /p, /t/ and ‘sh’ with monkey tree game” (Helen+Naomi-S5-V3)
- “Sorting... sounds into cups” (Helen+Naomi-S5-V3)

In three of the episodes, the SLT described activities as “input modelling activities” or “input modelling therapy” (Sarah-Louise, Helen-Thomas and Helen-Hannah). In four of the episodes (Laura-Ellie, Helen-Thomas, Helen-Hannah and Helen-Naomi), the videoing aspect of therapy was described as “video-therapy” or “therapy DVD.” In two of the episodes (Helen-Thomas and Helen-Hannah), activities were described as “adult-led”. Also appearing across the episodes were explicit descriptions of the SLT’s role in the activities, for example:

- “Modelled /t/ in isolation and /t/ words in play” (Sarah+Louise-S4-V4)
- “Blowing bubbles using an oral airflow” (Sarah+Louise-S5-V5)
- “Bubble wobbling with /t/” (Helen+Thomas-S4-V3)
- “Repetitive modelling of /b/ in ‘baby’ in puzzle with baby animals” (Helen+Hannah-S6-V4)
- “Adult models of sounds as monkeys hung up” (Helen+Naomi-S4-V2)

Table 6.2 gives details of the types of activities that took place.

Table 6.2: Types of activities in the MSIVT sessions

	Sarah-Louise	Laura-Ellie	Helen-Thomas	Helen-Hannah	Helen-Naomi	Totals
No. of activities	35	23	17	22	25	122
% toy-based	71 (25/35)	100 (23/23)	100 (17/17)	86 (19/22)	80 (20/25)	85 (104/122)
% paper-based	11 (4/35)	0	0	14 (3/22)	8 (2/25)	7 (9/122)
% pictures	37 (13/35)	9 (2/23)	24 (4/17)	18 (4/22)	68 (17/25)	33 (40/122)

Note. No.=number. %=proportion, shown as a percentage. Row number four gives the proportion of activities that featured pictures.

All 122 MSIVT activities across the five MSIVT episodes of care were play-based and therapist-led. The analysis shows the majority (85%; 104/122) of activities were toy-based, ranging from 71% (Sarah-Louise) to 100% (Laura-Ellie and Helen-Thomas). Therapists used a variety of toys across the episodes of care and some featured multiple times within the same episode. For

example, Wooden Click Clack Track featured in 5/7 of Sarah-Louise's MSIVT sessions. Bubble play featured in four of the five episodes (Sarah-Louise, Laura-Ellie, Helen-Thomas, Helen-Hannah), and often in several sessions, e.g. 4/6 of Laura-Ellie's sessions. Only 7% (9/122) of MSIVT activities were paper-based, ranging from none at all (Laura-Ellie and Helen-Thomas) to 14% (Helen-Hannah). Examples of activities categorised as paper-based are: painting pictures of balloons (Sarah+Louise-S3-V3-A4); sticking stickers (for drips) on a picture of a tap (Sarah+Louise-S4-V4-A1); drawing sound symbols on a picture of a hand (Helen+Naomi-S8-V6-A2); and sticking stickers (for spots) on a picture of a ladybird (Helen+Naomi-S3-V1-A2). Five of the activities in Sarah-Louise's therapy episode could not be categorised as toy- or paper-based. Three activities involved blowing a piece of tissue across the table or off the hand (to demonstrate oral airflow); one involved the game Hide and Seek, where the child was asked to find pictures hidden around the room; and one involved the game Speech Stickers on the iPad. Pictures featured in 33% (40/122) of activities. In Sarah-Louise and Helen-Naomi's episodes, this included pictures to represent the novel speech sound stimuli, e.g. sleeping baby to represent [ʃ] (Helen-Naomi), dripping tap to represent [t̪] (Sarah-Louise), as well as pictures of real word items, such as things that end in /ʃ/, e.g. wash, brush (Helen-Naomi), things that begin with /t/, e.g. t-shirt (Sarah-Louise).

6.2.3 Nature of target selection and stimulus design

Speech sound targets were identified in 26/29 (90%) of MSIVT sessions and 108/122 (89%) of MSIVT activities within them. It was not possible to identify phoneme targets in 14/122 (11%) of activities. In three of these (all in the Sarah-Louise episode), the SLT documented "blowing... using an oral airflow", suggesting the target was an oral airstream and not a specific phoneme, although it may have been a precipitator for the phoneme /f/. In the remaining 11 videos, the SLT produced multiple stimuli without reference to a specific phoneme. Table 6.3 gives details of the speech sounds targeted in the MSIVT sessions and the stimuli SLTs used.

Table 6.3: Speech sound targets and stimuli

	Sarah-Louise	Laura-Ellie	Helen-Thomas	Helen-Hannah	Helen-Naomi	Totals
Phoneme targets	/f t d s/	/f s θ ʃ/ (/ð z tʃ/)	/p b t d n l n ʃ/	/p b f t d s ʃ/	/p f t ʃ k/	/p b f θ t d l n s ʃ k/ (/ð z tʃ/)
Modal no. of targets per MSIVT session (range)	2 (2-3)	2 (2-3)	No mode (0-7)	1 (0-3)	3 and 5 (1-5)	2 (1-7)
Modal no. of targets per MSIVT activity (range)	1 (1-3)	1 (1-3)	1 (1-4)	1 (1-2)	1 (1-5)	1 (1-5)
Sound stimuli	[f: φ: t̥ t̥ t̥ t̥ d̥ d̥ s: s:]	[f: φ θ̥ s̥ ʃ: tʃ:] Often weakened	[p b f: t̥ t̥ d̥ d̥ n̥ l̥ s: ʃ:]	[p b f: t̥ t̥ d̥ s: ʃ:] Weakened fricatives	[p f: t̥ t̥ ʃ: k] Often weakened	Weakened and lengthened fricatives; interdental and linguolabial placements; weakened plosives and fricatives
Linguistic level: S; SS; IW; RW	S, SS, RW – all sounds IW – [f:]	S, SS, RW – [θ̥ s̥ ʃ:] S – [f:]	S – [p t̥ t̥ ʃ:] RW – all sounds	S – [t̥ t̥ d̥ ʃ:] SS – [ʃ:] RW – all sounds	S – all sounds SS – all sounds, except [k] RW – all sounds except [f:]	S, SS, RW, IW

Note. No.=number. MSIVT=multisensory input video-therapy. S=single sound; SS=sound sequence; IW=invented word; RW=real word. Sounds have been transcribed using IPA (2018) and ExtIPA (Ball et al., 2018).

6.2.3.1 Target selection

All five MSIVT episodes of care featured multiple phoneme targets. There was little variation between the episodes in the number of phonemes targeted. Eleven different phonemes were targeted in the 29 MSIVT sessions across the five episodes of care: /p b f θ t d l n s ʃ k/, ranging from four (Sarah-Louise) to eight (Helen-Thomas). Three of the phonemes listed as featuring in Laura-Ellie's episode are in brackets: /ð z tʃ/. These phonemes were not targets of therapy, but the therapist inconsistently produced them with adaptive articulation when they appeared in non-target words. There was also little variation in the types of phonemes targeted. The most common phonemes were /f t ʃ/, which each featured in four of the five episodes. The least common phonemes were /l n k/, which each featured in just one of the five episodes (Helen-Thomas and Helen-Naomi). Nine of the 11 (82%) phonemes are obstruents and include a variety of voiced and voiceless plosive, fricative and affricate consonants across multiple places of articulation. Voiceless consonants outnumbered voiced consonants; seven compared to four. Only two of the phonemes targeted are sonorants, /l n/, and they featured with very low frequency, in just one (3%) of the 29 videos (Helen-Thomas-S3-V2). Within the episodes of care there were some common class relationships between phoneme targets. For example, four of the seven phonemes in Sarah-Louise's episode, /p b t d/, share the same manner of articulation (they are all plosives). In Laura-Ellie's episode, all four phonemes belong to the fricative class of consonants and all are voiceless. The phonemes targeted in Helen-Naomi's episode are all voiceless consonants but represent five different places of articulation, ranging from bilabial to velar.

Most commonly across the five episodes of care, two phonemes were targeted per MSIVT session, though this ranged from none (where a phoneme target could not be identified; Helen-Thomas-S1-V1 and Helen-Hannah-S1-V1) to seven (Helen-Hannah-S3-V2). Within a session, SLTs typically targeted one phoneme per MSIVT activity; this was the case in 84% (91/108) of activities across the episodes. Nearly all (30/32; 94%) of the activities in Sarah-Louise's episode had one target. Only two activities in this episode involved multiple phoneme targets: fishing game with /f t s/ single sound pictures in video three (Sarah+Louise-V3-A5) and hide and seek with /t d/ single sound pictures in video seven (Sarah+Louise-V7-A1). More variability was observed in the other four episodes, though one target per activity was still the most frequently observed phenomena: 15/23 (65%) in Laura-Ellie; 10/17 (59%) in Helen-Thomas; 14/22 (64%) in Helen-Hannah; and 18/25 (72%) activities in Helen-Naomi.

6.2.3.2 Stimulus design

In all five MSIVT episodes of care SLTs used adaptive articulations in their stimulus productions. The types of articulation features they used and the stimuli they applied them to were also similar across episodes. Interdental voiced and/or voiceless plosives featured in all five episodes of care and interdental fricatives in two (Sarah-Louise and Laura-Ellie). Linguolabial plosives featured in three of the episodes, often interchangeably with interdental and, in some cases, alveolar plosives. Linguolabial fricatives did not feature in any of the therapy episodes. Lengthened fricatives featured in all five episodes of care. At times, a combination of dental or interdental articulation and lengthening was observed, e.g. use of the stimuli [ʃː] and [ʒː] in Sarah-Louise and Laura-Ellie's episodes, respectively. These specific articulation features make certain aspects of sounds more salient, e.g. oral airstream (lengthened articulation), place of articulation (linguolabial placement). Weakened articulation was observed in three of the five episodes (Laura-Ellie, Helen-Hannah and Helen-Naomi), typically on fricative production but also on production of the voiceless plosive [p]. However, whilst all of the video recordings were good audio quality, SLTs Laura and Helen used an external microphone during filming, and this did produce some perceptible differences on playback. Transcription of weakened articulation in this dataset is therefore not deemed to be reliable. In all five episodes of care, SLTs made cued articulation (Passy, 1993) finger gestures alongside stimulus production. They accompanied single sounds, e.g. [ʃː] (Laura-Ellie), sounds within sequences, e.g. [ʧɑ] (Sarah-Louise) and sounds within real words, e.g. [p] in /tɒp/ (Helen-Naomi).

The use of multiple speech sound stimuli also featured in all five MSIVT episodes of care. Excluding weakened articulations, 19 different speech sound stimuli were identified across the five episodes of care, which orientated to 11 different phonemes. Some phonemes had one stimulus, e.g. stimulus [b] was used for phoneme /b/. For other phonemes, e.g. /t/, the author observed multiple stimuli across the sessions, e.g. [t̪ t̪̥ t̪̥ t̪̥ t̪̥]. In these instances, stimuli varied subtly on one feature, e.g. place of articulation, as seen in the linguolabial, interdental, dentalised and alveolar plosive stimuli used for phoneme /t/. It was not possible to delineate the specific phoneme targets for all of the stimuli in Laura-Ellie's episode because the voiceless interdental fricative stimulus [θ̪̥] was used for multiple phonemes: /s z θ ð/ in target and non-target words. For example, in Laura+Ellie-V1-A2, this stimulus is produced as a single sound in a bubble wobbling activity, and the same stimulus is then produced for the phonemes /θ s z/ in the words 'thank you', 'lots', 'bubbles', and 'zebra'.

6.2.3.3 Linguistic level

The analysis revealed that, across all five episodes of care, SLTs produced stimuli at multiple linguistic levels. Only a quarter (7/29; 24%) of sessions featured just one level only. Therapists also combined different levels within the same activity, though this was less typical. Single sound level featured in the majority (25/29; 86%) of MSIVT sessions and in all five episodes of care, as shown in Table 6.3. Within these, it featured in approximately half (61/119; 51%) of all activities (excluding the three ‘oral airstream’ activities). This made single sound the most common level of production. Stimuli corresponding to the following eight phonemes were produced at single sound level: /p f θ t d s ʃ k/. Real word level productions also featured in a majority 24/29 (83%) of MSIVT sessions and in all five episodes of care. Also, similarly to single sound level, it featured in approximately half (56/119; 47%) of all activities. Therapists produced stimuli corresponding to the following 10 phonemes at real word level: /p b f θ t d l n s ʃ/. In real word level productions, stimuli occurred in a variety of different word positions; see Table 6.4.

Table 6.4: Position of stimulus in real word level productions

	Word position		
	Initial	Medial	Final
Number of sessions	19/29	11/29	18/29
Proportion (%)	66	38	62

Some sessions featured more than one different position, hence the number of sessions in each word position adds up to more than 29. Word-initial and word-final positions featured at a similar level (approximately two thirds of all sessions) and were more popular than medial position, which featured in approximately one third of sessions. Stimulus productions at sound sequence level featured in four of the five episodes of care though in fewer than half (12/29; 41%) of MSIVT sessions and in just 16/119 (13%) of activities. Stimuli corresponding to the following six phonemes were produced at sound sequence level: /f θ t d s ʃ/. Stimulus productions in invented words featured less frequently still, in just one session and in one activity in the entire therapy dataset (Sarah+Louise-V5-A4).

6.3 Variation in the delivery of MSIVT

6.3.1 Structure of MSIVT episodes of care

The analysis revealed variation in the duration of sessions and activities within the five episodes of care. The mean duration of sessions ranged from 11:57 (Helen-Thomas) to 31:53 (Helen-Naomi). This does not appear to relate to the SLT since the same SLT features in both of these episodes. The three shortest sessions featured in the episodes Helen-Thomas (08:52), Helen-Hannah (09:12) and Laura-Ellie (10:23). The three longest sessions featured in the Helen-Naomi (37:00) and Sarah-Louise (29:46) episodes. This variation may relate to the age of the child since the three children who received the shortest sessions (Thomas, Hannah and Ellie) started therapy at a younger age (1;6-2;0) than the two children who received the longest sessions (Louise and Naomi, both 2;3). Although the mean duration of activities within sessions was similar across the episodes of care, the duration of individual activities also varied, with a range of 0:47 (Helen-Thomas) to 12:20 (Helen-Naomi). Again, this does not appear to relate to the SLT since the same SLT featured in the shortest and longest activities. It may relate to the age of the child; the three youngest children (Thomas, Hannah and Ellie) had the shortest activities and the two oldest children (Louise and Naomi) had the longest activities.

6.3.2 Nature of MSIVT activities

The nature of activities varied in respect of materials used. Paper-based activities were not used at all in Laura-Ellie and Helen-Thomas's episodes of care, whereas they made up a small proportion of each of the other three episodes. Some of this variation may relate to which SLT provided the therapy since none of the activities led by Laura were paper-based. However, SLT Helen, who did not use paper-based activities with Thomas, did use them with Hannah and Naomi. Similarly, the analysis revealed a range of 9% (Laura-Ellie) to 68% (Helen-Naomi) in relation to the use of pictures in MSIVT activities. These two episodes feature two different SLTs (Laura and Helen), so this may be an accounting factor. Helen used fewer pictures in her two other episodes featuring Thomas and Hannah. These children were both younger when they started therapy (1;6) than Naomi, who was 2;3, which may have been a factor in Helen's picture preferences.

6.3.3 Nature of target selection and stimulus design

Whilst there were similarities in respect of the number and type of phoneme targets and stimuli, SLTs varied in how they introduced them across the episodes of care. See Table 6.5 for within-episode analysis of the sequence of target and stimuli selection.

Table 6.5: Sequence of target and stimuli selection in the five episodes of care

Video	Activity 1	Activity 2	Activity 3	Activity 4	Activity 5	Activity 6
Sarah-Louise						
1	t	t	s	s	Airflow	
2	t	t	s	s	Airflow	
3	t	t	f	f	f t s	
4	t	t	s	s		
5	Airflow	f	f	f	s	s
6	s	s	t	t		
7	t d	d	d	d	d	
Laura-Ellie						
1	f θ/s	f θ s	f θ s ʃ			
2	f θ/s ʃ	f θ s	θ/s			
3	ʃ	θ s	θ s	ʃ		
4	θ/s	ʃ	θ s ʃ	θ s	s	
5	θ s	ʃ	ʃ	ʃ	ʃ	
6	ʃ	ʃ	θ s			
Helen-Thomas						
1	Variety	Variety				
2	Variety	Variety	Variety			
3	t	t	p t	t		
4	b	ʃ	b	b	b	
5	ʃ	ʃ	ʃ			
Helen-Hannah						
1	Variety	Variety	Variety			
2	Variety	Variety	p			
3	Variety	Variety	ʃ			
4	b	b	b			
5	t	t	t	p t	t	
	ʃ	ʃ	ʃ	ʃ	ʃ	
Helen-Naomi						
1	p f t ʃ k	p	p	p	p	
2	p f t ʃ k	p t	t	t	t	
3	p t ʃ	p t ʃ	ʃ	ʃ	ʃ	ʃ
4	p ʃ k	p k	k	k	k	k
5	ʃ	ʃ	ʃ	ʃ		

Note. Sounds given as symbols using the IPA (2018) and ExtIPA charts (Ball et al., 2018). θ/s=either sound was the target. Airflow=oral airflow.

Therapists Sarah, Laura and Helen (in Helen-Naomi) exhibited similarities in their explicit identification of phoneme targets at the outset of therapy. Sarah and Laura selected targets representing phonemes affected by specific active cleft speech characteristics in the children's speech: backing to velar (Louise); and turbulent active nasal fricatives (Ellie). Therapist Helen identified an absence of pressure consonants in Naomi's speech (suggesting underlying velopharyngeal insufficiency) and described therapy as 'diagnostic'. In Sarah-Louise's episode, speech sound stimuli for the voiceless phonemes /t s f/ were introduced in the first and second videos and they appeared quite systematically throughout the episode. For example, video one focused on the voiceless alveolar sounds /t s/ and video five focused on the fricative sounds /f s/ following an oral airflow activity. The speech sound stimulus for voiced phoneme /d/, affected by the same backing to velar pattern as /t/, was observed for the first time in the last video (Sarah+Louise-S8-V7); it was introduced in a hide and seek activity alongside the stimulus for /t/. The majority (94%) of activities in Sarah-Louise's episode focused on a single sound and it was rare for Sarah to introduce multiple sounds within the same activity. A different approach was seen in Laura-Ellie's episode: speech sound stimuli for all four phoneme targets /f θ s j/ were introduced in video one and /θ s j/ featured in every video throughout the episode. The stimulus for phoneme /f/ did not feature beyond video two. Laura appeared to favour a strategy of one or two sounds in every activity and mixed interchangeably between these. In Helen-Naomi's episode, Helen appeared to favour a strategy of opening MSIVT sessions with one or two activities focusing on multiple sounds, sometimes as many as five, and then used the rest of the session to focus on one sound. The Helen-Thomas and Helen-Hannah episodes featured the two youngest children, who were 1;6 when they started therapy. In these episodes, also described as 'diagnostic', Helen did not identify explicit phoneme targets at the outset of therapy. The first half of these episodes featured a variety of 'potential' targets, identified by the use of adaptive articulations and cued articulation gestures. The second half of the episodes then focused on specific phonemes and stimuli.

There was variation within and between episodes in the number of activities that SLTs allocated to different speech sound targets and stimuli. In Sarah-Louise, SLT Sarah focused primarily on stimuli for the phonemes /t s/. Over the course of the episode, over two thirds of the activities (22/32; 69%) featured stimuli for one or both of these phonemes. A similar pattern was observed in Laura-Ellie's episode, where SLT Laura also appeared to prioritise two of the phonemes, /θ/s/ and /j/. All 23 (100%) activities featured stimuli for one or both of

these phonemes. Therapist Helen in Helen-Naomi, who was similar to therapists and Laura in other aspects of target and stimuli selection, allocated activities more evenly across four of the five phonemes: /p t ʃ k/.

6.3.3.1 Linguistic level

The analysis revealed that whilst all SLTs produced stimuli at the two most popular levels, single sound and real word, they allocated different amounts of time to the different levels and did not follow the same sequence through them. Therapists Sarah and Laura, whose episodes targeted specific characteristics diagnosed at the outset of therapy, demonstrated a preference for producing speech sound stimuli at single sound level and tended to introduce sounds at this level first. They did produce stimuli in real words but in fewer activities and generally more in the second halves of the episodes. In contrast, therapist Helen in the Helen-Thomas and Helen-Hannah episodes of care focused less on single sounds and showed a preference for producing stimuli in real words. Thomas and Hannah presented with absent pressure consonants at the outset of therapy and Helen described their therapy as 'diagnostic'. Helen did produce single sound stimuli in these episodes, though this occurred more in the second halves, i.e. the reverse of the pattern seen in the Sarah-Louise and Laura-Ellie episodes. There was marked variability in therapists' use of sound sequence stimuli. This level featured primarily in Sarah-Louise, Laura-Ellie and Helen-Naomi episodes, which involved the three oldest children. Helen used this level quite consistently throughout the Helen-Naomi episode, yet she did not use it at all in the Helen-Thomas episode and used it in only one video in the Helen-Hannah episode.

Close inspection of the metadata revealed that Sarah, Laura and Helen (in Helen-Naomi) produced speech sound stimuli for all phoneme targets at single sound level first. The progression from single sound level showed variability. Sarah progressed /f t s/ to real word level but she did this alongside sound sequence level and in the last three videos she moved systematically up the levels from single sound to real word. Laura followed a similar approach for the phonemes /θ/s/ and /ʃ/, moving systematically from single sound to sound sequence and then to real words, and Helen did something similar in the Helen-Naomi episode, moving from single sound with/without sound sequence and on to real word level. In the Helen-Thomas and Helen-Hannah episodes, Helen used a different approach. In Helen-Thomas, speech sound stimuli for all phonemes were first produced at real word level and only three of these (stimuli for /p t ʃ/) were later produced at single sound level. The same occurred in the

Helen-Hannah episode, although some of the stimuli (for /p t d/) were produced at single sound level before real word level. The sequence of movement through the levels was less systematic in the Helen-Thomas and Helen-Hannah episodes.

6.4 Child stimulus productions

Table 6.6 shows the number and proportion of MSIVT sessions and activities in which the child made one or more stimulus productions.

Table 6.6: Number and proportion of activities containing child stimulus production

	Louise	Ellie	Thomas	Hannah	Naomi	Total
Number of sessions	6/7	6/6	1/5	2/6	5/5	20/29
Proportion (%)	86	100	20	33	100	69
Number of activities	20/32	15/23	1/17	9/22	17/25	62/119
Proportion (%)	63	65	6	41	68	52

All five children made stimulus productions. They occurred in more than two thirds (20/29; 69%) of MSIVT sessions and approximately half (62/119; 52%) of MSIVT activities. However, the proportion of sessions and activities in which they occurred varied markedly from episode to episode. Louise, Ellie and Naomi demonstrated the highest rates of production. They made productions in 86-100% of their sessions and in more than 60% of their activities. Thomas and Hannah demonstrated the lowest rates of production, in 20% and 33% of their sessions, respectively, and 6% and 41% of their activities. Figure 6.1 shows when productions occurred in the five episodes.

Figure 6.1: Number of activities containing child stimulus productions

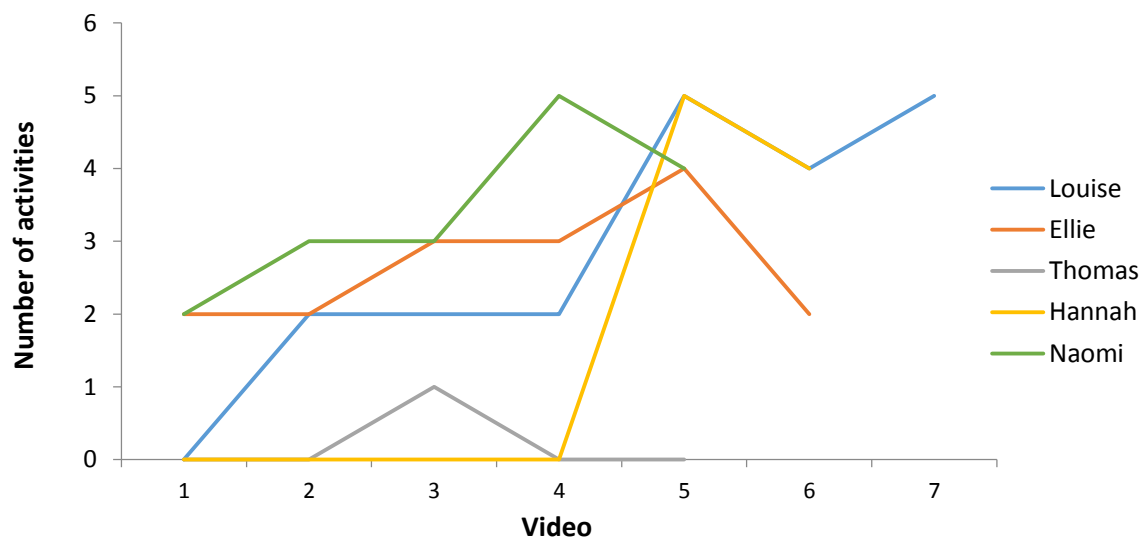


Figure 6.1 illustrates variation in the pattern in which productions emerged and continued throughout the episodes. Louise, Ellie and Naomi made stimulus productions in as early as the second and first video sessions, respectively, and continued to make productions throughout the episodes. In contrast, Thomas only made productions in one video (video three) half-way through the episode and Hannah only made productions in the final two videos (videos five and six). Thomas and Hannah were the two youngest children at the outset and throughout therapy.

Analysis of the metadata revealed variation between children in how many of their phoneme targets they made stimulus productions for, and at what levels. With some parallels to the analysis of how SLTs delivered therapy in the five episodes of care, there were similarities between Louise, Ellie and Naomi, and separately between Thomas and Hannah. Louise made productions for all four of her target sounds /f t d s/ at the three most popular levels: single sound; sound sequence; and real word. She also produced stimulus [f:] (for phoneme /f/) in invented real words. Ellie similarly made productions for all four of her target sounds /f θ s ʃ/ at single sound level and productions for the sounds /θ/s ʃ/ at sound sequence and real word levels. Naomi made productions for all five of her target sounds /p f t ʃ k/ at single sound level and all but one, /k/, at real word levels. She also produced the stimulus for target /ʃ/ at sound sequence level. In contrast, Thomas and Hannah made productions for one of their target sounds: /p/ at real word level (Thomas); and /t/ at single sound and real word levels (Hannah).

6.5 Conclusions

This chapter has presented the analysis of five MSIVT episodes of care as they were implemented in the NHS by three SLTs in two specialist SLT services. The children ranged in age from 1;6 to 2;3 at the start of therapy and 1;11 to 2;11 when it concluded. An episode of care consisted of between six and nine therapy sessions over a period of five to 11 months. Sessions took place, on average, once a month. The majority (76%; 32/42) of sessions involved video-therapy. Twenty-nine (90%) of these MSIVT sessions were available for analysis. The average duration of an MSIVT session was 20 minutes, though this ranged from nine to 37 minutes. Session duration may relate to the age of the child because the youngest children received the shortest sessions. One parent was present in every session. An MSIVT session consisted of, on average, five therapy activities. The average duration of an activity was four and a half minutes, though this ranged from under a minute to 12 minutes. Activity duration may too relate to the age of the child because the shortest activities featured in the shortest sessions. Activities were therapist-led, play-based and toy-based. Pictures were used in a third of activities.

The analysis found episodes of care targeted multiple speech sounds represented by multiple speech sound stimuli. Selected targets usually related to each other by one or more feature, e.g. place or manner of articulation. Multiple target sounds were targeted within a session, though individual activities typically focused on one target sound at a time. Across the episodes, all SLTs produced speech sound stimuli with adaptive articulations, a feature for making certain aspects of sounds more salient, e.g. linguolabial placement to highlight place of articulation (Harding-Bell & Howard, 2011). They also made cued articulation finger gestures alongside stimulus production. In some cases, more than one stimulus was used for the same target phoneme. Therapists produced speech sound stimuli at different levels, though most commonly as single sounds and in real words. Sound sequence stimuli featured infrequently, although in the episodes featuring the three oldest children (Sarah-Louise, Laura-Ellie and Helen-Naomi) it was used within a systematic approach. Invented word stimuli featured in just one activity. There was variation in the delivery of MSIVT in these episodes of care in relation to the sequence in which different sounds were targeted and the levels at which stimuli were produced. This may reflect the different speech characteristics in children's speech and also the targeted versus diagnostic nature of therapy. However, the lack of a consistent target and stimuli selection strategy across the episodes may also reflect the lack of prior investigation

into the MSIVT approach. It was interesting to find that not all speech sound stimuli were produced as single sounds prior to real words as this has implications for how the child may process the stimuli. The analysis found that all five children in this study made stimulus productions during therapy, though there was variation in how soon into therapy productions emerged and the levels in which they were made.

Chapter 7: Interactional Features of MSIVT (Phase 2)

7.1 Introduction

This chapter presents the results of Phase 2. This phase of the study aimed to analyse the features of therapist-child interaction in multisensory input video-therapy (MSIVT) sessions in order to identify and describe the ways speech and language therapists (SLTs) establish a child's attention and stimulate their awareness and production of speech sounds during therapy.

The analysis of 120 minutes of MSIVT interaction identified a total of 634 stimulus routine sequences (162 featuring Sarah-Louise; 179 featuring Laura-Ellie; 100 featuring Helen-Naomi; and 193 featuring Sam-James). Analysis of these sequences focused on the action accomplished in therapists' first-pair part turns. It revealed a distinction between actions that elicited a response from the child (invitations) and a type of action that did not (demonstration to the camera). In an invitation, the therapist directs her attention to the child and invites the child to do something relating to the speech sound stimulus. This definition of invitation is similar to Drew's (1981) more general use of the term in relation to initiating turns that invite self-correction, and is different to Schegloff's (2007) narrower usage of invitation as a social action where, for example, one person invites another person out for a coffee. In a demonstration to the camera, the therapist directs her attention to the camera and does not invite a here and now response from the child sat beside him/her. Key findings from the analysis are the ways therapists delivered invitations, how they orientated to the speech sound stimuli and targets at play, and the range of nonverbal as well as verbal responses they framed for the child. Four macro categories of action emerged from the analysis:

- A. Demonstration to the camera**
- B. Invitation to attend to the stimulus**
- C. Invitation to participate in the stimulus routine**
- D. Invitation to produce the stimulus**

Invitations took a variety of forms, ranging from demonstrations with one or more features to mobilise a response, to explicit verbal requests for a response. Explicit invitations resemble the verbal requests that are characteristic of the speech and language therapy discourse in Prutting et al. (1978), and are similar to directives, a characteristic feature of teacher-pupil

interaction (Sinclair & Coulthard, 1975). These four macro categories are distinct by their turn design features and the types of child responses they engendered. Within them, the analysis revealed nine sub-categories of action, also distinct by their design and the responses they engendered. They represent a hierarchy from least to most demanding action. In this chapter, I use extracts from the data to illustrate and describe each of the nine sub-categories of action and the consequences they had for the child and continuing trajectory of talk. I illustrate each sub-category of action with at least two extracts (except *D3. Nose-closing to elicit production*, which has just one extract). Table 7.1 summarises the nine sub-categories, their design features, the responses they engendered, and what therapists did in third turn position to accept or reject these responses.

Table 7.1: Summary of therapist actions, turn design features and the nature of subsequent turns

Therapist action	Design features	Child response	Therapist evaluation
A. Demonstration to the camera			
A1. Demonstration for later	Stimulus production with/without saliency features; camera-directed gaze	Engagement with the activity/absence of shared gaze (Extracts 1&2)	Absence of pursuit (Extracts 1&2)
B. Invitation to attend to the stimulus			
B1. Demonstration for attention	Stimulus production with saliency features; child-directed gaze during stimulus production	Absence of mutual gaze (Extract 3) Mutual gaze (Extract 4)	Pursuit for attention (Extract 3) Praise for attention (Extract 3) Reinforcement of stimulus (Extract 4)
B2. Verbal request for attention	Verbal directive or request (question) containing a verb associated with attention; singular or plural pronoun	Absence of mutual gaze (Extracts 5&6) Mutual gaze (Extracts 5&6) Stimulus production (Extract 3)	Pursuit for attention (Extracts 5&6)
C. Invitation to participate in the stimulus routine			
C1. Gesture to elicit participation	Nonverbal directive involving a toy, picture or finger gesture associated with the stimulus or routine	Participation in the stimulus routine (Extracts 7/8) Absence of participation (Extract 8)	Praise for participation (Extract 7) Absence of pursuit (Extract 8) Pursuit for participation (Extract 8)
C2. Verbal request for participation	Verbal directive or request (question) containing a verb associated with participation; singular or plural pronoun	Participation in the stimulus routine (Extract 9) Absence of participation (Extract 10) Stimulus production (Extract 10)	Pursuit for participation (Extract 10)
D. Invitation to produce the stimulus			
D1. Demonstration for imitation	Stimulus production with saliency features; sustained child-directed gaze and pause following stimulus production	Stimulus production (Extracts 11&12)	Praises and accepts production (Extracts 11&12) Rejects production and initiates repair (Extract 12)
D2. Incomplete demonstration	Incomplete verbal or nonverbal stimulus utterance; interrogative prosody; child-directed gaze and pause in the stimulus slot	Absence of production (Extract 13) Stimulus production (Extract 14)	Absence of pursuit for production (Extract 13) Praise for production (Extract 14)
D3. Nose-closing to elicit production	Nose-closing (gesture associated with stimulus production) on the child; child-directed gaze	Absence of production (Extract 15)	Absence of pursuit (Extract 15)
D4. Verbal request for production	Verbal directive or request (question) containing a verb associated with production; singular or plural pronoun; child-directed gaze; may or may not include a stimulus production	Absence of production (Extracts 16&17)	Pursuit for production (Extract 16) Absence of pursuit (Extract 17)

7.2 A. Demonstration to the camera

7.2.1 A1. Demonstration for later

A1. *Demonstration for later* belongs to the macro category A. *Demonstration to the camera*. In A1. *Demonstration for later*, the therapist produces the stimulus with camera-directed gaze with or without features to make auditory, visual or tactile properties of the sound stimulus salient. So called saliency features include adaptive articulation (Calladine & Vance, 2019; Harding & Bryan, 2000), a verbal label, e.g. ‘the bubble sound’ (Harding & Bryan, 2000; Vivanti, Hocking, Fanning & Dissanayake, 2016), an accompanying nonverbal cued articulation finger gesture (Passy, 1993), and/or a brief pause before the stimulus. The pause disrupts the fluency of the stimulus utterance and signifies that something of importance is coming (Gardner, 1994). These features help to identify the stimulus, as does the gaze to the camera, especially when it represents a change in gaze direction.

The first example of this action is illustrated in Extract 1. Therapist Sarah and child Louise are washing plastic toy animals in a toy bath. Louise’s mother and sister are also present. The activity features the recurring steps: 1. Therapist produces the stimulus (demonstration) → 2. Therapist and child wash the animal.

Extract 1: Washing toy animals (Sarah+Louise-V4-A2)

001 T Ok, so, where’s that [elephant
002 M [Here he is ((picks up the elephant))
003 C [((lifts gaze to the elephant))
004 T [Do you want to put him [in the wa- ((turns gaze [to camera)) wɔ̃tə
005 C [((takes hold of the elephant))
006 M [Put him in
007 ((pointing to the bath))
008 C ((drops elephant in the bath; gaze to the bath))
009 M Ooh
010 T Here’s your little brush ((hands brush to C; gaze to C))
011 C ((takes hold of brush; [reaches it into the bath; gaze to bath))
012 T [And we can wash his↑((turns gaze to camera))
013 (.) t̥əu ((washing with brush; lowers gaze))
014 C ((washing with brush; gaze to bath/elephant))
015 T Shall we wash his ((turns gaze to camera)) (.) t̥əu
016 Ah, there we go ((gaze to C)) you’re washing his (.) t̥əu
017 Ah, C, shall we wash ((turns gaze to camera)) his (.) t̥ɛt̥
018 ((gaze to C)) can you wash his ((turns gaze to camera)) (.) t̥ɛt̥

019 There we go ((washing with brush))
 020 C ((follows T's direction with brush))
 021 T There's his ((turns gaze to camera)) (.) **tɪrɪ**
 022 (3.0)
 023 T and what about his ((turns gaze to camera))(.) **tʊmi** ((gaze to bath))
 024 can you wash his (.) **tʊmi** in here ((washing with brush))
 025 C ((follows T's direction with brush))
 026 T There's his ((turns gaze to camera)) (.) **tʊmi**
 027 Ha ((gaze to C)) he's **wet** isn't he ((turns gaze to camera)) he's **wet**
 028 Shall we take him out and choose somebody else
 029 C ((gentle nod))
 030 T ((reaches into the bath and takes out the elephant))
 031 C ((reaches to the elephant))
 032 T Okay, put him on here to dry ((puts elephant down on the paper
 033 towels/on the table))

The sound stimulus is a voiceless interdental plosive [$t̪$] to represent the target phoneme /t/ in a variety of positions in five real words *water* (line 004), *toe* (lines 013, 015, 016), *tail* (lines 017, 018, 021), *tummy* (lines 023, 024, 026) and *wet* (line 027). Therapist Sarah produces the sound stimulus with adaptive articulation in the form of interdental placement and introduces the stimulus word with rising prosody (line 012) and a brief pause (e.g. lines 013, 016). She makes 12 stimulus productions in this extract, and almost all of them (10/12; 83%) are marked by a brief pause before the word. The first production is in line 004 (highlighted in grey). Sarah cuts short her initial start at the real word *water*, shifts her gaze to the camera, and then restarts it. This is a clear example of camera-directed gaze. Louise simultaneously reaches for the elephant and drops it into the bath. Her gaze tracks the elephant into the bath and then moves to the brush as she takes it from the therapist. At no point in this sequence (lines 001 to 011) does Louise's gaze turn to Sarah, and Sarah neither invites nor pursues it. Louise's compliance with her mother's instruction, in lines 006 and 007, to put the elephant in the bath indicates she is engaged with the activity, and the absence of her displaying visual attention to the stimulus is not treated as problematic. This is an example of *A1. Demonstration to the camera*. It is the first of a series in this extract, with further examples in the lines highlighted in grey. Sarah's stimulus productions are embedded within meaningful utterances that relate to the action and objects of shared attention, washing different parts of the elephant. Some are presented as directives, e.g. "*can you wash his...tail*" (line 018), and others as announcements, e.g. "*there's his...tummy*" (line 026) and "*...he's wet*" (line 027). Throughout the extract, we see

evidence of Louise’s engagement as she washes the elephant and follows Sarah’s lead to different parts of its body.

In Extract 2, therapist Laura and child Ellie are building a tower with coloured plastic bricks. Ellie’s dad is also present. The activity consists of the recurring steps: 1. Child chooses a coloured brick → 2. Therapist produces the stimulus → 3. Child builds the tower.

Extract 2: Building a brick tower (Laura+Ellie-V4-A1)

001 T There we go ((holds brick near mouth)) and the next one goes
 002 [((turns gaze to camera)) **θ̥i:** [((moves gaze to C; leans closer))
 003 C [((turns gaze to T)) [((turns away))
 004 T **θ̥i:** ((gaze to camera)) [((puts brick on the table)) can that one go
 005 on top ((pointing to the tower))
 006 C [((gaze to brick))
 007 ((gentle nod; picks brick up and puts it on top of the tower))
 008 T And then we’ve got some:: [pink ones to come
 009 C [((turns gaze to T/new bricks))
 010 T [This one goes ((holds brick near mouth; gaze to camera)) **θ̥ai:**
 011 ((turns gaze to C during production))
 012 C [((lifts gaze to T; [reaches for brick and puts it on the tower))
 013 T [((turns gaze to camera)) **θ̥ai:**
 014 C ((gaze to the tower)) [it goes up ((taps
 015 the tower))
 016 T It [is going up, up, up, up
 017 C [((turns gaze to T, then back to the tower))
 018 T And [I’ve got a green one that goes ((brick by mouth; gaze to camera))
 019 (.) **θ̥u:** ((moves gaze to C during production; brick with-held))
 020 C [((gaze to T/brick; reaches for the brick))
 021 ((takes brick when it is released))
 022 T Good watching
 023 C ((smiles))
 024 T ((turns gaze to camera)) **θ̥u:** ((finger to corner of mouth))
 025 C ((puts the brick on the tower))

Here, the stimulus is a voiceless interdental fricative in three consonant-vowel (CV) sound sequences, [θ̥i θ̥ai θ̥u] (lines 002, 004, 010, 013, 019, 024). Therapist Laura produces the fricative with adaptive articulation in the form of lengthened articulation and visible, but not exaggerated, anterior tongue placement. She also displays objects and gestures alongside her stimulus productions, e.g. holds a brick by her mouth (lines 001, 010, 018), and holds a finger by her mouth (line 024). The first stimulus routine sequence commences in line 001. Laura’s

verbal utterance “*and the next one goes...*” introduces her stimulus production in line 002. Although Laura’s gaze is to the camera for the utterance, she moves her gaze to Ellie and moves closer to her as she produces the stimulus, so this is not a demonstration to the camera. The stimulus production in line 004 (highlighted in grey) is made with camera-directed gaze, so this is the first example of *A1. Demonstration for later*. This may have been triggered by Ellie’s withdraw of mutual gaze in line 003. Ellie’s attention is re-attained by the brick and her response (lines 006-007) is evidence that she is engaged with the tower building activity. Further examples of *A1. Demonstration for later* occur in lines 013 and 024 (also highlighted in grey). Interestingly, all three examples within this extract occur following demonstrations with child-directed gaze (lines 002, 010, 018), suggesting a recurring sequence: stimulus production for the child *now*, followed by stimulus production for the child *later*. In all cases of *A1. Demonstration for later*, Ellie’s gaze is not to Laura. For example, in line 014, Ellie’s gaze is to the brick tower and her utterance “*it goes up*” overlaps Laura’s stimulus production in line 013, suggesting her attention is not on the stimulus. Laura does not pursue Ellie’s attention; instead, she positively reinforces her observation that the tower she is building “*...is going up, up, up*” (line 016). This is in contrast to the following sequence commencing in line 018, which introduces the second sub-category, *B1. Demonstration for attention*. We revisit this in the following section.

Extracts 1 and 2 have presented eight examples of the therapist action, *A1. Demonstration for later*. In none of these examples was the child’s, Louise in Extract 1 and Ellie in Extract 2, gaze to the therapist; their gaze was usually to the toys being used in the activity. Furthermore, on no occasion did either therapist, Sarah in Extract 1 or Laura in Extract 2, display any pursuit for the child’s attention. As such, these examples portray these demonstrations as single turn actions that do not engender a here and now response. This action has similarities with the restart-relevant directives in the Reed et al. (2013) study.

7.3 B. Invitation to attend to the stimulus

7.3.1 B1. Demonstration for attention

B1. Demonstration for attention belongs to the macro category *B. Invitation to attend to the stimulus*. In the previous extract, Extract 2, lines 018-019, we saw therapist Laura introduce a new brick and new sound sequence stimulus [ʊ̥u:]. Although her gaze was to the camera while she produced the utterance “*and I’ve got a green one that goes...*”, and for the onset of

stimulus production, she moved her gaze to Ellie whilst producing the stimulus and withheld the brick. Ellie reached her hand to the brick but did not take it and maintained eye gaze with the therapist/brick until Laura completed her production and released the brick. Laura's response in line 022 is an evaluation, in which she praises Ellie for her visual attention ("*good watching*"). In redirecting her gaze to Ellie, and by withholding the brick until Ellie meets her gaze, Laura has made Ellie more accountable to attend to what she is doing in the here and now. Her demonstration has become an invitation. Laura establishes Ellie's visual attention with her utterance "*and I've got a green one that goes...*" and introduction of the brick in line 018, and she sustains it by withholding the brick while she delivers the important stimulus information. By holding the brick (object of interest) close to her mouth, Laura helps to ensure Ellie sees her tongue placement. In this combination, these multimodal behaviours serve (successfully) as attention-mobilising features. In the absence of any verbal directive to watch or listen to the stimulus, this is an example of an implicit invitation to attend to the stimulus. As soon as Laura completes her stimulus production, and praises Ellie for watching, she breaks her gaze with Ellie and moves it to the camera. Along with the specific nature of her praise ("*good watching*") and the absence of a pause between stimulus production and releasing the brick, (which allows Ellie to shift her attention to the tower), this strongly suggests Laura is not seeking a stimulus production from Ellie. This is an important observation because it distinguishes between stimulus productions set up for attention (*B1. Demonstration for attention*) and those set up as models for imitation (*D1. Demonstration for imitation*).

B1. Demonstration for attention is therefore an implicit form of *B. Invitation to attend to the stimulus*. It is characterised by a stimulus production with features that mobilise the child's attention, namely child-directed eye gaze during (but not sustained beyond) production, as well as features such as body leaning and saliency features that draw attention to the stimulus. If attention is not established, we may see a pursuit from the therapist. We do not see a pause following stimulus production; instead, the therapist will facilitate progression with the activity, unless they are dealing with a pursuit for the child's attention. We may observe the therapist explicitly praising the child for demonstrating good listening or watching. Characteristically, if the child does not produce (imitate) the stimulus, it is not treated as problematic.

Extract 3 features the same therapist-child dyad, Laura-Ellie. In this activity, they are threading coloured cotton reels onto a piece of string. Ellie's dad is present. The activity consists of the recurring steps: 1. Child chooses a coloured reel → 2. Therapist produces the stimulus → 3. Child threads the reel.

Extract 3: Threading cotton reels (Laura+Ellie-V1-A1)

001 T Another one [((gaze to C)) shall we have more?
002 C [((gaze to T)) yeah
003 T What colour, shall we have green or yellow ((presents reels))
004 C Yellow ((gaze to T))
005 T Yellow ((gaze to C))
006 **fɪə** [**fɪə** ((holds reel by mouth; gaze to C))
007 C [() ((lowers gaze; plays with string))
008 T Ready↑
009 C () ((lifts gaze to T))
010 T We're gonna listen first ((points to ear)) **fɪ** [**ɪ** ((reel by mouth;
011 gaze to C))
012 C ((lifts gaze to T/reel)) [No, no, I ((takes
013 reel and puts it on the table))
014 T Good girl, well done, on it goes
015 C ((plays with the string))
016 T ((moves gaze to camera)) on it goes **fɪ** [**fɪ** ((turns gaze to C)) **fɪ**
017 C [I wanna try ((tries to
018 thread the reel; drops the string))
019 T Oh dear
020 C ((bends down and [picks up the string)) I want to
021 T [It fell **af** ((gaze to camera))
022 C ()
023 T It **fel** ((gaze to camera)) **af** ((gaze to the reel/string in C's hand))
(1.09)
024 T This one goes **fɪ** ((holds reel by mouth; gaze to camera))
025 C ((brief gaze to T; plays with string and reel))
026 T ((turns gaze to C)) **fɪ** [**ɪ** ((reel by mouth))
027 C ((gaze to T)) [ha () ((gaze to string/reel))
028 T [C, listen first ((points to ear; gaze to C)) **fɪ** [**ɪ** ((reel by mouth))
029 C [((turns gaze to T)) [ϕ ((reel by her
030 mouth))
031 T Good girl ((points to C)) [you did ((points)) it too, well done
032 C [((turns round to her dad))
033 [() **Fɪ**:
034 T [On it go- ((offers the reel; gaze to C))
035 Ah, you showed daddy, that was a clever sound ((gaze to C))

The stimulus in Extract 3 is a voiceless labiodental fricative [f] produced with soft and lengthened articulation. It features as a single sound and in the real words *off* and *fell* (lines 021, 023). Throughout the extract, we see therapist Laura hold cotton reels beside her mouth while she produces the stimulus. This resembles her enactment with the bricks in Extract 2.

Laura's first stimulus production occurs in line 006 (highlighted in grey). Her gaze is to Ellie but there is no explicit direction for Ellie's attention; as such, it is an example of *B1. Demonstration for attention*, an implicit invitation to attend to the stimulus. Although Ellie's gaze is to the therapist for the first production, she does not display sustained attention; she starts to utter something as she breaks her gaze and plays with the piece of string in her hand. Laura's verbal utterances in lines 008 ("*ready*" produced with interrogative prosody) and 010 ("*we're gonna listen first...*" with accompanying gesture) represent a pursuit for Ellie's attention, which we did not see following the therapists' demonstrations to the camera in Extracts 1 and 2. Although Ellie does gaze to Laura/the reel (line 012), this display of visual attention is brief and she once again overlaps Laura's stimulus production with her next utterance. Despite this, Laura praises Ellie (line 014) and allows the activity to continue. In this extract, it is noteworthy that Laura does not pursue Ellie for a stimulus production; she pursues her attention. Her stimulus productions in line 006, 010 and 016 (made salient with features such as adaptive articulation) appear to be set up for Ellie to listen or attend to, and not as models for Ellie to imitate.

We observe a similar enactment later on in the same activity. Laura's stimulus production with child-directed gaze (line 026) is another implicit invitation to attend to the stimulus. Ellie again overlaps Laura's production (027), breaks their mutual gaze, and starts to manipulate the string and reel in her hand. As in the first example, Laura treats this as problematic. One scenario is that the problematic nature of Ellie's response is the absence of a production, which would be the case if Laura's stimulus production were a model for imitation and not a *B1. Demonstration for attention*. However, Laura's pursuit in line 028 suggests it is the loss of visual attention characterised by Ellie's lowered gaze that was problematic. Laura pursues Ellie's (auditory) attention ("*C, listen first...*" line 028) with a verbal directive and delivers a new production characteristic of *B2. Verbal request for attention*. This pursuit provides internal evidence that Laura's action was an invitation requiring Ellie's attention in the here and now. As well as re-establishing mutual gaze, Ellie responds with a stimulus production (line 028). Laura's surprised response (line 029) suggests this is over and above what she was seeking from Ellie, and is therefore further evidence to support categorisation of her action in line 026 as an invitation for attention.

Extract 4 features the therapist-child dyad Sarah-Louise. They are sticking stickers (to represent drips of water) onto a picture of a tap. The activity has the recurring steps: 1. Child chooses a sticker/'drip' → 2. Therapist produces the stimulus → 3. Child sticks the 'drip' on the picture.

deliver a fifth and final production. Louise's sustained gaze suggests Sarah has her attention, and this creates opportunity (in line 014) for Sarah to provide reinforcement of the stimulus articulation with the articulatory descriptor "*with my tongue*". Sarah then nods and utters "...*that's it*" while simultaneously lowering her gaze from Louise and to the picture (line 016). This serves to close the sequence and inform Louise it is time to stick the sticker on her picture. There is no evidence to suggest Sarah's stimulus production was a model requiring imitation.

In each one of Louise's turns, we see evidence of her engagement with the activity. She indicates which sticker she wants (line 002), nods to affirm Sarah's receipt (line 004), takes the sticker when it is offered (line 006), holds it on her thumb and displays patience, establishes mutual gaze with Sarah as she produces the stimulus (line 008), sticks the 'drip' only when Sarah's turn is complete (line 017), and cooperates in the closing sequence of the activity (lines 020-022). Sarah does not pursue Louise for a stimulus production, and her utterance "*that's it*" in line 016 brings the stimulus routine to a close. Sarah's evaluative turn (lines 016, 018) suggests Louise's visual attention to the stimulus (lines 010-011) was an acceptable response to her invitation.

7.3.2 B2. Verbal request for attention

In Extracts 3 and 4, we saw examples of the therapist action *B1. Demonstration for attention*, an implicit action belonging to the macro category *B. Invitation to attend to the stimulus*. Its key defining feature was the therapists' displays of child-directed gaze during stimulus production. In Extract 3, featuring therapist-child dyad Laura-Ellie, Laura, on two occasions, made an explicit pursuit for Ellie's auditory attention when her *B1. Demonstration for attention* was unsuccessful. These came in the form of verbal directives: "*we're gonna listen first*" (line 010); and "*(child's name), listen first*" (line 028), both with an accompanying nonverbal gesture (*points to ear*). Laura only embarked on another stimulus production when Ellie re-established eye gaze with her, suggesting she also sought her visual attention. These are examples of the present action, *B2. Verbal request for attention*. This is an explicit action within the same macro category *B. Invitation to attend to the stimulus*. The explicit verbal request typically precedes a stimulus production. As such, the therapist request and child response (eye gaze to signify attention) form a pre-sequence prior to the main stimulus routine sequence. This resembles the attention-seeking sequence in Ronkainen's (2011) study of interaction during auditory-verbal therapy.

Extract 5, featuring therapist-child dyad Sam-James, illustrates a similar example. Sam and James are posting pictures of bubbles (representing the sound stimulus) into a toy post-box. The activity consists of the recurring steps: 1. Therapist produces the stimulus (demonstration) → 2. Child posts the picture.

Extract 5: Posting bubble pictures (Sam+James-V1-A1)

001 T Do some more↑ [((presents bubble pictures; gaze to C)) [some more↑
 002 C [((lifts gaze to T; nods; hands on post-box))
 004 [((lowers
 005 gaze; tries to open the post-box))
 006 T Are you ready C ((presents pictures; gaze to C)) got ((shakes the
 007 pictures)) some more look
 008 C ((gaze to post-box; tries to get into it))
 009 T ((gaze to post-box)) let's leave them in there ((taps post-box))
 010 ((lifts gaze to C)) [are you ready, the bubbles go:↑ **p p p** ((picture
 011 by face))
 012 C [((lifts gaze to T))
 013 T Shall we do it on your hand? ((reaches for C's hand))
 014 C ((allows T to take his hand))
 015 T Ah, ready **pppp** ((gaze to C; makes sounds close to C's hand;
 016 happy expression on face))
 017 (1.0)
 018 C ((takes his hand away))
 019 T And this one ((takes hold of C's other hand; gaze to C))
 020 C ((lifts gaze to T; allows T to take his hand))
 021 T **pppp** ((on C's hand; gaze to C))
 022 C ((takes hand and offers his other hand))
 023 T And that one! ((takes C's hand)) **pp** ((on C's hand; gaze to C))
 024 C ((takes hand; smiles; offers his other hand))
 025 T And that one! ((takes hand; smiles)) **pppp** ((on C's hand; gaze C))
 026 C ((takes hand away))
 027 T ((smiles; happy expression; gaze to C))
 028 C ((smiles))
 029 T There you go ((gives picture to C)) [good boy
 030 C [((takes picture; posts
 031 it))
 032 T ((turns gaze to camera)) [the bubble sound ((moves finger around
 033 lips)) **ppppp** ((moves gaze between camera and C))
 034 C [((posts the picture in the post-box))

The stimulus in Extract 5 is a voiceless bilabial plosive produced with weakened pressure (salient adaptive articulation) at single sound level. Therapist Sam makes multiple productions

in sound strings, uses bubble pictures to symbolise the stimulus, which she holds by her face during stimulus production, and names the stimulus with the verbal label “...the bubbles...” (line 010). Sam’s display of the bubble picture alongside stimulus production (lines 010-011) occurs recurrently throughout this activity and, as such, becomes part of the stimulus routine (Step 1 in the activity).

In line 006, Sam’s verbal directive “are you ready, (child’s name)?” which is accompanied by a nonverbal offering of the bubble pictures, is an explicit invitation for James to attend to the stimulus symbol, the bubble picture. Letts (1985) describes ‘ready’ as an ‘attention-getter’. Sam is not successful in establishing James’ shared attention; he instead continues to try to open the post-box. She tries again with a repeat of her utterance “are you ready?” in lines 010-011. James simultaneously shifts his gaze to Sam and she goes on to produce the stimulus. James maintains visual attention throughout the stimulus productions and all the way until Sam gives James the picture to post (line 029), thereby closing the stimulus routine sequence. Previously, in Extract 3, Ellie’s overlapping talk (lines 007, 027) suggested an absence of auditory attention and, in her pursuit, therapist Laura used a verb explicitly associated with this form of attention, “...listen...” This is not what we see with Sam and James in the present extract, Extract 5. Sam’s phrase “are you ready” seems to articulate attention on a broader level, and her pursuit suggests James’ visual attention is at least one form of attention she seeks.

In line 013, we see Sam introduce a new step in the stimulus routine (Step 1); she produces the sound stimulus on James’ hand. This is a technique associated with giving tactile stimulation of the oral plosion generated from aspirated stimulus production. James allows Sam to take his hand, and in lines 022 and 024 (highlighted in yellow) we see him initiate this himself by offering his hand without invitation. Mutual eye gaze is maintained for all 14 stimulus productions (lines 015, 021, 023, 025).

In Extract 6, therapist Helen and child Naomi are playing Puff the Pop-Up Dragon. Naomi’s mother is also taking part. There are flags on the table, each one sitting on top of a picture card that represents a sound stimulus. The activity has the recurring steps: 1. Child chooses a flag/picture → 2. Therapist (or mother) produces the stimulus → 3. Child (or mother) inserts the flag.

Extract 6: Puff the Pop-Up Dragon (Helen+Naomi-V3-A1)

001 T [So, [mummy's turn ((moves flags towards C)) which will mummy do
002 C [((gaze to flags))
003 [((reaches for a flag/picture))
004 T ((gaze to C))
005 C Um ((reaches for a flag/picture; moves it towards M))
006 [()
007 T [Yellow, thank you
008 C I can [do this one ((picks up green flag; reaches picture to T))
009 M [What have I got ((picks up the picture))
010 Oh: another, look ((holds baby sleeping picture; gaze to T))
011 T [So let's look ((points to mother)) at mummy's card first
012 Thank you ((whispered; takes picture from C))
013 C ((brief gaze to T))
014 M [((taps C on the arm))
015 T [Let's look at mummy's card ((points to M))
016 C ((moves gaze to M's picture))
017 M ((lifts gaze to camera; holds picture facing camera))
018 C ((moves gaze to camera))
019 M ((moves gaze to C)) baby sleeping [((moves picture closer to C))
020 C [((gaze picture; then lowers it))
021 M [ʃ: ((makes cued artic gesture))
022 C [Ha, mummy got the baby sleeping ((moves gaze and head to T))
023 ()
024 M ((puts picture on the table and moves it towards T))
025 T Mummy got the baby [sleeping ((picks up the picture; moves gaze))
026 to camera)) that goes ʃ: ((makes cued artic gesture))
027 C [((gaze to flag; reaches toward the castle))

The sound stimulus in this extract is a single sound voiceless post-alveolar fricative represented by the picture of a baby sleeping. Therapist Helen produces the stimulus as a lengthened fricative (salient adaptive articulation), makes a nonverbal cued articulation finger gesture alongside (line 026) and verbally labels it "...baby sleeping..." (line 022). The picture of a baby sleeping symbolically represents the sound stimulus. Helen's verbal directive to Naomi "...let's look at mummy's card first" and accompanying nonverbal gesture (points to mother) in line 011 is an explicit invitation for the child's (visual) attention to the stimulus. Naomi gazes briefly to Helen (line 013), but not to her mother. Helen's repeat utterance in line 015 ("let's look at mummy's card") is a pursuit for Naomi's visual attention. Naomi responds appropriately to Helen's instruction; we see her gaze move from Helen on her left to her mother's picture card on her right (line 016). She then gazes briefly to the camera before returning her gaze to the

picture in line 020. However, this attention is brief. By the time Naomi's mother embarks on stimulus production (line 021), Naomi has already lowered her gaze, suggesting loss of visual attention, and her overlapping utterance in lines 022-023 suggests a loss of auditory attention. Despite the earlier pursuit, Helen, at this stage, does not pursue Naomi's attention to her mother's stimulus production. Instead, she picks up the picture card and makes a stimulus production with camera-directed gaze (*A1. Demonstration for later*) while Naomi and her mother proceed with the activity.

Extracts 3, 4, 5 and 6 illustrate therapist turns that invite the child to display visual and/or auditory attention to the sound stimulus, sometimes directly (as in Extract 3), sometimes indirectly via its symbolic representation (as in Extract 6), and sometimes both (as in Extract 4). These actions belong to the macro category *B. Invitation to attend to the stimulus*. Sub-category *B1. Demonstration for attention* represents an implicit form of invitation, and sub-category *B2. Verbal request for attention* is an explicit invitation. We saw that both types of action might lead to a pursuit when the therapist has not established or maintained the child's attention. Sometimes, the pursuit is more explicit than the initial invitation, like in Extract 3.

I will now move on to illustrate two actions belonging to the macro category *C. Invitation to participate in the stimulus routine*. In these, we will see that there is a more demanding requirement for the child to perform one or more of the behaviours that make up the stimulus routine. An important prerequisite for this macro category of invitation is that the composition of behaviours that make up the stimulus routine has already been performed at least once, giving the child *a priori* knowledge. This has similarities with Letts' (1985) description of a 'signal'. As in *B. Invitation to attend to the stimulus*, the analysis revealed that *C. Invitation to participate in the stimulus routine* also comes in implicit and explicit forms.

7.4 C. Invitation to participate in the stimulus routine

7.4.1 C1. Gesture to elicit participation

Extracts 7 and 8 illustrate implicit invitations, which are characterised by a gestural rather than verbal request by the therapist to elicit the child's participation. Importantly, the specific gesture must have already been associated with some part of the enactment of the stimulus routine.

Extract 7 is of therapist Sarah and child Louise playing with the Click Clack Track game. There are pictures of snakes (symbolising the sound stimulus) on the surface of the table. The activity features the recurring steps: 1. Child chooses a car and puts it on the track → 2. (Car lands on a picture symbolising the stimulus and) therapist produces the stimulus.

Extract 7: Wooden Click Clack Track with snake pictures (Sarah+Louise-V4-A3)

001 T Which one's next ((gaze to C; moves cars closer to C))
 002 C ((points to green car))
 003 T Oh:: the green one ((hands car to C))
 004 C ((takes car and puts it on the top of the track))
 005 M Push it up there, C ((points to the top of the track))
 006 C ((moves car further up the track))
 007 T That's it [good girl, whoa::: wee:::
 008 C L((releases the car))
 009 ((gaze follows car; smiles))
 010 T Oh [((turns gaze to C)) which one's it got
 011 C L((gaze to T; reaches and points to the top of the track))
 012 T It went all the way down [it went really fast look ((points to
 013 the picture underneath the green car))
 014 C L((nods; turns gaze to T))
 015 T [It's caught this one ((moves picture/car towards C))
 016 C L((gaze to picture/car))
 017 T There we go ((picks up the car; reveals the picture underneath))
 018 C ((picks up the picture; holds it facing camera; gaze to camera))
 019 T Ah [((gaze to C)) and that snake say::s↑
 020 (1.0)
 021 C L((turns gaze to T))
 022 T s:: ((gaze to camera)) s:: ((makes cued artic gesture; gaze to C))
 023 C ((lowers gaze to picture))
 024 T Doesn't he ((takes picture)) good girl right ((presents the cars))
 025 you're showing that camera all your sounds
 026 C ((points to the blue car))
 027 T Oh, blue, there we go ((moves blue car towards C))
 028 C ((picks up the blue car; puts it on the top of the track))

The sound stimulus in Extract 7 is a voiceless dentalised fricative produced at single sound level. The dentalised tongue position and lengthened production are features of adaptive articulation. Therapist Sarah uses a snake picture to symbolise the stimulus and makes a nonverbal cued articulation gesture alongside stimulus production. The selected extract takes us two minutes into the activity. In earlier sequences, Sarah has verbally directed Louise to

hold or *show* the snake picture to the camera while Sarah produces the sound stimulus. This makes the picture display part of the stimulus routine (Step 2), similar to Sam's enactment with the bubble pictures in Extract 5. Sarah embodies the stimulus production in a composition of behaviours that centre on the stimulus picture, i.e. the stimulus routine. Sarah's nonverbal behaviour in line 015, in which she moves the selected picture/car towards Louise, is therefore an implicit invitation by Sarah for Louise's participation in the routine. Louise responds by gazing to, picking up, and then holding the picture so that it is facing the camera (line 018). This display is evidence of Louise's *a priori* knowledge of the stimulus routine, and validates Sarah's action as an invitation. In line 019, Sarah steps up her demands on Louise by implicitly inviting her to produce the stimulus (a type of action described later in this chapter). Louise establishes mutual gaze but does not produce the stimulus (line 021). Sarah goes on to produce the stimulus herself and praises Louise for "...*showing that camera all your sounds*" (line 025). Louise's responses in this extract suggest Sarah is praising her for her participation in the stimulus routine (holding pictures up for the camera), and perhaps also her attention to the stimulus, evidenced by appropriate eye gaze and an absence of overlapping talk.

Extract 8 provides another example of *C1. Gesture to elicit participation*. This extract features therapist-child dyad Sam-James. James' mother is also present. The same activity, at an earlier stage, featured in Extract 5. Therapist Sam has established a recurring stimulus routine featuring a stimulus production alongside nonverbal display of a bubble picture (and sometimes a wooden letter) by her face with child-directed gaze.

Extract 8: Posting bubble pictures (Sam+James-V1-A1)

001 T p p ((gaze to C; holds picture by face))
 002 C L((gaze to the post-box; posts letter))
 003 T ((picks up a bubble picture, offers it to C)) here we go
 004 C ((lowers gaze then returns gaze to the post-box))
 005 T Here you go ((offers picture))
 006 ((puts picture down on the table))
 007 C ((opens the post-box door))
 008 T ((picks up new letter)) I'll hold it for you ((holds door open))
 009 C ((reaches inside post-box; starts to remove pictures and letters))
 010 T LGood boy
 011 (4.0)
 012 T all these bubbles ((gaze follows C))
 013 M Wow, can you do it?
 014 T All these bubbles
 015 M Here you are ((offers C a bubble picture)) here you are

016 T Have you got them, okay ((turns gaze to C))
 017 (2.0)
 018 T ((holds picture and letter by face; gaze to C))
 019 C [((picks up a picture and a letter))
 020 M [That's it
 021 C ((holds picture by face; reaches letter to T's face; gaze to T))
 022 T **ppp** ((takes letter from C)) **p** ((gaze to C))
 023 Do you want that one? ((offers a letter to C))
 024 C ((takes letter but shakes head and hands it back))
 025 T No, for me, for me ((takes letter from C))
 026 C ((holds picture by face; gaze to T))
 027 T ((holds letters and picture by face; gaze to C)) **pppppp**
 028 C **pppp**
 029 T **(p) (p)** good boy ((nods; smiles))
 030 C ((smiles; turns gaze to post-box; starts to post the picture))

Extract 8 opens with therapist Sam producing the stimulus with child-directed gaze (line 001). The direction of James' gaze and manipulation of the letter and post-box (lines 002, 004) suggest his attention is not on the stimulus. Sam's offering of a bubble picture in lines 003 and 005 looks like a pursuit for James' attention and is a gestural invitation for him to take part in the stimulus routine. There is no evidence of James acknowledging the picture and his attention remains with the post-box. There is no evidence to suggest Sam treats this as problematic; we do not see a further pursuit for James' attention (to the stimulus) or his participation. Instead, she assists James, permitting a period of child-directed play ("I'll hold it for you ((holds door open))" line 008). James' mother's interjections in lines 013 and 015 constitute a pursuit for James' participation and he responds accordingly, picking up a bubble picture and holding it by his face (lines 019, 021). This display by James indicates his *a priori* knowledge of this behaviour within the stimulus routine as he is not, in this local sequence, instructed to hold the picture in this way. From here to the end of the extract, James' eye gaze and manipulation with the bubble picture (lines 019, 021, 026) are evidence that his attention is on the stimulus and, in line 028, we see him make a stimulus production. Sam's imitation in line 029, verbal praise "good boy" and nonverbal affirmation (*nods; smiles*), serve as an acceptance of James' accurate (matched) productions.

7.4.2 C2. Verbal request for participation

In C2. *Verbal request for participation*, the therapist's invitation is explicit. Her request is characterised by a verb associated with enacting one or more behaviours in the stimulus

routine. It may occur in conjunction with a gestural cue. The sequential context and orderliness of this type of action is similar to *B2. Verbal request for attention*. The therapist request and child response (action to signify participation) typically form a pre-sequence before the main stimulus routine sequence.

In Extract 9, therapist Sarah and child Louise are playing with the Wooden Click Clack Track game. Pictures of a tap, symbolising the sound stimulus, are arranged on the table. The activity consists of the recurring steps: 1. Child chooses a car and puts it on the track → 2. (Car lands on a tap picture and) therapist produces the stimulus.

Extract 9: Wooden Click Clack Track with drippy tap pictures (Sarah+Louise-V1-A2)

001 T Would you like (1.0) a blue [car or a green car ((offers cars;
 002 gaze to C))
 003 C [((turns gaze to cars))
 004 That one ((points to and takes the blue car))
 005 T The blue one
 006 C ((puts the car on the table and pushes it))
 007 T Right, pop it on top ((points to the top of the track))
 008 M Up there ((points to the top of the track))
 009 C [((puts car on the top of the track))
 010 T [See if we can catch another picture
 011 C ((releases the car))
 012 T Whoa::: wee:: [wow you've caught this one ((points to picture))
 013 C [((reaches for the car))
 014 T ((picks up the car and picture)) [shall we hold it up to the camera
 015 ((gaze to C; holds picture facing C))
 016 C [((gaze to the picture))
 017 ((nods; brief gaze to the camera))
 018 T Go on then ((offers picture to C)) [do you want to hold it up so
 019 that so that the camera can see?
 020 C [((gazes to and takes picture))
 021 ((holds picture facing camera; gaze to camera))
 022 T ((turns gaze to camera)) ah there it is and it ((points to picture))
 023 say::s↑ **ttttt** (...) **ttt** ((makes cued artic gesture in synchrony))
 024 ((turns gaze to C)) with my tongue
 025 C ((turns gaze to T))
 026 T **ttt** ((gaze to C))

In Extract 9, the sound stimulus is a voiceless interdental plosive made with visible tongue placement; it orientates to the phoneme /t/. Therapist Sarah typically produces the sound multiple times in a sound string. She makes the nonverbal cued articulation finger gesture for

/t/ alongside, and in synchrony, with her productions. In earlier sequences within this activity, the enactment of displaying tap pictures to the camera has been associated with stimulus production, i.e. it has become part of the stimulus routine. A directive by Sarah inviting the child to perform this behaviour is therefore an example of *C. Invitation to participate in the stimulus routine*. Extract 9 illustrates an explicit form of this action, and is an example of sub-category *C2. Verbal request for participation*. Therapist Sarah uses an action eliciting verb *hold* and refers directly to the camera; see lines 014 and 018-019. Although she does not name the picture or stimulus, her hold on the picture suggests this is what the pronoun *it* relates to. Louise's shift in gaze to the picture supports this; see lines 016 and 020. In Sarah's first invitation in lines 014-015, she uses the plural pronoun *we*. However, her body position, direction of her gaze (to Louise) and offering of the picture, suggest Louise is the target respondent. Louise nods in immediate response to Sarah's request (line 017) and turns her gaze to the camera, though she does not take hold of the picture. Whilst this brings into question the action of Sarah's request, her pursuit in lines 018-019 provides evidence that Louise was the target respondent. Note how in her pursuit, Sarah uses a singular pronoun referring directly to Louise: "*go on then, do you want to hold it up*" This does not mean that Louise's response in lines 016-017 was inappropriate. It is entirely reasonable to suggest, given her age and language levels, that she responded to the latter half of Sarah's request "...to the camera" only. Louise's rapid response to Sarah's more explicit request in lines 018-019 does support this. Louise's display of the picture and sustained camera-directed gaze while Sarah makes multiple stimulus productions, demonstrate her engagement with the routine.

In Extract 10, therapist Sam and child James are posting pictures into a toy post-box. The pictures are of big and little balls, which represent the sound stimulus, a voiced bilabial plosive [b]. The activity consists of the recurring steps: 1. Child chooses big or little ball → 2. Therapist produces the stimulus → 3. Child posts the picture. Earlier on in the activity, the enactment of showing pictures to the camera (referred to as 'telly') with accompanying camera-directed gaze has formed part of the stimulus routine.

Extract 10: Posting ball pictures (Sam+James-V14-A2)

001 T C, big one ((displays picture of big ball)) or little one
 002 ((displays picture of little ball; gaze to C))
 003 C ((holds post-box door open; gaze inside post-box))
 004 ((releases the door)) () ((gaze to post-box))
 005 T We're going to put them in there((gestures posting))aren't we
 006 C ((turns gaze to the picture))

007 T [Big one ((displays big ball picture)) or little one ((displays
 008 little ball picture))
 009 C [(gaze to pictures))
 010 () ((reaches for little ball picture))
 011 T [Little one ((keeps hold of picture))
 012 C [(gaze to T))
 013 T Are you going to show the [telly for me ((turns gaze; points to
 014 the camera))
 015 C [(gaze to camera))
 016 ((takes picture; reaches it towards the post-box; gaze follows))
 017 T ((gaze to C; reaches hand to occlude her nose))
 018 C, ((releases nose; reaches other hand towards the picture in C's
 019 hand)) show the telly first
 020 ((gaze to camera; [reaches hand to occlude nose))
 021 C [m[m m m ((gaze to camera; smiles))
 022 ((turns gaze to M))
 023 T [b b b ((holds nose; gaze C))
 024 M [Good boy, well done
 025 T [Good boy [(taps C on arm)) C
 026 ((turns to M)) so if I pinch his nose, that'll come out as a **bə**
 027 C [(posts the picture))
 028 T Good boy ((gaze to camera; occludes nose)) **bbb**
 029 ((turns gaze to C)) the bouncy ball [goes **bb** good boy, in there
 030 C [()

Therapist Sam's verbal utterance in line 013, involving singular pronoun *you*, is an explicit invitation for James' participation in the stimulus routine. However, the verb *show* is more ambiguous than the verb *hold* that we saw in Extract 9. Sam may have been inviting James to show the picture to the camera, i.e. seeking his participation, but she may instead have been inviting him to show the stimulus it represents to the camera, i.e. seeking production. Sam's subsequent behaviours in lines 017-020 suggest the former. In occluding her own nose, Sam is setting herself up to produce the stimulus. When James takes the picture to post it rather than display it, she treats this as problematic. In her pursuit (in lines 018-019), she reaches for the picture and asks again "*show the telly first*", which suggests the missing response that she seeks from James is a display of the picture and not a stimulus production. She simultaneously resumes nose closing, which further suggests stimulus production is going to be her role in the routine. In response, we see James turn his gaze to the camera and embark on four stimulus productions (lines 021). The absence of a prior stimulus production by Sam in this local sequence means James' output was processed from stored knowledge about the ball picture,

i.e. it is not an imitated response from non-lexical processing. Of course, James' [m] production is not a phonetic match to the stimulus [b], but Sam's positive praise to James (line 025), and explanation to his mother (line 026) tells us she accepts it as an accurate motor program. The hypothesis here is that the phonetic mismatch is passive in nature, due to velopharyngeal insufficiency and/or palatal fistula, and is not an active mismatch caused by an inaccurate motor program (Calladine & Vance, 2019). Sam draws this sequence to a close with two stimulus productions, one with camera-directed gaze (line 028) and another with child-directed gaze (line 029), using this opportunity to reinforce the verbal label and accurate phonetic template.

So far, this chapter has presented five sub-categories of therapist action explicated by detailed analysis of therapy talk within the MSIVT activities: *A1. Demonstration for later*; *B1. Demonstration for attention*; *B2. Verbal request for attention*; *C1. Gestural request for participation*; and *C2. Verbal request for participation*. *A1. Demonstration for later* does not place any demand on the child for a here and now response. *B1. Demonstration for attention* and *C1. Gesture to elicit participation* invite nonverbal responses from the child in the here and now and are therefore more demanding than *A1. Demonstration for later*, but not as demanding as the explicit forms, *B2. Verbal request for attention* and *C2. Verbal request for participation*. The final category of action revealed by the analysis is *D. Invitation to produce the stimulus*. Close examination of therapist-child interaction revealed four sub-categories of action.

7.5 D. Invitation to produce the stimulus

7.5.1 D1. Demonstration for imitation

Therapists' turns fulfilling this sub-category of action have design features in common with turns that function as *B1. Demonstration for attention*. For example, the therapist produces the stimulus with child-directed gaze and features associated with optimising saliency, e.g. adaptive articulation, holds toys and pictures by the face, and makes cued articulation finger gestures alongside production. However, additional features project stimulus productions as models for the child to imitate. The most distinctive appears to be sustained child-directed gaze, accompanied by a period of silence, beyond stimulus production. If the therapist displays a toy or picture alongside stimulus production, it is usually withheld.

In Extract 11, therapist Sam and child James are posting pictures of bubbles and toy letters into a toy post-box. We visited this activity previously in Extracts 5 and 8, as illustrations of the actions *B2. Verbal request for attention* and *C1. Gesture to elicit participation*, respectively. Extract 11 takes us further into the activity. The stimulus is a voiceless bilabial plosive produced with weakened articulation [p̥]. Therapist Sam typically produces it multiple times in a sound string rather than as one single sound. She refers to it with the verbal label ‘bubbles’ and symbolises it with a picture of bubbles. Throughout the activity, she uses the articulatory descriptor ‘with my lips’ and a finger gesture (motioning around her lips) to draw attention to where in the mouth the sound is made (place of articulation), though these latter features do not appear in this particular extract.

Extract 11: Posting bubble pictures (Sam+James-V1-A1)

001 C ((opens the post-box; gaze to post-box))
 002 T In there ((offers picture; gaze to C)) shall we put mine in
 003 C ((takes picture; reaches to post it))
 004 T ((gaze to C; postures articulatory placement for the stimulus))
 005 C ((retracts picture; brings it to side of his face; gaze to T))
 006 (p) [pp (p) ((maintains gaze during stimulus production))
 007 T [((nods; mutual gaze))
 008 C ((reaches to post picture in the post-box))
 009 T Good boy ((smiles; turns gaze to M)) good boy ((smiles; nods))
 010 C [((opens the post-box door; gaze to post-box))
 011 T [Open up [wow, what’s it in there?
 012 C [((reaches inside post-box; removes picture; gaze inside))
 013 T ((gaze to C))
 014 C [((reaches hand inside post-box; [removes another picture)
 015 T [Bubbles [all these bubbles
 016 C ((puts a picture on the table; [closes the post-box door))
 017 T [Close it up ((gaze to C))
 018 C ((gaze to the post-box; reaches with the picture; turns it around))
 019 [((reaches and picks up a letter))
 020 T [((reaches and picks up a picture; letter in other hand; gaze to C))
 021 C [((brings picture to hold it by his face))
 022 T [I’ve got some [((holds letter and picture by face; gaze C; postures
 023 articulatory placement for stimulus))
 024 C [((lifts gaze to T))
 025 (3.0)
 026 T ppp [p ((sustained gaze to C; picture/letter by face))
 027 C [((lifts a letter; holds it by face)) pp [p (p) (p) ((gaze to T))
 028 T [((smiles; nods))
 029 pp [pp ((smiles; turns gaze to M)) good boy ((whispered))

030 C L((reaches to post the picture and the letter))
 031 T In there ((gaze to letter as it is posted))
 (0.56)
 032 C ((gaze to the letter and pictures on the table))
 033 T Have some more↑ ((gaze follows C))
 034 ((holds picture and letter by face; gaze to C))
 035 C ((reaches to open the post-box))
 036 T Can you get some more↑
 037 C ((opens the post-box door))
 038 T Get them out ((removes pictures/letters; gaze to pictures/C))
 039 C ((holds door open; gaze to the pictures))
 040 T Good boy, are you helping?, is that it, yeah, that's it ((gaze C))
 041 I've got so::me ((holds picture and letter by face; gaze to C))
 042 (1.0)
 043 C ((lifts gaze to T; reaches and picks up a letter))
 044 T **pp** **pppp** ((picture/letter by face; maintains display))
 045 C L((lowers gaze to the pictures)) (noise) ((picks pictures up))
 046 T ((brief gaze to M; winks; maintains picture display))
 047 ((turns gaze back to C)) **p** ((picture/letter by face))
 048 [(4.0)
 049 C L((manipulates picture and letter in his hands; then lifts them up))
 050 [(holds picture/letter by his face; brief gaze T))
 051 T **pppp** ((gaze C; picture/letter by face))
 052 (0.1)
 053 C **(p) (p) (p) (p) (p) (p)** ((gaze to camera/M; picture/letter by face))
 054 T **ppppp** ((gaze to C; smiles; nods))
 055 C [(takes letter and picture from T; puts them on the table; reaches
 056 for pic))
 057 T L((sustained gaze to C; picture by face; postures placement))
 058 Ooh and that one ((releases picture; laughs))
 059 C ((puts picture on table; brings other picture to side of his face;
 060 lifts gaze to T))
 061 T ((meets C's gaze; holds letter by face)) **pp** [(p
 062 C L((lowers gaze; posts
 063 picture))

James makes the first stimulus productions in this extract; see line 006. Just prior to this, in line 005 (highlighted in yellow), he initiates the stimulus routine with his nonverbal display of the stimulus picture and letter and therapist-directed eye gaze. This sequence of behaviours shows Sam that James is engaged in the activity and willing to participate in stimulus production. The first example of *D1. Demonstration for imitation* by Sam occurs in line 026. Interestingly, it is James again who initiates this particular routine (line 021) when he lifts the picture and holds it

by his face. Sam mirrors this nonverbal display with her own letter and picture (line 022). Her utterance “I’ve got some↑ ((*postures articulatory placement for stimulus*))” accompanied by sustained gaze and pause is an example of *D2. Incomplete demonstration*, which I will discuss in the following section. James maintains his display posture and establishes mutual gaze with Sam, evidence that he is engaged and attending, but he does not produce the stimulus. Sam goes on to produce the stimulus herself in line 026. Unlike in turns operating as *B1. Demonstration for attention*, here Sam maintains the nonverbal display and child-directed gaze for several seconds beyond stimulus production, which indicate to James that the sequence is not yet complete. With these additional features, Sam’s turn fulfils a *D1. Demonstration for imitation* and mobilises a production response from James (line 027). James’ response provides internal evidence for the action and the sequence is brought to a close. This type of action resembles the ‘model’ and ‘augmented model’ in Gardner (1994) and McCartney’s (1989) studies.

A second series of turns fulfilling action *D1. Demonstration for imitation* occurs in lines 044-047 and 051-052. This time, it is Sam who initiates the routine; see line 041. James does not imitate the stimulus from Sam’s first productions (line 043); instead, he gazes to the pictures in his hands (lines 045, 049). In line 050, we see James enact the picture/letter display without explicit invitation from Sam and, in line 053, we see him produce a string of six stimulus productions. Sam’s sustained nonverbal display and child-directed gaze (from line 041) alongside repeated stimulus productions (lines 044, 047, 051) provide strong evidence that her productions are set up to elicit imitation. James’ productions match the stimulus on placement and manner, but are audibly silent. Sam’s nodding response in line 054 suggests she accepts James’ productions as acceptable approximations.

At no point in this extract, which lasts 91 seconds, does Sam verbally request production from James, yet he produces the stimulus 15 times in three separate sequences of interaction (lines 006, 027, 053). It is interesting that, on all three occasions, James produces the stimulus in sound strings, not as single sounds, mirroring the typical nature of Sam’s productions. Similarly, Sam does not verbally request that James participates in the nonverbal picture display, yet he both imitates it (line 027) and initiates it (highlighted in yellow) in several separate sequences. James’ verbal and nonverbal participation in this activity demonstrate a high level of shared attention and engagement.

In Extract 12, therapist Sarah and child Louise are playing with a fishing game. Pictures of a drum to represent the sound stimulus sit underneath wooden fish on the table surface. Louise is catching the fish with a toy fishing rod and returning them to the pretend water. The recurring steps are: 1. Child catches a 'fish' and puts it in the 'water' → 2. Child displays the drum picture → 3. Therapist (and child) produces the stimulus.

Extract 12: Fishing game with drum pictures (Sarah+Louise-V7-A2)

001 T What's the next one ((gaze to picture/fish))
 002 C ((catches fish with rod))
 003 T Ooh ((gaze to C))
 004 C ((pulls fish off the rod))
 005 T Very good ((points to the water/box)) in it goes
 006 C ((puts fish in the water/box))
 007 T This one can be (.) **di:** ((sustained gaze to C))
 008 C ((picks up drum picture; holds it facing the camera; gaze to
 009 camera; turns gaze to T)) **ji:**
 010 T **di:** ((gaze to C))
 011 C ((puts picture in the water/box))
 012 T Good girl ((gaze to 'water'; removes picture; puts it on table))
 013 C ((catches another fish))
 014 T And the last one ((gaze to C))
 015 C ((pulls fish off))
 016 T Ooh ((gaze follows the fish))
 017 C ((reaches with fish towards the 'water'; brief gaze to T))
 018 T ((gaze to C)) that's a big pull, isn't it
 019 C ((drops fish in the 'water'; [picks up the drum picture; gaze to
 020 picture then camera; holds picture facing camera))
 021 T [Let's do that one again
 022 Watch me ((reaches towards C's arm; gaze to C))
 023 C ((turns gaze to T; holds picture facing camera))
 024 T [**di:** ((makes cued artic gesture in synchrony; sustained gaze to C))
 025 C [**da:** ((makes part cued artic gesture; sustained gaze to T))
 026 T Oh, you're doing **da:** ((makes cued artic; gaze to C; smiles))
 027 C ((sustained gaze to T; nods))
 028 T I wanted it to say↑ ((takes picture from C)) (.) **di:** ((makes cued
 029 artic; sustained gaze to C))
 030 [((postures artic))
 031 C [((lowers gaze)) **gi:** ((lifts gaze to T))
 032 T Ooh ((points to mouth)) [with our tongue
 033 C [((postures linguolabial tongue placement))
 034 T [**di:** ((makes cued artic gesture; sustained gaze to C))
 035 C [**di:** ((gaze to T))

The stimulus in this activity is the voiced linguolabial plosive [d̥] in consonant (C) and vowel (V) sound sequences. The CV stimuli are symbolised by pictures of drums and the sound stimulus orientates to the phoneme /d/. In this extract, V = [i:] or [α:]. Therapist Sarah introduces the stimulus with a brief pause (lines 007, 028), rising prosody (line 028) and alongside some, though not all, productions makes a nonverbal cued articulation finger gesture (lines 028-029). In Extract 4, lines 010 to 014, featuring the same therapist-child dyad, we saw Sarah produce multiple repetitions of the sound stimulus with child-directed gaze, and then re-direct her gaze to the camera and terminate the sequence. In the present extract, it is Sarah's sustained gaze beyond her stimulus productions in lines 007 and 028-029 that indicates the sequences are not yet complete.

The first *D1. Demonstration for imitation* occurs in line 007. Louise's response (lines 008-009) provides evidence of Sarah's action. We see her display the drum picture to the camera (line 008), establish mutual gaze with Sarah and produce the stimulus (line 009). The author transcribed Louise's production as a voiced palatal plosive, which matches the stimulus on voicing and manner, but is a mismatch on placement. In relation to psycholinguistic theory, this mismatch represents an inaccurate motor program. Sarah's sustained gaze as she repeats the stimulus (line 010) suggests she has perceived the mismatch, but she does not reveal this to Louise. When Louise terminates the sequence by putting the fish in the water, Sarah provides verbal praise and allows the activity to continue. Sarah's subsequent invitation in lines 021-022 supports the hypothesis that she perceived the (place of articulation) mismatch in Louise's production in line 009: "*let's do that one again... watch me*" (lines 021-022). This represents a delayed other-initiated repair and Louise's overlapping production in line 025 shows that she treats it as a *D4. Verbal request for production*. The absence of therapist stimulus production in this sequence means Louise's production in line 025 is generated from stored representations. Louise's production involves a different vowel to the sound sequence demonstrated in the previous stimulus routine sequence (see line 007) but her dentalised production is a closer match to Sarah's linguolabial stimulus. The change in articulatory placement from Louise's production in line 009 to her production in line 025 suggests she had pre-existing knowledge of the placement requirement for 'the drum sound', and access to this may have been stimulated by Sarah's verbal invitation to 'watch me' (line 022). Sarah's reaction (line 026) suggests she was not expecting Louise to generate a production from stored representations. Again, this provides further evidence that her productions in this extract are demonstrations fulfilling the action of inviting stimulus imitation.

We see another *D1. Demonstration for imitation* in lines 028-029. Louise's production in line 031 reinforces the nature of Sarah's action. In this sequence, we see another mismatch of similar nature (inaccurate placement) to Louise's earlier production in line 009. Unlike last time, this time Sarah immediately reject Louise's production and implicitly invites her to repair her production using a combination of verbal and nonverbal features: "ooh: ((points to mouth)) with our tongue [*di*:] ((makes cued artic gesture; sustained gaze to C))." In line 035, we see Louise effectively revise her production; this production is a clear match to Sarah's earlier stimulus productions (lines 007, 010, 024, 028, 034). Interestingly, despite Sarah's specific verbal articulatory descriptor, the timing of Louise's articulatory posturing in line 033 and her subsequent production in line 035, which overlaps Sarah's stimulus production, mean they are not responsible for the success of Louise's repair in this sequence. Instead, Louise's response (lines 033 and 035) further supports the suggestion that she had pre-existing knowledge of the stimulus and therapy process that equipped her to be able to revise her production effectively from Sarah's rejection and non-specific element of her repair initiation, "ooh" (line 032). We will revisit this in Chapter 8.

7.5.2 D2. Incomplete demonstration

Incomplete utterances are typically associated with eliciting a response (Koshik, 2002). In this type of action, the therapist introduces stimulus production with a lead-in utterance, but pauses in the stimulus slot, (typically utterance-final position), creating opportunity for the child to produce the stimulus and complete the demonstration. Other turn design features are child-directed gaze (in the stimulus slot) and interrogative prosody, which may be rising or flat with an elongated vowel. In addition, the therapist may withhold the stimulus picture or toy. Other studies have also found these features mobilise response (Gardner 1994; Ronkainen, 2011; Stivers & Rossano, 2010;). Characteristic of this sub-category of action is that the absent element in the therapist's utterance is a stimulus production. As such, it is therefore a moderately explicit invitation to the child to produce the stimulus. Similar to the sub-category of actions in *C. Invitation to produce the stimulus*, the present sub-category of action requires that the stimulus routine has been enacted in previous sequences so that the child has pre-existing knowledge of the missing element. The first example of the action *D2. Incomplete demonstration* is illustrated in Extract 13, which features therapist Laura and child Ellie. Laura is introducing pictures for Ellie to post in a toy post-box. Lines 001 to 010 illustrate the recurring steps of the activity: 1. Therapist displays and names a picture representing a real word stimulus → 2. Child posts the picture.

Extract 13: Posting word-final /ʃ/ pictures (Laura+Ellie-V4-A1)

001 T Oh: let's see first picture ((gaze to camera)) I've found to post
002 [((reaches picture towards camera))
003 C [((gaze to picture then follows picture to camera))
004 T I:s ((moves picture by face)) a **bʌʃ:** ((finger to mouth))
005 C [((plays with post-box))
006 T ((soft gaze to C)) **bʌʃ:** ((puts picture on table)) can you put it in
007 the post-box
008 C ((picks up the picture; [posts it))
009 T [bʌʃ: ((finger to corner of mouth; gaze
010 to camera)) good posting
(0.15)
011 T And oh look (.) he's [having a: [:(((holds a picture facing C; gaze
012 to C))
013 C [((gaze to picture))
014 [((reaches for picture))
015 (2.0)
016 T ((moves picture and gaze to camera)) let's show the camera
017 C ((gaze follows picture/to camera))
018 T ((picture by face)) **wɒʃ:**
019 [((finger to corner of mouth))
020 C **wɒʃ:** ((plays with post-box))
021 T **wɒʃ:** ((shift in gaze toward but not to C; finger to mouth))
022 ((turns gaze to C)) that's a long sound from my mouth ((points))
023 C ((turns gaze to T; reaches hand towards picture in T's hand))
024 T **wɒʃ:** ((on back of her hand)) you can feel it tickle again
025 ((takes C's hand)) **wɒʃ:** ((on C's hand))
026 C ((gaze to T))
027 T ((lifts gaze to C))
028 C **wɒʃ:** [((takes T's hand))
029 T [Tickle tickle
030 C **wɒʃ:**
031 T Good try, good try ((nod; smiles))
032 ((turns gaze to camera)) **wɒʃ:** in it goes ((offers picture to C))
033 C ((takes picture and posts it))

The stimulus in Extract 13 is a lengthened post-alveolar fricative [ʃ] in real words *brush* and *wash*. Laura makes a finger gesture alongside stimulus production, where she points to the corner of her mouth (lines 004, 009, 019, 021). She also provides verbal articulatory description (line 022) and tactile stimulation (line 025). In the first sequence, Laura introduces real word *brush* in lines 001-002. Her manipulation of the picture, displaying it to the camera

rather than to Ellie, is not unlike *A1. Demonstration for later*. She then turns to Ellie and repeats the stimulus (line 006). Ellie's gaze and manipulation of the post-box suggest her attention is not on the stimulus, though Laura does not treat this as problematic. She instead instructs Ellie to post the picture. Ellie complies (line 008) and Laura re-directs her gaze to the camera and repeats the stimulus (line 009-010). Laura praises Ellie (line 010) to conclude this sequence. We re-join the activity 15 seconds later. Laura is introducing a new real word stimulus, *wash* (lines 011-012, 015). This turn contains the response mobilising features of child-directed gaze, elongated vowel ("a:" line 011) and a pause in the stimulus slot (line 015). Ellie's gaze is to the picture, but she does not take up the invitation to complete Laura's utterance. Laura does not treat this as problematic, going on to produce the stimulus herself (line 018). Although Ellie did not take up the invitation to produce in line 017, we see her imitate Laura's production in line 020. Ellie produces an active nasal fricative, [ŋ^{h}]. The nature of this mismatch with Laura's stimulus production suggests lexical processing has occurred, resulting in implementation of an inaccurate motor program. Laura's evaluation (lines 021-022, 024-025), which includes repeat stimulus productions, a specific articulatory descriptor and tactile stimulation, mobilises an (unsuccessful) repair by Ellie in lines 028 and 030.

A second example is illustrated in Extract 14, featuring therapist Helen and child Naomi, in the activity that we visited previously in Extract 6. Helen, Naomi and Naomi's mother are playing Puff the Pop-Up Dragon. The activity has the recurring steps: 1. Child chooses a flag/picture → 2. Therapist (or mother) produces the stimulus → 3. Child (or mother) inserts the flag.

Extract 14: Puff the Pop-Up Dragon (Helen+Naomi-V3-A1)

001 T Which one are you picking ((points to flags/pictures))
 002 C Umm ((puts hand on a flag/picture and moves it toward herself))
 003 ((turns gaze to T))
 004 T ((meets C's gaze))
 005 C ((moves flag/picture towards T))
 006 T Okay a blue ((reaches [to flag/pic]) flag and what's on your picture
 007 C [((moves flag/pic away))
 008 Hmm ((picks up the flag and picture; gaze to picture))
 009 The baby baby ((reaches picture to T))
 010 T ((takes picture; turns it round to face C; gaze to C))
 011 [And when the baby's sleeping [we have to say↑
 012 (.)
 013 C [((turns to face the castle; flag in hand))
 014 [((posts the flag)) ↴ ((grimace))
 015 M Good girl ((nods))

016 T ((lifts gaze to M; smiles)) really good, we do
017 C I'm gonna put this one here ((posts the flag; turns gaze to T))
018 T No pop ((gaze to C))

In line 002, we see Naomi choose a flag to reveal a hidden picture and Helen ask her “...*what’s on your picture?*” Since the picture symbolises the sound stimulus, this may have been an attempt by Helen to elicit a stimulus production. Naomi’s response in line 009 is not a stimulus production, but it is an appropriate response; she correctly identifies that the picture is of a baby. Naomi remains engaged in the activity as she hands Helen the picture and turns to post the flag. Helen’s turn in lines 010-012 is an example of the present action, *D2. Incomplete demonstration*. Her gaze and body orientation are towards Naomi and she produces the lead-in utterance “*and when the baby’s sleeping we have to say...*” with rising prosody and an overt pause in the stimulus slot. The pause is only brief since Naomi completes the demonstration without hesitation. It is noteworthy that Helen uses a plural rather than singular pronoun in her utterance (line 011) but all other features of the utterance and its delivery suggest she is inviting Naomi to produce the stimulus and complete the utterance. Naomi’s gaze is to the castle and her physical attention appears to be on posting the flag, yet she still produces the sound stimulus, and this is accepted by both Helen and her mother. Naomi successfully posts the flag and the sequence draws to a close (lines 017-018). Naomi’s response is evidence that the design of Helen’s utterance made it an invitation for a stimulus production and, since Naomi’s gaze was not to Helen as she delivered it, it would seem to be the prosodic and linguistic features, rather than Helen’s eye gaze, that are key.

7.5.3 D3. Nose-closing to elicit production

Similar to *C2. Gesture to elicit participation*, the present sub-category of action is characterised by a nonverbal behaviour that symbolises stimulus production. Nose-closing is a diagnostic and therapeutic technique associated with demonstrating or facilitating accurate, oral sound production (Piggott, 2014; Piggott & Jeyes, 2014). In a demonstration, the therapist closes her own nose by gently occluding her nares with her thumb and finger, or with a finger from each hand, as she produces the stimulus. In a facilitation, the therapist occludes the child’s nares during stimulus production. It enables the child to experience oral pressure consonant production without loss of airflow or pressure nasally and without nasal resonance. It is particularly useful for children with inaccurate nasal motor programs for oral consonants (to support formation of new, accurate motor programs) and children with an impairment in

motor execution affecting the production of motor programs for oral sounds (to provide optimum auditory feedback) (Calladine & Vance, 2019). When a child has existing knowledge of this technique and its association with stimulus production, the therapist's gesture of closing the child's nose represents an invitation for production.

In Extract 15 we return to therapist Sam and child James, who we saw in Extract 10. They are taking part in an activity that has the recurring steps: 1. Child chooses big or little ball → 2. Therapist produces the stimulus → 3. Child posts the picture.

Extract 15: Posting ball pictures (Sam+James-V14-A2)

001 T Now then, here we are, do you want a big [ball ((displays picture
 002 of big ball)) [to put in there ((gestures posting)) or a
 003 [little ball ((displays picture))
 004 C ((gaze to picture))
 005 ((gaze follows picture))
 006 ((gaze to picture))
 007 T Big one ((shakes picture)) or little one ((shakes picture))
 008 C ((reaches for big ball picture))
 009 T Big one! ((withholds picture))
 010 C (p) (p) [(p) (p) ((gaze to T; holds picture in T's hand))
 011 T [Good boy ((points to C)) it does it loud ((occludes own
 012 nose)) b [bb like that
 013 C ((lowers gaze to picture; takes picture))
 014 T But you used ((points)) your lips didn't you good boy ((gaze to C))
 015 C ((posts picture in post-box))
 016 T In there ((waves))
 017 [(turns gaze to camera)) bb
 018 C [(opens post-box and peers inside))
 (2.07)
 019 T Shall we have some more ((gaze to C))
 020 Look, there's [more ((displays ball pictures))
 021 C ((turns gaze to pictures)) m m m ((smiles))
 022 ((turns gaze to M))
 023 T You're right! ((taps C on arm)) clever boy!
 024 C ((reaches for picture))
 025 T That's the bouncy ball sound ((holds picture facing C; points to
 026 pic; gaze to camera)) let's show the telly ((points to camera))
 027 C ((turns to post-box; hand on picture))
 028 T Let's show the telly ((withholds the picture))
 029 ((occludes C's nose; gaze to C)) bb
 030 (1.0)
 031 b

032 C ((turns gaze to T; scrunches nose))
 033 T Are my hands cold ((laughs; gaze to C))
 034 C ((turns away; [opens the post-box door])
 035 T [((holds door)) there we are put it in bye bye
 036 ((turns gaze to camera)) ball

Therapist Sam's gestural invitation to produce the stimulus occurs in line 029 (highlighted in grey), and is categorised as such because of the prior demonstration in line 011-012. Although Sam produces the stimulus herself, she does pause between her second and third productions, providing opportunity for James to imitate. James tolerates nose-closing but does not produce the stimulus and this absent response is not treated as problematic; we see no further pursuit and James is permitted to post the picture. In a previous interactional sequence, beginning in line 026, we see Sam produce the utterance "let's show the telly" followed by a pursuit in line 028. There is ambiguity in the verb *show* in this context. *Show* may mean 'make the ball sound for the camera', in which case Sam's utterance serves as an invitation to produce. However, it may instead mean 'take hold of the picture and display it for the camera', serving as an invitation to participate. Sam's repeat of the nose-closing gesture in her proceeding turn (lines 029-031) provides evidence to support the former.

7.5.4 D4. Verbal request for production

This is the most explicit form of *D. Invitation to produce the stimulus* observed in the therapy data. It is also the most demanding type of action, requiring the child to produce the stimulus him- or herself. Turns operating as explicit invitations are designed like directives; they either instruct or request. The turn may or may not include a stimulus production: if it does, it functions as a model for the child to imitate; if it does not, the child is required to formulate output from stored lexical representations (Stackhouse and Wells, 1997). Extracts 16 and 17 illustrate this type of therapist action.

In Extract 16, therapist Helen, child Naomi and Naomi's mother are playing with the game *Monkey Business* (produced by Early Learning Centre). They each have turns in the activity. The activity consists of the recurring steps: 1. Spin the spinner to select the number of monkeys to hang → 2. Choose monkey colour(s) → 3. Produce the stimulus → 4. Hang the monkey(s). We join the activity on Naomi's turn part way through step 1; her spin has selected number two. Although therapist Helen directs Naomi to choose a colour (Step 2), she already has an orange monkey in her hand.

Extract 16: Monkey Business (Helen+Naomi-V5-A1)

001 T So what colour are you gonna do then ((gaze toward C))
002 C ((holds monkey in her hands; gaze to monkey))
003 T Are you gonna do orange as well
004 C Hmm ((gaze to tree)) [() ((stands up; wobbles; hangs
005 the monkey))
006 T [((postures cued artic gesture; gaze camera))
007 Oops ((turns gaze and body towards the tree))
008 Can we hear your sound then
009 C ((sits down; gaze to T; smiles))
010 T ((gaze to C; smiles)) he's next [((points)) to mine and they both
011 went↑ ((postures cued artic gesture; sustained gaze to C))
012 (2.0)
013 C [((turns gaze to the monkey))
014 Hang ((gaze to T))
015 T They did both hang ((gaze C)) and they both said↑ ((postures
016 artic; makes cued artic gesture))
017 (1.0)
018 C [ʃ: [((accompanying facial grimace; turns gaze to monkey))
019 T [ʃ: ((makes cued artic gesture)) lovely

The sound stimulus is the voiceless post-alveolar fricative [ʃ] produced as a lengthened fricative and with accompanying nonverbal cued articulation gesture. The present action *D4. Verbal request for production* occurs in line 008, though it emerges in line 006 as therapist Helen postures the cued articulation gesture in readiness. Although this is in accordance with the sequence of steps that has been established, i.e. Step 3. Stimulus production, it appears to be mistimed as Naomi's visual and auditory attention are on the tree; she is stood up attempting to hang the monkey, giving a narrative of what she is doing (line 004). Helen acknowledges this when Naomi returns to her seat, "he's next to mine..." (line 010), but her proceeding utterance is characteristic of a pursuit, this time in the form of *D2. Incomplete demonstration*. Helen's cued articulation posture is further evidence that her invitation is for production, but there is no evidence that Naomi observes this as her eye gaze is to the monkey tree. Naomi's response, "hang" (line 014), is not linguistically inappropriate, and Helen acknowledges this, "they did both hang" (line 015). In lines 015-016, we see Helen make a further pursuit with a turn that has characteristics of *D2. Incomplete demonstration*. This time, Helen uses a verb more explicitly related to production, *said*, and when she resumes her cued articulation gesture, she appears to have mutual gaze with Naomi. Naomi's production response in line 018 supports Helen's *D2 Incomplete demonstration* and suggests the

increasing explicitness that she incorporated into her pursuits was effective. Naomi's production is not an exact match for the stimulus; there is perceptible accompanying audible nasal emission and observable facial grimace. However, the nature of these features is passive, and the hypothesis evident in Helen's evaluation (in line 019) is that Naomi has implemented an accurate motor program for the stimulus.

A second example is shown in Extract 17. Therapist Laura and child Ellie are engaged in a pretend play activity involving toy figures. Therapist Laura invites Ellie to decide whether the figure is going to play or go to sleep. The sound stimulus is a lengthened voiceless post-alveolar fricative [ʃ:], which Laura typically produces with an accompanying finger gesture. The stimulus is symbolised by the acts of 'telling' the figures to be quiet when they are playing because others are sleeping, and saying 'night night' when putting figures to sleep. The recurring steps in the activity are: 1. Choose (and name) a figure → 2. Decide if the figure is going to play or go to sleep → 3. Tell others to be quiet/produce the stimulus.

Extract 17: Pretend play with toy people (Laura+Ellie-V6-A1)

001 T ((holds bag open))
 002 C ((gazes inside the bag; reaches hand inside; gaze to T)) it B
 003 ((pulls a figure out of the bag; gaze to figure))
 004 ((turns to D)) daddy that's ((points to figure)) you ((points to D))
 005 D (laughs) ((pats side of C's face with his hand; gaze to C))
 006 C ((turns around; stands the figure on the table))
 007 T Where's daddy gonna go is he gonna go and play or go to sleep
 008 ((gaze to C))
 009 C Play
 010 T He's gonna play oh so w- what do we have to say to daddy to make
 011 sure he doesn't wake up mummy ((sustained gaze to C))
 012 [(1.0)
 013 C [((points to the figures))
 014 T We have to say: [((gaze to camera; makes finger gesture))
 015 C [And that ((points to figures))
 016 T ʃ: daddy
 017 C [And that [and that and that ((picks up daddy figure))
 018 T [((turns gaze to C))
 019 C ((moves figure toward T))
 020 T ((gaze to figure)) and that one ((makes finger gesture)) ʃ:
 021 Don't ((turns gaze to camera)) wake up mummy ((makes finger
 022 gesture)) ʃ: ((turns gaze to C))
 023 C And (Laura)
 024 T Don't wake up (Laura) ((shakes head; smiles; gaze to camera))

025 Definitely don't wake up (Laura)
026 C ((plays with the figures))
027 (4.0)
028 T There you go daddy's gonna go and play ((moves daddy figure away))
029 There he is
030 C And he can go play ((moves another figure away))
031 T He's gonna go and play ((opens the bag)) who else

We join the activity as Ellie selects a figure from Laura's bag. She initially identifies it as her brother, B (line 002), but then changes this to her father (line 004). Laura asks her if daddy is "...gonna go and play or go to sleep" (line 007). Ellie chooses "play" (line 008) and Laura confirms her decision (line 010). Laura goes on to ask Ellie "what do we have to say to daddy to make sure he doesn't wake up mummy" (lines 010-011) with child-directed gaze. Laura's utterance features an explicit verb relating to production, *say*, and, since this instruction to the figures has been associated with the sound stimulus, this is an explicit invitation for stimulus production, i.e. *D4. Verbal request for production*. Laura sustains her gaze to Ellie and pauses at the end of her question, creating opportunity for Ellie to produce the stimulus. Laura does not produce the stimulus so if Ellie does it would be from stored representations. On this occasion, Ellie does not produce the stimulus; she points to the figures and her gaze is to the figures. Laura goes on to produce the stimulus herself with camera-directed gaze (lines 014, 016). That is, rather than pursue Ellie for a production she steps down the demands on her in the local here and now by delivering an *A1. Demonstration for later*. She also allows Ellie to direct play by responding to her manipulation of the daddy figure and her verbal commentary (see lines 017-026). There is no evidence in this continuing trajectory of turns to suggest that Laura treats the absence of a stimulus production (by Ellie) as problematic. In line 028 we see Laura draw this play sequence to a close and, in line 031, initiate a new one; she opens the bag and invites Ellie to choose a new figure.

7.6 Conclusions

This chapter has presented the findings from detailed analysis of 120 minutes of video recordings from 12 MSIVT sessions involving four therapist-child dyads. Therapy took place in three specialist SLT services in the NHS. In Phase 1, analysis revealed that MSIVT sessions typically consist of five therapy activities. In Phase 2, conversation analysis of therapist-child interaction during MSIVT activities examined the structure of such activities as a series of therapist-initiated stimulus routine sequences. Central to these sequences are stimulus

productions by the therapist and/or the child. A key finding from the analysis was the nature of therapists' verbal and nonverbal behaviours alongside stimulus productions, and the toys and pictures they used. These behaviours and materials established symbolic associations with the sound stimulus and embedded the stimulus production as one element of a stimulus routine.

Detailed analysis of therapy talk sequences found that therapists' initiation turns fulfilled four macro categories of action, distinct by their turn design features and the types of child responses they engendered. Therapists exhibited both a range and myriad of verbal, nonverbal and prosodic features in their turns, which appear to work well to establish the child's active engagement in the activities as well as make important features of the speech sound stimuli salient. They also have response-mobilising properties, which, either in isolation or in combination, distinguished the four macro categories of action and nine sub-categories within them. *A. Demonstration to the camera* did not engender a response from the child in the here and now of the activity. The therapist produced the stimulus with camera-directed gaze and absence of attention from the child was not treated as problematic. *B. Invitation to attend to the stimulus* and *C. Invitation to participate in the stimulus routine* typically engendered nonverbal responses from the child; either their attention to the stimulus or their participation in nonverbal elements of the stimulus routine. However, in Extracts 3 and 10, we saw that explicit forms of these invitations elicited stimulus productions from the child. The fourth category of action, *D. Invitation to produce the stimulus*, characteristically engendered a verbal response. Within these broad categories of action, the analysis uncovered nine sub-categories, ranging from implicit forms, where responses were mobilised by subtle nonverbal or prosodic practices, to explicit verbal requests, softened by subtle linguistic practices, such as plural pronouns and ambiguous verbs.

One of the distinctive features of the interaction was that therapists did not always pursue a response to their invitations when one was not forthcoming. This framed a range of accepted responses from the child and gave them freedom to choose whether and how to respond. This reflects the principle of the MSIVT approach to provide the child with opportunities to respond rather than elicit responses with a directive style of interaction. All four therapists interacted explicitly with the camera during MSIVT activities and used it as an interactive tool to support the delivery of therapy. Therapist stimulus demonstrations with camera-directed gaze did not demand, or typically elicit, a here and now response from the child. This somewhat unique type of action, *A1. Demonstration for later*, represents a demonstration for future use; for the child to attend to at a future moment in time. The camera was instrumental in facilitating the

sub-category of action *C. Invitation to participate in the stimulus routine*; this type of action gave the child opportunity to take part in meaningful nonverbal enactments with the toys and objects as well as engage in stimulus production. The camera also appeared to work well as a tool for downgrading demands on the child when they chose not to respond.

Chapter 8: MSIVT Interactions over an Episode (Phase 3)

8.1 Introduction

This chapter presents the results from Phase 3. This phase of the study aimed to examine therapist-child interaction during MSIVT from a longitudinal perspective in order to identify and describe change over the course of an episode of care.

The analysis focused on the first and last videos of therapist-child dyad Sarah-Louise's episode of care. These correspond to MSIVT sessions that took place seven months apart. Louise was aged 2;3 in the first session (V1) and 2;10 in the last session (V7). The analysis revealed differences in the nature of action Sarah initiated at these two stages of therapy, the responses that were engendered, and the effects this had on the sequential structure of the interaction. At the onset of therapy, Sarah showed a preference for action that did not demand a here and now responses or was minimally demanding, in that it invited a nonverbal response. In contrast, the end of therapy was characterised by a profile of more demanding action, in that it invited more verbal responses. At both stages in therapy, Sarah showed a preference for implicit rather than explicit forms of action. Louise demonstrates her engagement in therapy at both stages, but the nature and timing of her responses in the final session give rise to more jointly accomplished action and a different pattern of interaction. I begin the chapter with the findings from analysing the nature and relative frequencies of the different sub-categories of action at the two stages of therapy. I then present the findings from detailed analysis of sub-categories of action common to both stages of therapy and examine some specific types of sequences that only occurred in Video 7. I use extracts from the data to illustrate these findings from the analysis.

8.2 Profile of therapist action in Video 1

The 10-minute excerpt of Video 1 featured three MSIVT activities; see Table 8.1 for a description.

Table 8.1: Activities and symbolic representations in Video 1

	Activity 1	Activity 2	Activity 3
Description of activity	Toy sink and drippy tap, which releases water when pressed	Releasing car from track to land on a picture of a tap	Feeding toy snakes pretend food
Symbolic representation for sound stimulus	Toy sink/tap	Tap picture	Toy snake
Nonverbal behaviours in the stimulus routine	Pressing the tap	Displaying the picture for the camera; cued articulation finger gesture	Feeding the snake; cued articulation finger gesture

Analysis of the 10-minute excerpt identified 56 stimulus routine sequences. Analysis of these sequences using the framework presented in Table 7.1 in Chapter 7 (see page 125) identified five different sub-categories of action: *A1. Demonstration for later*; *B1. Demonstration for attention*; *B2 Verbal request for attention*; *C2. Verbal request for participation*; and *D1. Demonstration for imitation*. These actions represent each one of the macro categories from the least demanding, *A. Demonstration to the camera* to the most demanding, *D. Invitation to produce the stimulus*. Despite this range of action, one particular action dominated, as shown by the quantitative analysis presented in Table 8.2.

Table 8.2: Nature and frequency of action in Video 1

Action	Number of initiations	Percentage (/56)
A. Demonstration to the camera	31	55
A1. Demonstration for later	31	55
B. Invitation to attend to the stimulus	11	20
B1. Demonstration for attention	10	18
B2. Verbal request for attention	1	2
C. Invitation to participate in the stimulus routine	13	23
C1. Gesture to elicit participation	0	0
C2. Verbal request for participation	13	23
D. Invitation to produce the stimulus	1	2
D1. Demonstration for imitation	1	2
D2. Incomplete demonstration	0	0
D3. Gesture to elicit production	0	0
D4. Verbal request for production	0	0

Note. /56=out of 56 total initiations to give percentage of total initiations

Table 8.2 reveals a profile of low demanding action. The majority (55%) of initiations belong to the macro category *A. Demonstration to the camera*, followed by *B. Invitation to attend to the stimulus* and *C. Invitation to participate in the stimulus routine*, which featured at similar levels: 20% and 23%, respectively. Only 2% (1/56) of initiations belong to the macro category *D. Invitation to produce the stimulus*.

Sub-category *A1. Demonstration for later* was the most prominent action in Video 1. The second most prominent action was *C2. Verbal request for participation*, which represented 20% of the initiations. This involved explicit invitations for the child's participation in pressing the tap (V1-A1), displaying pictures to the camera (V1-A2) and feeding toy snakes (V1-A3). *B2. Verbal request for attention* and *D1. Demonstration for imitation* featured infrequently in Video 1. The majority (42/56; 75%) of Sarah's initiations were implicit forms of action: stimulus productions to the camera (*A1. Demonstration for later*) or stimulus productions with features that mobilised a nonverbal response (mostly *B1. Demonstration for attention*, plus one *D1. Demonstration for imitation*). A quarter (14/56; 25%) of Sarah's initiations were explicit forms of action and they all invited nonverbal responses: *B2. Verbal request for attention*; and *C2. Verbal requests for participation*. The most demanding and explicit sub-category of action, *D4. Verbal request for production*, did not feature at all in Video 1.

8.3 Profile of therapist action in Video 7

The 10-minute excerpt of Video 7 featured three MSIVT activities; see Table 8.3 for a description.

Table 8.3: Activities and symbolic representations in Video 7

	Activity 1	Activity 2	Activity 3
Description of activity	Finding pictures hidden around the room	Fishing game to reveal picture of a drum	Releasing car from track to make it go 'down'
Symbolic representation for sound stimulus	Tap and drum pictures	Drum picture	Toy car
Nonverbal behaviours in the stimulus routine	Displaying the picture for the camera; cued articulation	Displaying the picture for the camera; cued articulation	Releasing the car; cued articulation

Analysis of the 10-minute excerpt identified 45 stimulus routine sequences. Analysis of these sequences using the framework presented in Table 7.1 in Chapter 7 (see page 125) identified seven different sub-categories of action: *A1. Demonstration for later*; *B1. Demonstration for*

attention; B2 Verbal request for attention; C2. Verbal request for participation; D1. Demonstration for imitation; D2. Incomplete demonstration; and D4. Verbal request for production. As in Video 1, there are actions to represent each of the four macro categories. Table 8.4 shows the frequency and distribution of the different sub-categories.

Table 8.4: Nature and frequency of action in Video 7

Action	Number of initiations	Percentage (/45)
A. Demonstration to the camera	3	7
A1. Demonstration for later	3	7
B. Invitation to attend to the stimulus	4	9
B1. Demonstration for attention	2	4
B2. Verbal request for attention	2	4
C. Invitation to participate in the stimulus routine	7	16
C1. Gesture to elicit participation	0	0
C2. Verbal request for participation	7	16
D. Invitation to produce the stimulus	31	69
D1. Demonstration for imitation	25	56
D2. Incomplete demonstration	1	2
D3. Gesture to elicit production	0	0
D4. Verbal request for production	5	11

Note. /45=out of 45 total initiations to give percentage of total initiations

Table 8.4 reveals a profile of more demanding action. The majority (69%) of initiations belong to the macro category *D. Invitation to produce the stimulus*. In terms of sub-categories of action, one stood out in Video 7, *D1. Demonstration for imitation*, which represented 56% (24/45) of all initiations. This is an implicit form of action in macro category *D. Invitation to produce the stimulus*. In relation to turn design features, *D1. Demonstration for imitation* is distinct from other sub-categories of action by sustained child-directed gaze and silence following stimulus production. In contrast, for example, to *B1. Demonstration for attention*, which features lowered gaze and progression with the activity following stimulus production. A striking finding from the analysis, which I present later in this chapter, was that Louise's responses defined this action more strongly than its turn design features.

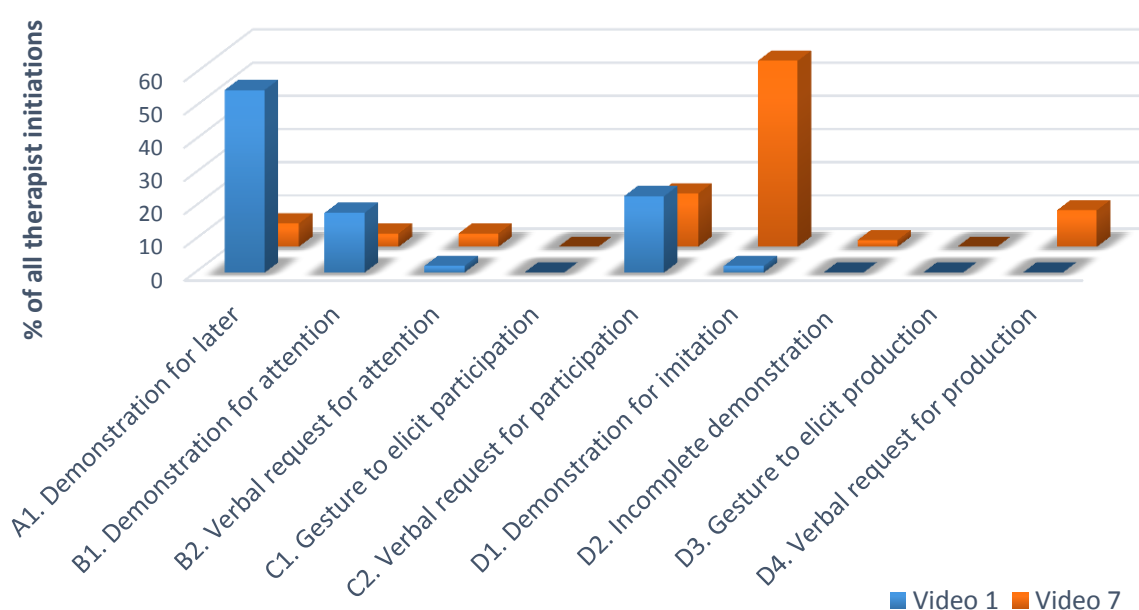
Implicit forms of action outweighed explicit forms: *A1. Demonstration for later*; *B1. Demonstration for attention*; *D1. Demonstration for imitation* and *D2. Incomplete demonstration* made up nearly three quarters (31/45; 69%) of all initiations. Explicit forms represented 31% (14/45) of initiations.

8.4 Comparative analysis of Videos 1 and 7

8.4.1 Nature and relative frequencies of action

Comparative analysis of the action profiles of Videos 1 and 7 revealed differences in the nature and relative frequencies of action at these two stages of therapy, as illustrated by Figure 8.1.

Figure 8.1: Nature and frequencies of action in Videos 1 and 7



Five of the nine sub-categories of action featured in both videos: *A1. Demonstration for later*; *B1. Demonstration for attention*; *B2. Verbal request for attention*; *C2. Verbal request for participation*; and *D1. Demonstration for imitation*. However, with the exception of *C2. Verbal request for participation*, the relative occurrence of different sub-categories differed markedly in the two videos. Two actions, *D2. Incomplete demonstration* and *D4. Verbal request for production*, which belong to the macro category *D. Invitation to produce the stimulus*, only featured in Video 7. These are highly demanding actions because they invite a production response. *A1. Demonstration for later* featured prominently in Video 1, representing 55% of all action, compared to 7% in Video 7. Whereas, *D1. Demonstration for imitation* featured prominently in Video 7, representing 56% of all action, compared to just 2% in Video 1. Interestingly, Sarah displays a preference for implicit action at both stages of therapy. Yet, the

overall nature of action in Video 7 is more demanding due to the greater presence of category *D. Invitation to produce the stimulus* initiations: 69% (31/45) of action in Video 7 compared to just 2% (1/56) in Video 1.

These findings illustrate differences in the nature and relative frequencies of different sub-categories of action in Videos 1 and 7, but they do not articulate the full extent of the interactional differences between these two videos. I will now present a series of extracts, along with detailed turn-by-turn analysis, to show qualitative differences in the practices Sarah and Louise used to accomplish some of the actions common to both stages of therapy.

8.4.2 Practices for initiating and responding to the same action

8.4.2.1 A1. *Demonstration for later*

A1. Demonstration for later is the least demanding action in the therapy data. It is characterised by one or more stimulus productions, typically made salient by adaptive articulation, and camera-directed gaze. In Phase 2, we found that this type of action did not engender a response from the child. One of the key findings from the analysis in Phase 3 was the recurrent nature of this action in Video 1; a practice that did not occur in Video 7. This practice bears resemblance with the clustered restart-relevant directives in the Reed et al. (2013) study.

Extract 1 features Sarah and Louise playing with a toy sink, which is positioned on the table in front of them. The recurring steps in the activity are: 1. Sarah presses the tap (to release the water) → 2. Sarah simultaneously produces the stimulus. Louise's mother (M) and sister (S) are also present and during the activity they each have turns at pressing the tap.

Extract 1: Toy sink and drippy tap (Sarah+Louise-V1-A1)

```

001 T [A want a turn?
002 S [(presses tap))
003 C ((gaze to sink))
004 T Hah, S's having a turn:: hah, can you see the water in there
005 ((gaze to C; points inside the sink; moves it closer to C))
006 There's water ((gaze to camera)) can you see ((gaze to C;
007 tilts sink toward C)) in there
008 C ((gaze to sink))
009 T Hah ((lifts gaze from sink/C to the camera))
010 [And when we make the taps [(taps the tap)) drip ((gaze to C))
011 C [(lowers gaze to the table)) [(gaze to the sink))
012 T It says: ((finger on tap; turns gaze to camera)) ttt

```


produces the stimulus as a string of multiple productions (e.g. lines 023, 036, 051), whereas in real words they feature as single productions (not as strings) and are embedded in short phrases, e.g. “*there’s water*” (line 006), “*I’m getting wet*” (line 048).

A1. Demonstration for later features 11 times in this extract (lines 006, 010/012-013, 018, 023, 025/027, 029, 036, 037-038, 043-044, 048, 050-051; highlighted in grey). The first, in line 006, features real word *water* produced with camera-directed gaze. Louise’s gaze is to the toy sink and tap, the object of shared attention, which her sister continues to press. The second occurs in lines 012-013, this time as a string of single sound productions and alongside her own pressing of the tap. Again, Louise’s gaze is to the sink/tap. In line 014, Sarah re-recruits Louise’s sister (S) as a participator in the activity, inviting her to have a turn with the tap, and produces a second string of stimulus productions (line 016). The change in gaze direction from the camera to Louise during this string of productions increases Louise’s accountability to attend, so this particular action is *B1. Demonstration for attention*. Mutual gaze for these three productions suggests Louise’s attention was successfully established. On completion, Louise’s gaze returns to the toy sink/tap (line 017). Sarah’s production of the real word stimulus *water* in lines 018-019 is another *A1. Demonstration for later*, strongly supported by Sarah’s gaze change just prior to the sound stimulus in the middle of the target word *water*. Another *A1. Demonstration for later*, featuring six stimulus productions, follows in line 023. In the *A1. Demonstration for later* actions in lines 025/027 and 036, Sarah introduces her stimulus productions with verbal utterances that are characteristic of *B2. Verbal request for attention*. Yet, on both occasions her gaze is to the camera, not to Louise, and there is no opportunity for Louise to respond.

This extract illustrates a predominance of *A1. Demonstration for later*, accomplished by a recurrent series of initiations by Sarah, in which she makes repetitive stimulus productions often alongside the nonverbal act of pressing the tap. These initiations do not invite a here and now response from Louise. However, the absence of any overlapping talk, and evidence of intermittent therapist-directed gaze by Louise, suggest she is displaying some degree of attention to what Sarah is doing. As well as fulfilling the dominant action *A1. Demonstration for later*, Sarah does create some local opportunities for Louise to engage in the activity, e.g. with frequent and intermittent child-directed gaze, verbal commentary, and occasionally inviting Louise to decide if the activity should continue. See, for example, lines 043-045, where Sarah asks Louise “*shall we do some more*” (line 044). Louise displays a gentle nod in response (line 045). The prominent nature of this action in Video 1 gave Louise the opportunity to observe a high level of stimulus production. In the 10-minute excerpt selected for analysis,

Sarah produced the stimulus 97 times as part of her 31 *A1. Demonstration for later* initiations, 87 at single sound level (typically as a string of multiple productions) and 10 in real words.

This action, *A1. Demonstration for later*, featured far less frequently in Video 7. Rather than it forming a series of recurrent action, it occurred in-between sequences of different action. Sarah used it as part of her turns in the activity and, it appeared, as a way of reinforcing stimulus productions during Louise's turns, as shown in Extract 2. In this extract, Sarah, Louise and Louise's mother (M) are taking turns with the Wooden Click Clack Track. The recurring steps: 1. Player chooses a car → 2. Sarah or player produce the stimulus → 3. Player puts the car on the track.

Extract 2: Wooden Click Clack Track (Sarah+Louise-V7-A3)

001 T Whose turn is it this time? ((gaze C))
002 C ((points to herself briefly; gaze T)) you ((points to T))
003 T My turn oh I'll have the red ((holds red car near face))
004 C L((nods; gaze T))
005 T I'm gonna make the car go↑
006 [(1.0)
007 L((turns gaze to the camera)) d[au:n ((makes cued artic gesture))
008 C L((turns gaze to track; points
009 to the top of the track)) dau:
010 T ((turns gaze to C)) [((puts car on top of track))
011 C L((gaze to the car))
012 T Ready, steady:↑
013 [(1.0)
014 T L((turns gaze to the camera)) d[au:n ((releases the car))
015 C ((gaze follows car))
016 T [Wo::w ooh ((stops the car))
017 C L((smiles; gaze to car))
018 T ((turns gaze to C)) nearly fell off the table then
019 [Oh is it mummy's go ((picks up blue car; gaze C))
020 C [I want () ((gaze to car in T's hands))
021 ((lifts gaze to T; nods))
022 T ((nods)) I think it's mummy's turn isn't it ((reaches car to M))
023 What's mummy gonna do? ((releases the car; gaze M))
024 M L I'm gonna make it go ((takes car))
025 d[au:n ((puts car on top of track))
026 C Ldau: ((gaze to M; points to track; gaze follows car))
027 M ((puts car on top of track and releases it))
028 C ((gaze follows car))
029 T It's going d[au:n whao::: ((gaze follows car))

The stimulus in this activity is a voiced linguolabial plosive [d̥] for the target phoneme /d/ in real word *down*. We join the extract on Sarah's turn to choose a car (line 003) and produce the stimulus. Her utterance in lines 005-006 "*I'm gonna make the car go...*" has prosodic turn-design features consistent with *D2. Incomplete demonstration*, but her explicit change in gaze direction during the silence, from Louise to the camera, provides evidence that this is an *A1. Demonstration for later*. However, Louise's production response in line 009 treats this as a *D1. Demonstration for imitation*. A similar thing happens in lines 022 to 026 on Louise's mother's turn with the car, when Louise once again produces the stimulus. Her mother's utterance "*I'm gonna make it go*" (line 024) does not have the rising prosody and silence features of Sarah's utterance in lines 005-006. Louise is showing orientation to the symbolic association made on prior turns between releasing a car down the track and producing the stimulus. The next *A1. Demonstration for later* in this extract occurs in line 014 as Sarah releases the car down the track. Her utterance "*ready steady...*" in line 012, is also produced with rising prosody and a pause in the stimulus slot, but once again, Sarah's explicit change in gaze-direction during the silence identifies this as *A1. Demonstration for later*. This time, it does not engender a verbal response from Louise.

These two examples of *A1. Demonstration for later* in Video 7 and the 11 examples we saw in Video 1 share the feature camera-directed gaze that defines this sub-category of action. However, the analysis shows differences in how and what they accomplish. In Video 1, the recurrent nature of this initiation did not invite or engender a response from Louise. In Video 7, Sarah's use of rising prosody and silence created opportunity for Louise to respond and, in one of the two examples in Extract 2, she took up the opportunity with an imitation of the stimulus.

8.4.2.2 C2. Verbal request for participation

A defining criteria for this action is that the specific participatory behaviour that the therapist requests has been associated with stimulus production. The nature of the behaviour and association will vary from activity to activity, as we saw in Chapter 7. Earlier in this chapter, we learned that this action represented 23% of initiations in Video 1 and 16% in Video 7. One specific initiation, inviting Louise to display stimulus pictures to the camera, featured in one activity in both videos, allowing comparative analysis of the specific turn design features and consequences. This action occurred six times in Video 1-Activity 2 and seven times in Video 7-Activity 1. To facilitate comparative analysis, I examined the first six instances in Video 7 and the six instances from Video 1; see Table 8.5.

Table 8.5: Turn design features of C2. Verbal request for participation initiations and second-pair part responses

	Therapist's verbal and nonverbal talk	Verb	Pronoun	Nature	Child response
Video 1-Activity 2					
1	Shall we hold it up and show the camera? (points) Show the TV	Hold, show	Plural	Request, directive	(nods; holds picture facing camera)
2	Shall we hold it up to the camera? (picks up the picture)	Hold	Plural	Request	(nods)
3	<i>Go on then, do you want to hold it up so the camera can see? (gives picture)</i>	Hold	Singular	Directive, request	(takes picture; holds it facing the camera)
4	Would you like to hold that picture up to the TV? (withholds toy)	Hold	Singular	Request	(nods)
5	<i>Go on then, you hold that one up, show the TV</i>	Show	Singular	Directive	(picks up the picture; holds it facing the camera)
6	Thank you, hold it up to the TV (points to the picture), let's have a look	Hold	Plural	Directive	(picks up the picture; turns it round to face the camera)
Video 7-Activity 1					
7	That was quick, right, shall we show the TV	Show	Plural	Request	(nods; turns picture to face the camera)
8	Shall we show it to the TV? (holds picture facing camera)	Show	Plural	Request	(nods; takes hold of picture; holds it facing the camera)
9	Do you want to show it to the TV?	Show	Singular	Request	(walks to the camera; holds picture facing camera)
10	Shall we show it to the s-camera? (points to the camera)	Show	Plural	Request	(turns round to face the camera/TV; holds picture facing the camera)
11	Hold it down to the camera (points to the camera)	Hold	NA	Directive	(picture in hand; turns round to face the camera)
12	Lovely, shall we show it to the TV? (points to the camera)	Show	Plural	Request	(holds picture facing camera)

Note. Text in brackets is nonverbal talk. Italics=pursuit utterance.

In both videos, Sarah used singular and plural pronouns and did not show a preference for one type over the other. She showed a slight preference for requests (questions) over directives (instructions) but there was little difference between the two videos. Requests give a recipient more opportunity to decline. There was some difference in verb preference. In Video 1, Sarah showed a preference for the more literal and explicit verb *hold*, whereas she showed a preference for the more ambiguous verb *show* in Video 7. *Show* could mean ‘produce the sound stimulus’ as well as ‘display the picture’. All Sarah’s C2. *Verbal request for participation* initiations in Video 1 engendered a nonverbal response from Louise. However, two of the initiations were pursuits for a more specific response; see #3 and #5 in Table 8.5. Louise’s responses treat Sarah’s initiations as yes/no questions rather than directives to display the picture. Both of Sarah’s pursuits begin with the directive “*go on then...*” and successfully elicit the target response from Louise. With the exception of examples #2 and #4, there is little other difference in Louise’s responses in the two videos. Extract 3 illustrates examples #4 (lines 003-004) and #5 (lines 006-007) from Video 1-Activity 2.

Extract 3: Wooden Click Clack Track with drippy tap pictures (Sarah+Louise-V1-A2)

```

001 T I think you've caught this one ((points to picture))
002 C [ ((picks up the car and puts it on the track))
003 T [Would you like to hold that picture up to the TV ((takes car
004 from C))
005 C ((nods))
006 T Go on then you hold that one up ((picks up picture)) show the TV
007 ((gaze to picture))
008 C ((takes picture from T; turns it around to face the camera;
009 gaze to camera))
010 T [ [ ((makes cued artic; gaze camera)) it says [ [ [ [ ((makes cued
011 artic gesture; gaze camera))
012 ((turns gaze to C)) doesn't it [ [ [ [ ((makes cued artic))
013 C [ ((lowers picture and gaze))
014 T [ ((gaze follows picture))
015 C ((reaches picture to T))
016 T ((takes picture from C))

```

In Extract 3, Louise has caught a tap picture with the car by releasing it down the track. The pictures symbolise the sound stimulus [t]. Sarah’s initiations in lines 003-004 and lines 006-007 are examples of C2. *Verbal request for participation*. Louise’s first response in line 005 indicates

she is engaged in the activity, but the more explicit pursuit by Sarah in lines 006-007 is required to accomplish the target action. Louise maintains her display of the picture while Sarah produces the stimulus eight times with camera-directed gaze (lines 008-011) and only lowers it as Sarah closes her turn in line 012 with three final productions, this time with her gaze to Louise.

The pursuits in Video 1-Activity 2 suggest Louise needed a little more support than she did in Video 7-Activity 1. This is despite Sarah's preference in that video for the more explicit verb. In Video 7-Activity 1, Sarah displayed a preference for the ambiguous verb *show*. Extract 4 illustrates example #7 in Table 8.5.

Extract 4: Hide & Seek and posting with tap and drum pictures (Sarah+Louise-V7-A1)

001 T Can you get one- oh you've got one: [good girl that was quick
 002 C [((picks up picture; reaches
 003 it toward T))
 004 T ((reaches to the picture but does not take it)) right shall we
 005 show the TV ((gaze to materials on the table; moves papers))
 006 C ((nods; turns gaze to the camera; [turns picture to face camera))
 007 T [And we can say what it-
 008 ((turns gaze to C)) oh ((turns gaze to camera))
 009 C [((postures interdental placement and cued artic))
 010 T [There we go ((turns gaze to C)) what sound is that?
 011 C [[ttt] ((gaze to T; makes cued artic gesture))
 012 T [[tttt] ((gaze to C; cued artic gesture)) you're using ((turns
 013 gaze to camera)) your tongue [((gaze to C)) brilliant
 014 C [((breaks gaze; steps toward the
 015 post-box))
 016 T [[ttt] ((gaze camera; cued articulation gesture))
 017 ((gaze C)) [do you want to pop it in
 018 [((points to post-box))

In Extract 4, Louise has found a tap picture that was hidden around the room and has brought it back to the table ready to post in the post-box. Sarah's initiation in lines 004-005 is an example of *C2. Verbal request for participation* and is typical of the *C2* initiations in this activity. Louise responds with an appropriate nonverbal response in line 006; she nods and turns her gaze and the picture to the camera. As she is doing this, Sarah begins to initiate a *D4. Verbal request for production* (line 007). However, she does this somewhat subtly with an absence of child-directed gaze and different focus of attention (see line 005), and she cuts her utterance short when her

gaze reaches Louise (see lines 007-008). As Sarah's gaze moves again to the camera, we see Louise posturing both the articulatory placement for stimulus production and the hand position for the corresponding cued articulation finger gesture (see line 009), and she goes on to produce the stimulus with accompanying finger gesture in line 011. Although the sequential position of Louise's posturing follows Sarah's emerging *D4. Verbal request for production*, the focused nature of her gaze and picture display, and the absence of these features in Sarah's initiation, suggest it may have instead been part of her response to Sarah's *C2. Verbal request for participation* in line 006. Even though Sarah initiates a new *D4. Verbal request for production* in line 010 "what sound is that?" Louise's productions that follow in line 011 can clearly be associated with the posturing that preceded it.

8.4.2.3 D1. Demonstration for imitation

This therapist action is characterised by a stimulus production and sustained child-directed gaze. It is an implicit form of invitation belonging to the most demanding category, *D. Invitation to produce the stimulus*. This action occurred just once in the 10-minute excerpt of Video 1, representing just 2% of all Sarah's initiations. In contrast, it featured prominently in the 10-minute excerpt of Video 7, representing 56% of initiations. The low level of occurrence in Video 1 provides little data for comparison, but the analysis did reveal some interesting findings. In Video 1, the *D1. Demonstration for imitation* appeared after a series of low demanding initiations in the form of *A1. Demonstration for later* and *B1. Demonstration for attention* and it is not a strong example of this form of action. In contrast, in Video 7, it formed a dominant part of the interaction and was a systematic part of a recurrent stimulus routine. Its appearance in the interaction resembled the dominant *A1. Demonstration for later* initiations in Video 1. A key finding from the analysis was the high level of output these initiations elicited from Louise in Video 7 and the collaborative nature in which this action was accomplished. Extract 5 illustrates the example of this action that occurred in Video 1. Extracts 6, 7 and 8 illustrate examples from Video 7. This action featured in all three activities in this video, so the extracts present at least one example from each activity.

Extract 5: Toy sink and drippy tap (Sarah+Louise-V1-A1)

001 T I think it's [C's ((gaze S)) turn
 002 M [((rolls up C's sleeve))
 003 It's [C's go
 004 T [Ready [C ((gaze C))

003 C ((nods))
004 T |Where's some more ((gestures 'where'))
005 C ((bends down))
006 T Oh under the |ta:ble
007 C |((stands up; holds picture; both facing T))
008 T Oh, another drum sou:nd d|d d d d ((cued artic; sustained gaze C))
009 C |((releases picture from hand and moves
010 into cued artic finger gesture))
011 |d d d d ((cued artic; facing T))
012 T You're using your tongue beautifully
013 |((gazes and points to the post-box)) good girl
014 C |((leans forward and |posts the picture in post-box))
015 T |Pop it in

In Extract 6 from Video 7, Louise is looking for pictures that Sarah has hidden around the therapy room. In lines 005 and 007, we see her find a picture under the table and return to the table with it in her hand. The picture is of a drum, which symbolises the stimulus, a voiced linguolabial plosive [d̥]. Sarah initiates a *D1. Demonstration for imitation* in line 008. She acknowledges Louise's picture, labels it "...drum sound" and produces a string of five stimulus productions with accompanied and synchronous cued articulation gestures and sustained child-directed gaze. Louise's overlapping production in line 011 is evidence of this action. There is no opportunity in this sequence to identify the nature of action from the design features that distinguish it from *B1. Demonstration for attention* (sustained gaze and pause following stimulus production) because of the timing of Louise's imitation. Louise begins to imitate after just one of Sarah's productions. She simultaneously releases the picture from her right hand to adopt the hand position for the corresponding cued articulation gesture, which she does in synchrony with her own productions, mirroring Sarah's enactment. Sarah continues with her string of productions alongside Louise and does not show any element of surprise toward Louise's response, providing further evidence of the action. This example shows Louise fulfilling a collaborative role in defining and accomplishing the action, which we did not see in Video 1. Sarah's specific feedback in line 012 ("you're using your tongue beautifully") accepts and affirms Louise's productions and ends the sequence. Extract 7 takes us into the second activity of Video 7.

Extract 7: Fishing game with [d] CV sound sequences (Sarah+Louise-V7-A2)

001 T And I |((takes pic from C)) can make up a little word

002 C L((gaze follows picture))
003 T ((turns picture around to face C))
004 I can say [((intake breath; gaze to C)) d[ɜ: ((cued artic; holds
005 picture; sustains gaze))
006 C L((lifts gaze to T)) L[ɜ: ((gaze to T; makes
007 finger gesture))
008 T d[ɜ: ((higher pitch)) very good
(0.50)
009 C [((reaches for the picture))
010 T L((reaches for the picture))
011 L((intake breath)) so this one can [say:: ((sustained gaze to C))
012 (2.0)
013 L((allows C to take picture))
014 C ((gaze to picture; turns it round to face the camera))
015 T [d[ɜ: ((makes cued artic gesture))
016 C L((lifts gaze to T))
017 L[ɜ: ((makes finger gesture))
018 T Oh is it saying d[ɜ: ((smiles; sustained gaze))
019 C ((nods; sustained gaze to T))
020 T Good girl [you used your sound ((relaxes gaze))
021 C L((lowers gaze))

Extract 7 features the same sound stimulus, this time at sound sequence level. Sarah's initiation in lines 001, 003-005 is an example of *D1. Demonstration for imitation*. Although Louise begins to imitate in line 006, (which appears to go unnoticed by Sarah), it is an incomplete production and she quickly goes on to produce a full consonant-vowel sequence in line 007 with accompanying finger gesture and mutual gaze. Sarah accepts this in line 008 with an affirming production and non-specific praise "very good". Another sequence takes place in lines 009 to 021. Sarah's *D1. Demonstration for imitation* initiation begins in line 010 (highlighted in grey) and continues to line 015. It begins as she reaches for the drum picture and says, "So this one can say..." (lines 010-011). As she does this, she allows Louise to take the picture from her. In line 014, we see Louise turn the picture round to face the camera. This nonverbal enactment is characteristic of the stimulus routine we saw in the previous activity (Video 7-Activity 1) and have seen in previous sessions, e.g. Video 1-Activity 2. As she does this, she lifts her gaze from the picture to Sarah and produces the stimulus with accompanying finger gesture. This response by Louise overlaps Sarah's production and gesture in line 015. Sarah's surprised reaction in line 018 provides evidence to support her initiation as *D1. Demonstration for imitation*, but the timing of Louise's production suggests she

has not treated it in this way. One hypothesis is that the beginning of Sarah's initiation (lines 010-012) served as *D2. Incomplete demonstration*, which seems plausible given the prosodic design of this utterance. An alternative hypothesis is that the picture itself or the act of taking it and displaying it for the camera initiated production. This type of sequence occurred recurrently in this activity and is interesting because it shows Louise using past experience to shape her current participation. The final examples, taken from the third activity in Video 7, are in Extract 8.

Extract 8: Wooden Click Clack Track (Sarah+Louise-V7-A3)

001 T Would you like red or green ((offers cars))
 002 C I want ((reaches for green car))
 003 T The green one ((withholds car))
 004 C ((releases hand from the car))
 005 T So can [you make the car go↑ ((gaze to C)) **d̥aun** ((makes cued artic))
 006 C L((lifts gaze to T))
 007 T [(sustained gaze to C)]
 008 C **dau:** ((nods; reaches for the car))
 009 L(2.0)
 010 T **d̥aun** ((makes cued artic; sustained gaze to C))
 011 C [dau: (n̥)] ((exaggerated artic for vowel; gaze to T))
 012 T I'm using my tongue ((points to mouth)) [d̥aun ((makes cued artic))
 013 C Ldau: ((exaggerated artic
 014 for vowel; gaze to T))
 015 T ((offers car)) lovely
 016 C ((takes car; turns to the track))

Extract 8 features the same sound stimulus, this time in real word *down*. We join the activity as Sarah offers Louise a choice of two cars to put on the track. Louise indicates that she would like the green one. The first *D1. Demonstration for imitation* occurs in lines 005 and 007. This is a clean example and it successfully mobilises a production response from Louise in line 008. Sarah's sustained eye gaze beyond imitation (line 009) suggests she rejects Louise's production and she initiates a repair sequence with another *D1. Demonstration for imitation* in line 010. A second repair sequence takes place in lines 012-015, featuring another *D1. Demonstration for imitation*. The timing and articulatory nature of Louise's revised responses in lines 011 and 013 indicate that she has *a priori* knowledge of stimulus production. However, her linguolabial articulation for the /n/ phoneme rather than /d/ (line 011) and exaggerated mouth shape for the vowels suggests she is still refining this. The timing of Louise's production (overlapping Sarah's repeat production)

suggests it was elicited by Sarah’s verbal articulatory descriptor “*I’m using my tongue*” and/or articulatory gesture (*points to mouth*).

Sarah’s *D1. Demonstration for imitation* initiations took a variety of forms in Video 7, as shown by the six examples in Extracts 6, 7 and 8. The presence of this action in all three activities in Video 7 allowed some within-session analysis; see Table 8.6.

Table 8.6: Analysis of *D1. Demonstration for imitation* in Video 7

Therapist’s talk in <i>D1. Demonstration for imitation</i>	Turn design features
Activity 1	
1 Oh, another drum sou:nd <i>d̥ d̥ d̥ d̥ d̥</i> (CA; sustained gaze)	Verbal lead-in; elongation; stimulus; CA; sustained gaze
Activity 2	
2 And I (takes picture) can make up a little word (turns picture around) I can say (intake of breath) <i>d̥z:</i> (CA; sustained gaze)	Verbal lead-in; intake of breath; stimulus; CA; sustained gaze
3 (reaches for picture; intake of breath) so this one can say:: (sustained gaze) (2.0) <i>d̥ɔ:</i> (CA)	Verbal lead-in; intake of breath; elongation; silence stimulus; CA; sustained gaze;
Activity 3	
4 So can you make the car go↑ <i>d̥aʊn</i> (CA; sustained gaze)	Verbal lead-in; rising prosody; CA; sustained gaze
5 <i>d̥aʊn</i> (CA; sustained gaze)	Stimulus; CA; sustained gaze
6 <i>I’m using my tongue (points to mouth) d̥aʊn</i> (CA)	Verbal artic descriptor and gesture; stimulus; CA

Note. C=child. Text in brackets is nonverbal talk. Italics=pursuit utterance. CA=cued articulation finger gesture.

Sarah’s turns display a variety of response-mobilising features, such as, elongation in examples 1 and 3 and an intake of breath or rising prosody in examples 2, 3 and 4. Features common to all turns are: stimulus production (made salient with adaptive articulation); sustained child-directed gaze; and cued articulation finger gesture. Excluding the two pursuit examples (#5 and #6), Sarah’s utterances also include a verbal lead-in utterance prior to stimulus production. The analysis did not detect any distinct differences in how Sarah designed or delivered her initiations across the

session, other than variation due to the nature of the activity, e.g. the silence in #3 appeared to represent the time Sarah needed to generate a target sound sequence for Louise to imitate.

The prominent nature of *D1. Demonstration for imitation* in the 10-minute excerpt of Video 7 gave Louise the opportunity to engage in a high level of stimulus production. In the 25 examples that existed in this excerpt, 12 featured the stimulus at single sound level in a string of multiple productions, seven as sound sequences, and six as real words. Every one of Sarah's initiations successfully elicited a stimulus production from Louise. She made 85 stimulus productions in response to these initiations alone, 72 at single sound level (all but one as a string of multiple productions), seven in sound sequences and six in real words.

8.4.3 Initiation-Response-Evaluation (IRE) and repair sequences

The different profiles of therapist action that characterised Videos 1 and 7 projected different opportunities for Louise in terms of the nature of appropriate responses. In Video 1, Sarah showed a strong preference for the sub-category of action *A1. Demonstration for later*, which did not invite or engender a here and now response from Louise. Her other initiations were typically *B1. Demonstration for attention* or *C2. Verbal request for participation*, which invited nonverbal responses. These were often successful, as we saw in relation to *C2* initiations earlier in this chapter. Video 7 featured a prominence of the sub-category of action *D1. Demonstration for imitation*. Whilst this would inevitably create different response opportunities for Louise, the key finding from the analysis was that Louise's participation in therapy shaped the action and influenced the sequences of action that followed. Louise frequently overlapped Sarah's productions with her imitations of the stimulus and accompanying finger gestures. This gave rise to a different pattern of interaction, resembling the IRE structure that is characteristic of teacher-pupil interaction (Sinclair & Coulthard, 1975) and the repair sequences in Gardner's (1994) data of therapist-child interaction during phonology therapy. Within these sequences, the analysis revealed times when Louise produced target stimuli without a prior demonstration by Sarah and other times when she produced on imitation. This was considered with reference to Stackhouse and Wells' (1997) theory of speech processing. Analysis of sequences of therapy talk that featured child productions gave insight into the nature of Sarah's repair practices and the consequences these had for Louise. I will illustrate these findings with three examples of *D4. Verbal request for production*, which did not feature in Video 1. Four initiations of this kind occurred in Video 7 and

each one successfully elicited a production response from Louise. Extracts 9 and 10 illustrate three examples.

Extract 9 features the fishing game and drum pictures that symbolise the stimulus [d̥] in consonant and vowel sound sequences. The table surface displays toy fish with drum pictures underneath. The recurring steps in the activity are: 1) Louise catches a fish (to reveal a drum picture) → 2) Sarah (and Louise) produce the stimulus.

Extract 9: Fishing game with drum pictures (Sarah+Louise-V7-A2)

001 T What's the next one? ((gives C the fishing rod))
 002 C ((takes the rod; gaze to fish; [catches a fish])
 003 T Hmm ((gaze to fish)) [oh!
 004 C ((lifts up the fish on the rod; gaze to fish))
 005 T Oh you've got a blue one shall we put it in the water
 006 C ((detaches fish from rod; puts it in the 'water'))
 007 [((reaches to pick up the picture))
 008 T [Lovely and [this one we're gonna say↑
 009 C [((picks up picture; gaze camera; turns picture around
 010 to face camera))
 011 T ((turns gaze to camera))
 012 C ((turns gaze to T; maintains picture display for camera))
 013 T **d̥:** ((cued artic; gaze camera then to C; maintains cued artic))
 014 C **d̥:** ((makes cued artic gesture; gaze to T))
 015 T **d̥:** ((cued artic; sustained gaze to C))
 016 C ((lowers gaze and [reaches picture toward the 'water'))
 017 T [Look I'm using my tongue C ((takes picture
 018 from C))
 019 C [((turns gaze to T))
 020 T [((points to mouth)) watch [**d̥:** ((cued artic; sustained gaze))
 021 C [**d̥:** ((gaze to T))
 022 ((sustained gaze; makes cued artic gesture))
 023 T [Hmm brilliant you did too
 024 C [((lowers gaze; picks up fishing rod))
 (0.08)
 025 C [((reaches for the picture; holds it facing the camera))
 026 T [Oh this one let's say [:::
 027 (2.0)
 028 C [((turns gaze to T; holds picture))
 029 T Um: ((gaze to C; postures artic placement)) **d̥:** ((sustained gaze))
 030 C **d̥:** ((cued artic; gaze to camera))
 031 T ((sustained gaze to C))

032 C ((turns gaze to T))
 033 T **d̥a:** ((gaze to C))
 034 C **d̥a:** ((gaze to T; holds picture in outstretched arm))
 035 T Very good ((takes picture from C)) I saw your tongue
 036 C [((lowers gaze; picks up the rod))
 037 T [Good girl you're using your drum sound

In Extract 9, there are two consecutive sequences initiated by *D4. Verbal request for production*, the first in lines 008-023 and the second in lines 026-035. Both initiations (highlighted in grey) feature the explicit verb *say* and less explicit plural pronoun, *we* in the first and *us* in the second. In both sequences, Sarah produces the target stimulus as part of her utterance. In relation to Stackhouse and Wells' (1997) speech processing model, this means Louise can process output without accessing stored lexical representations. I will present the analysis of each sequence in turn.

We join Extract 9 as Louise, on direction from Sarah, catches a fish to reveal another drum picture. Her nonverbal behaviour in lines 002, 004 and 006 illustrate her engagement with the activity; she takes the rod, catches the fish and puts it in the (pretend) water. In lines 007 and 009-010, we observe Louise reach to pick up the drum picture and position it in a display for the camera. This is a specific enactment that we saw Sarah invite from Louise in Videos 1 and 7 with explicit *C2. Verbal request for participation* initiations. Here, Louise initiates the enactment. Louise turns her gaze to Sarah and displays patience, appearing to wait for Sarah to deliver her *D4. Verbal request for production* including stimulus production. Louise imitates Sarah's production (in line 014) and her subsequent nonverbal behaviour (in line 016) suggests she interpreted Sarah's repeat production (in line 015) as an acceptance. The author perceived Louise's production as an alveolar plosive [d]. This is not a match for Sarah's linguolabial stimulus but is an accurate realisation for phoneme /d/ to which the stimulus orientates. However, Sarah's subsequent talk in lines 017-018 and 020 is evidence that she does not accept Louise's production. She sustains her gaze and explicitly invites Louise's attention to specific articulatory features of the stimulus, e.g. "*look I'm using my tongue*" (line 017) and (*points to her mouth*) (line 020). Notice how Sarah refers to how she makes the sound rather than how Louise made it. As part of her repair initiation, Sarah makes another stimulus production in line 020. Sarah's turn across lines 017-018 and 020 has the design features of *B2. Verbal request for attention* with the additional feature of sustained gaze. In response, Louise resumes mutual eye gaze (line 019) and makes a revised stimulus production

(line 021), this time with a more visible linguolabial tongue placement. Interestingly, the timing of Louise’s production suggests it was the verbal directives for Louise’s attention and/or their specific nature that initiated the repair rather than Sarah’s stimulus production and sustained gaze. Sarah accepts Louise’s revised production with positive verbal praise and explicit feedback for using her tongue (see line 023), and the sequence draws to a close.

The second sequence in Extract 9 occurs in lines 026 to 037. Louise has caught a boot to reveal another picture. As in the preceding sequence, Louise initiates the nonverbal enactment of the stimulus routine, picking up the picture and displaying it for the camera (see lines 025 and 028) whilst also establishing mutual gaze with Sarah. It would seem that the act of retrieving the picture might now be part of the stimulus routine. Sarah’s *D4. Verbal request for production* occurs in lines 026-027 and 029. Louise again shows patience as she waits for Sarah to produce the target stimulus. Louise responds with a stimulus production with accompanying cued articulation gesture (line 030). Sarah’s repeated production, at closer proximity and with sustained eye gaze (line 033), elicits a second production from Louise; evidence that she treated it as a repair initiation. The author did not detect a perceptible phonetic difference between Sarah’s two stimulus productions, yet Louise’s second production (line 034) has a more visible interdental tongue placement than her first production, indicating repair work. This indicates that Louise has *a priori* knowledge about the target articulation for the sound stimulus and Sarah’s repair initiation in line 033 functions as a *D1. Demonstration for imitation*. In lines 035 and 037, we see Sarah accept Louise’s revised production with a specific evaluation “*very good... I saw your tongue*” (line 035) and “*good girl you’re using your drum sound*” (line 037). The nature of Sarah’s repair initiation and subsequent acceptance suggests the target production for Louise in this activity is a linguolabial production. Extract 10 features the same activity.

Extract 10: Fishing game with drum pictures (Sarah+Louise-V7-A2)

001 C ((drops fish in the 'water'; [picks up the drum picture; gaze to
 002 to picture then camera; holds picture facing camera))
 003 T [Let's do that one again
 004 Watch me ((reaches towards C's arm; gaze to C))
 005 C ((turns gaze to T; holds picture facing camera))
 006 T [dɪ: ((makes cued artic gesture in synchrony; sustained gaze to C))
 007 C [dɑ: ((makes part cued artic gesture; sustained gaze to T))
 008 T Oh you're doing [dɑ: ((makes cued artic; gaze to C; smiles))
 009 C [((nods; sustained gaze))

010 T I wanted it to say↑ ((takes picture from C)) (.) **di:** ((makes cued
 011 artic; sustained gaze to C))
 012 [((postures artic))
 013 C L((lowers gaze)) **gi:** ((lifts gaze to T))
 014 T Ooh ((points to mouth)) [with our tongue
 015 C L((postures linguolabial tongue placement))
 016 T [**di:** ((makes cued artic gesture; sustained gaze to C))
 017 C L **di:** ((gaze to T))
 018 T Good girl [you ((taps C on arm)) used your tongue with that sound
 019 C L((lowers gaze))

This extract features an interesting example of *D4. Verbal request for production*. It occurs in Sarah's turn in lines 003-004 and 006. Sarah uses an explicit verbal directive "let's do that one again" (line 003) featuring verb *do* and plural pronoun *us*, followed by a second verbal directive "watch me" (line 004). She supplements her invite for Louise's attention with a tap on her arm and eye gaze and goes on to produce the stimulus in line 006. When Louise overlaps her production (line 007), Sarah's somewhat perturbed response in lines 008 and 010-011 suggests her initiation was not complete, i.e. her stimulus production in line 006 was part of her *D4. Verbal request for production*. Another sequence unfolds in lines 010 to 013, initiated by Sarah's *D1. Demonstration for imitation* in line 009. Louise responds with a velar production (line 013); this is not a match for the linguolabial stimulus or the phoneme /d/ and Sarah's rejection with specific articulatory descriptor in line 014 is evidence that she perceived this. The nature of Louise's production suggests she may have processed it from stored representations, which are not yet fully accurate. We know from Extract 6, presented on page 177-178, that Louise had established an accurate motor program for [d] in the first activity of this session. In lines 014 and 016, we see Sarah provide further scaffolding, in the form of a verbal and nonverbal descriptor "ooh (points to mouth) with our tongue" (line 014) and a stimulus production (line 016) to help Louise access her accurate motor program for the target sound sequence. Louise's response emerges after Sarah's gesture (see line 014-015) as she postures the target tongue placement and then goes on to make a revised production in line 017, simultaneously with Sarah's production. The timing of Louise's response indicates her output was processed from stored representations and confirms that accurate representations were emerging in her speech processing system (Stackhouse & Wells, 1997). The finding that Sarah's gesture appeared to initiate Louise's response reflects her specific

knowledge of stimulus production and of the therapy process, which have accumulated over the course of the episode of care.

8.5 Conclusions

This chapter has presented the findings from detailed analysis of two x 10-minute video recordings from the first and last MSIVT sessions in the episode of care featuring Sarah and Louise. The analysis has revealed two very different profiles of therapist initiation, accomplished action and patterns of interaction at these two stages of therapy. Video 1 was characterised by a high level of minimally demanding *A1. Demonstration for later* and *B1. Demonstration for attention* initiations, featuring a high frequency of therapist stimulus productions. This gave Sarah the opportunity to provide Louise with intense exposure to salient stimuli, some for her attention in the session and some for her attention later (during home practice). Louise's mother was engaged in the session, providing verbal narrative and assisting Louise to participate. Video 7 was characterised by a high level of more demanding action, primarily accomplished with *D1. Demonstration for imitation* initiations. This gave Louise frequent opportunity to practice speech output with the continuing support of Sarah's stimulus productions. Implicit action outweighed explicit action at both stages of therapy. In both videos, Sarah's explicit invitations were typically *B2. Verbal request for attention* or *C2. Verbal request for participation*. The most demanding explicit invitation, *D4. Verbal request for production*, featured in Video 7 but not Video 1.

By tracking the design and delivery of initiations that were common to both videos, and the types of responses they engendered, the analysis has shown how both participants orientate to the multimodal stimulus routine, which becomes a familiar part of the interaction as therapy progresses. For example, Sarah transitioned from explicit 'hold the picture' utterances in Video 1 to more ambiguous 'show the TV/camera' utterances in Video 7. As an embedded part of the stimulus routine, such ambiguous utterances implicitly invited a stimulus production from Louise. Louise engaged nonverbally in activities and stimulus routine enactments in both videos. However, in Video 1, she often required encouragement or support from Sarah or her mother to participate. This reflects the early stage of therapy where Louise is developing her understanding about what therapy is and what her role is in it. In Video 7, Louise participated on request with relative ease and made a high level of stimulus production responses with accompanying cued articulation finger gestures. A key finding was that she also initiated the stimulus routine herself. Louise's

participation and initiations in the final MSIVT session are evidence of her familiarity with therapy and illustrate the empowering effect of the stimulus routine.

Analysis of the action *A1. Demonstration for later*, which features a therapist stimulus production with camera-directed gaze, revealed a sophisticated practice by Sarah. In Video 1, she delivered these initiations recurrently, one after another, but switched her gaze to Louise intermittently in-between consecutive turns, as though to maintain her attention in the here and now. This action featured far less frequently in Video 7; here Sarah used it intermittently on her turns in an activity. This video was characterised by a very different profile of action that gave rise to a high level of stimulus production by Louise. Sarah's productions to the camera had the appearance of helping to maintain a level of input stimulation for Louise at a future moment in time. Both patterns of display required Louise to navigate what was for her now and what was for her later. The finding that Louise initiated stimulus enactments with camera-directed gaze suggests the camera, as well as facilitating the delivery and accomplishment of certain types of action, has an empowering effect on facilitating a child's awareness and production of speech sounds.

The high level of production responses in Video 7 provided opportunities for repair work to explore and refine Louise's output and it changed the pattern of interaction. Sarah's repair initiations were typically more explicit as she facilitated articulatory revisions. They featured verbal descriptions of articulatory features and accompanying gestures, typically with a repeat stimulus production. The nature and timing of Louise's revised responses provided evidence that she had developed accurate internal representations for target stimuli and had acquired specific knowledge of stimulus production and the work that is to be achieved in therapy sessions.

Chapter 9: Discussion

9.1 Introduction

This thesis has examined the nature of multisensory input video-therapy (MSIVT) for young children with cleft palate-related speech sound disorder (CPSSD) as it was delivered as part of clinical practice within the NHS.

The study employed different methods to identify and describe characteristic features of MSIVT, broadly in terms of the structure of episodes of care and the nature of activities, speech sound targets and stimuli, and at a micro level in terms of the nature and pattern of therapist-child interaction. The findings reveal the distinctive nature of this innovative intervention. This chapter discusses the findings in relation to the existing literature on CPSSD, MSIVT and clinical and non-clinical interaction, and introduces some of the implications for clinical practice and future research that will be expanded on in Chapter 10.

9.2 Pilot Study

The pilot study aimed to examine the author's own clinical practice in order to formulate ideas for how to analyse the therapy. It gave insight into the multiple nature of target selection and the use of adaptive articulation in stimulus productions. The methods used to identify and describe these features were subsequently used in Phase 1. The pilot study also gave insight into the verbal and nonverbal nature of therapist behaviours and child responses and provided empirical evidence of two children making stimulus productions during therapy when production was not explicitly requested. The findings that multiple behaviours often existed within therapist turns, and turns served a variety of functions, informed the focus on the characteristics of interaction, rather than behaviours, in Phases 2 and 3.

9.3 Phase 1: Delivery of MSIVT in the NHS

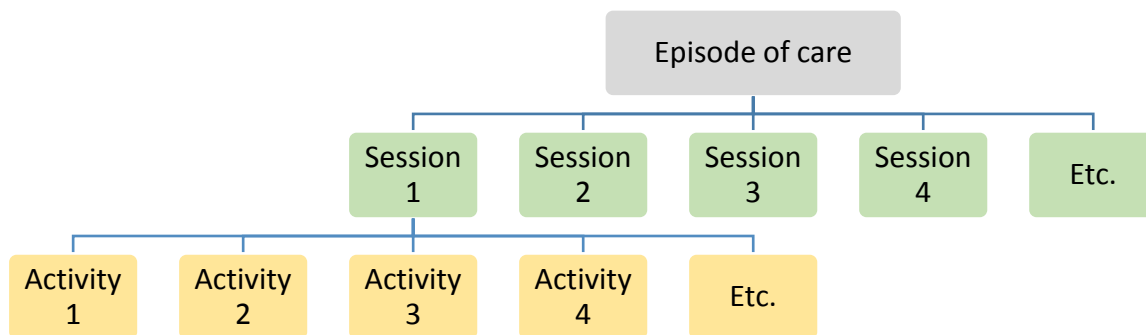
Phase 1 aimed to examine delivery of MSIVT as an episode of care in order to describe its structure and characteristic features relating to the nature of sessions, activities, speech sound targets and speech sound stimuli as well as any variation in delivery across therapist-child dyads. It found that episodes of care consisted of up to nine monthly and hospital-based therapy sessions. Sessions

involved approximately five play-based activities. Episodes of care featured multiple speech sound targets that were targeted at multiple linguistic levels. Therapists produced speech sound stimuli with adaptive articulations and made accompanying cued articulation finger gestures alongside productions. The analysis found variation in delivery in respect of session and activity duration and the sequence in which targets and stimuli were introduced. These findings are discussed below with reference to the existing literature on the MSIVT approach and, where appropriate, other speech interventions for children with and without CPSSD.

9.3.1 Structure of MSIVT episodes of care

Episodes of care were characterised by approximately nine monthly therapy sessions and lasted five to 11 months. The majority of sessions included video-therapy. These sessions lasted approximately 20 minutes and consisted of approximately five therapy activities each about four and a half minutes long. Figure 9.1 illustrates the relationship between these components.

Figure 9.1: Structure of MSIVT episodes of care



The number and frequency of sessions in the five episodes of care are different to those received by Sophie, one of the children studied by Calladine and Vance (2019), who received 18 fortnightly therapy sessions over a period of 18 months as part of her clinical care at a different specialist SLT service in the NHS. In Kaiser et al.'s (2017) study of the naturalistic approach enhanced milieu teaching with phonological emphasis (EMT+PE), children received 48 30-minute sessions bi-weekly for six months. This reflects the gap often seen between the dosage of therapy in efficacy studies and that seen in clinical practice (Dobinson & Wren, 2013; Wren, Harding, Goldbart & Roulstone, 2018). The optimum or minimum dosage of MSIVT required to have an impact on a child's speech processing skills is presently unknown.

Therapy took place in hospital settings and one or both parents were the main agent of change. Therapy in a clinical setting is routine practice in specialist SLT services in the NHS and it is common practice for parents to be the agents of change in speech intervention for preschool children (Enderby et al., 2009; Pamplona, Ysunza & Urióstegui, 1996; Scherer et al., 2008b; Treslove, 2014). Although the children studied by Calladine and Vance (2019) were seen in their home, therapy in that service is now provided from clinical settings.

There was variation in delivery in relation to the duration of sessions, which ranged from around nine minutes long (Helen-Thomas and Helen-Hannah) to over 25 minutes long (Sarah-Louise and Helen-Naomi). This raises the question of equity of care. However, the same SLT, Helen, delivered the episodes of care with the shortest sessions (Helen-Thomas and Helen-Hannah) and the longest sessions (Helen-Naomi), so this suggests the determining factor was not SLT preference or practice. Age of the child may have been a factor since Thomas and Hannah, who received the shortest sessions, were the youngest children in the study (1;6 at the onset of therapy), and Naomi, who received the longest sessions, was one of the two oldest children (2;3 at the onset of therapy). The findings from Phase 3 also raise possible interactional factors. Activities varied in duration from under one minute to over 12 minutes. Within every episode of care, activity duration varied with a range of more than five minutes between the shortest and longest activities. Varying activity duration was therefore a characteristic feature of the MSIVT episodes of care in this study, which cannot solely be explained by factors such as SLT or age of child. This has implications for managing parental expectations of therapy. The shorter activities typically involved the youngest children (Thomas, Hannah and Ellie) and the longer activities involved the two oldest children, Louise and Naomi, so age may have been a factor that influenced the duration of an activity.

9.3.2 Nature of MSIVT activities

The delivery of MSIVT was characterised by therapist-led, play-based and toy-based activities. A third (33%) of activities involved pictures to represent the speech sound stimuli, though there was variation across the episodes of care. Pictures featured prominently in the Sarah-Louise and Helen-Naomi episodes (in 37% and 68% of activities, respectively), but infrequently in Laura-Ellie's episode (9% of activities). This may reflect SLT or child preference rather than the age of the child, since picture-use was higher in the episodes involving the two youngest children, Thomas and Hannah (24% and 18%, respectively), than Laura-Ellie, which featured a different SLT. Paper-based activities were not a characteristic feature of MSIVT delivery in these episodes of care. These

findings reflect the young age of the children in this study and are consistent with Harding and Bryan (2000) and Calladine's (2009) descriptions of the MSIVT approach. The play-based nature of MSIVT activities is similar to that of EMT+PE (Kaiser et al., 2017; Scherer & Kaiser, 2010). However, Phase 2 of the study revealed a more structured manipulation of the toys than that which Kaiser and colleagues describe.

9.3.3 Nature of speech sound targets and stimuli

9.3.3.1 Target selection

The selection of multiple phoneme targets was a characteristic feature of MSIVT delivery in this study. The number of targets ranged from four (Sarah-Louise) to seven (Helen-Thomas and Helen-Hannah), though this appeared to be related to the age of the child and the nature of their speech characteristics. The type of phonemes targeted reflects the underlying principles of target selection for this client group (Harding & Grunwell, 1998; Harding-Bell & Howard, 2011). For example, in the Sarah-Louise and Laura-Ellie episodes, the SLTs identified explicit phoneme targets at the outset of therapy that were based on the diagnosis of specific 'active' cleft speech characteristics. Such characteristics represent impairment at the level of motor program (Calladine & Vance, 2019). Consistent with well-established principles of target selection, SLTs Sarah and Laura targeted all affected phonemes during the episode of care (Harding-Bell & Howard, 2011; Russell & Albery, 2005). They introduced targets systematically as the episodes progressed, though they used different strategies, e.g. Sarah favoured one target per activity, whereas Laura favoured one to three. The children Thomas, Hannah and Naomi presented with absent pressure consonants and more restricted speech sound inventories than Louise and Ellie. Such characteristics may be due to impairment at the level of motor program, but may also reflect impairment in motor execution (Calladine & Vance, 2019). Consistent with Naomi's presentation of suspected velopharyngeal dysfunction, SLT Helen selected voiceless targets only. This is a well-established principle of target selection when there are structural restrictions on sound production (Harding & Grunwell, 1998). Helen introduced targets systematically in the Helen-Naomi episode, but with a different strategy to Sarah and Laura, i.e. she typically introduced all targets in the first activity of a session and then focused on one target in each of the remaining activities. Helen used a different target selection approach in the Helen-Thomas and Helen-Hannah episodes. She did not identify specific target phonemes at the outset of therapy. Instead, she appeared to introduce a range of sounds as potential targets to inform selection rather than

the ‘selection and focus’ approach seen in the Helen-Naomi episode. Some of the variation in delivery seen in this study in respect of target selection is likely to reflect the limitations in existing descriptions of MSIVT and the lack of experimental research in this area.

9.3.3.2 Stimulus design

The use of adaptive articulations in stimulus productions was a characteristic feature of MSIVT delivery in this study, and this aspect of stimulus design aligns with the nature of CPSSD and principles of the MSIVT approach (Calladine & Vance, 2019; Harding & Bryan, 2000). Adaptive articulation is used to increase the saliency of speech sounds to maximise the quality of input processing, and to help the child perceive stimuli as novel so that speech processing occurs without access to stored motor programs and phonological representations (Calladine & Vance, 2019; Harding & Bryan, 2000). Across the episodes, SLTs used different stimuli for the same target phoneme, e.g. [t̥ t̄ t̃ t̂ t̃] for /t/. This may reflect flexibility by the SLT to find the ‘right’ type of stimulus. It also resembles the variability in phone production that Gardner (1994) observed in phonology therapy sessions. Gardner (1994) suggests this can present a challenge for the child, as they have to decipher which sound is the actual target. However, the use of such subtle variations in speech sound therapy is supported by exemplar theory (Cole, 2009; Foulkes & Docherty, 2006). An interesting finding in this study was that the reverse also occurred, i.e. several different phonemes represented by one stimulus. However, this was only evident in one of the episodes (Laura-Ellie) and was not a characteristic feature of MSIVT delivery. In this episode, an interdental fricative stimulus was used for the phonemes /s z θ ð/. This may have been favourable for the child in terms of speech processing. If the stimulus did not orientate singularly to any one phoneme, it may have helped to give it novel status and facilitate new processing. Introducing phonetic variations of the stimulus to facilitate updating of existing motor programs and phonological representations associated with all four phonemes would be required at a later stage. The enactment of cued articulation finger gestures (Passy, 1993) was also a characteristic feature of MSIVT delivery. This aligns with existing descriptions of MSIVT and the multisensory principles of the approach.

9.3.3.3 Linguistic level

The MSIVT sessions in this study were characterised by multiple different levels of stimulus production. This aligns with current descriptions of the MSIVT approach (Calladine & Vance, 2019;

Harding & Bryan, 2000; Treslove, 2014). Different levels also featured within a single activity, although it was more common for SLTs to focus on one level only in a single activity. In MSIVT, speech sound stimuli represent novel sounds that orientate to target phonemes. As therapy progresses, the goal is to support the child in integrating new motor programs into lexical processing so they are available for retrieval during everyday speech production (Calladine & Vance, 2019; Harding & Bryan, 2000; Treslove, 2014). The linguistic context in which a stimulus is produced therefore has implications for how it is processed by the child at different stages in therapy. The study found that the specific levels at which SLTs produced stimuli, and the sequence of progression through them, did not always align with the principles of MSIVT described by Harding and Bryan (2000) and Calladine and Vance (2019). Single sound and real word levels were the most common. Contrary to existing descriptions (Calladine, 2009; 2010; Calladine and Vance, 2019; Treslove, 2014), sound sequence level only featured in three of the episodes (Sarah-Louise, Laura-Ellie and Helen-Naomi) and in only 15/119 (13%) of the activities, and invented word level only featured in 1/119 (0.8%) activity. Therapists Sarah and Laura showed a preference for single sound level, whereas Helen displayed a preference for real words in all three of the episodes she provided. The children in these episodes, Thomas, Hannah and Naomi, all presented with absent pressure consonants and therapy was described as 'diagnostic'. These differences in relation to level of stimulus production may reflect the individualised nature of therapy tailored to the child's unique profile of speech processing characteristics, but may also reflect differences in SLT preference informed by previous knowledge and experience. All three SLTs displayed systematic progression through the levels for different stimuli, e.g. single sounds first then on to real words, but this did not happen consistently, and in the Helen-Thomas and Helen-Hannah episodes, stimuli were often produced in real words first. These findings have implications for the child in terms of speech processing. If real words known to the child are introduced before single sound stimuli, Stackhouse and Wells' (1997) psycholinguistic theory of speech processing suggests lexical processing is likely to occur. Although this study did not assess the impact of intervention, differences in the order in which therapists introduced targets at different levels, suggests MSIVT may have had different effects on the children's speech processing systems. This has implications for SLTs in terms of training, and for future research.

9.3.4 Child stimulus productions

Child stimulus productions occurred in all five MSIVT episodes of care and in approximately half (52%) of all activities, though the children varied in how many sessions and activities they made productions. Louise, Ellie and Naomi, who demonstrated the highest rates of production, were the oldest children in the study. Thomas and Hannah, who demonstrated the lowest rates of production, were the youngest children. This suggests the child's age or stage of development may have been a factor in their level of output in the MSIVT sessions. The finding that children made stimulus productions during MSIVT is consistent with Calladine (2009), who observed stimulus productions in all four two-year-old children in her study, and Calladine and Vance (2019). The present study did not examine the exact number of child stimulus productions, which would provide a more precise measure of rate of production. The findings from Phase 3 raise the possibility that interactional factors might influence the rate of production and this may provide a within-session outcome of therapy. Children produced stimuli at different levels, though examination of the data suggests this reflected the different levels of stimulus production they were exposed to by the therapist.

9.4 Phase 2: Interactional features of MSIVT

Phase 2 aimed to analyse therapist-child interaction during MSIVT sessions in order to identify and describe the ways SLTs establish a child's attention and stimulate their awareness and production of speech sounds. The findings reveal the structured nature of MSIVT sessions and the activities within them. Therapists led the sessions but gave the child opportunity to make choices that influenced the content of activities and controlled the pace of activities. Activities were characterised by recurrent sequences in which the therapist made salient productions of speech sound stimuli alongside pictures, finger gestures and toys, referred to as 'stimulus routine sequences'. One of the key findings from the analysis is the opportunities these stimulus presentations gave the child to be an active participant in therapy.

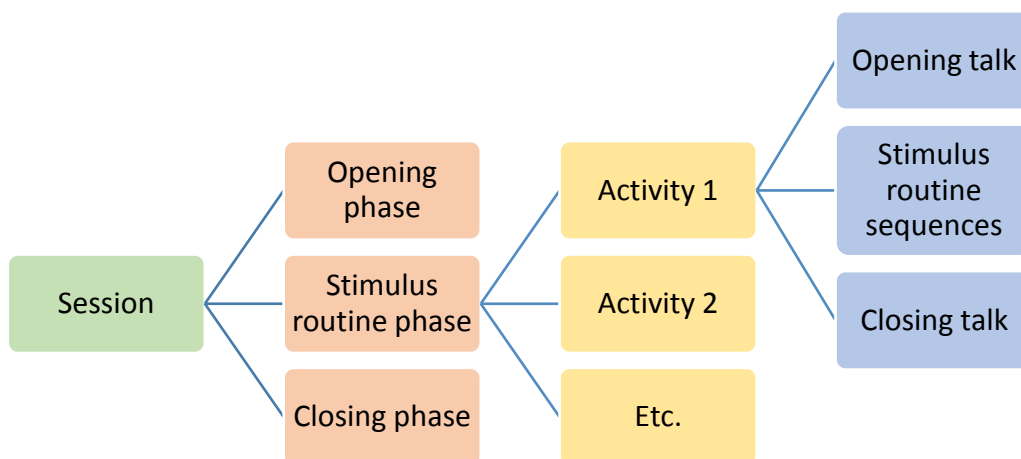
The analysis was particularly interested in the actions that therapist turns fulfilled, the practices therapists used to accomplish them, and the consequences they had for the child. It found that therapists used a variety of different types of action and these elicited range of responses from the child. Therapists did not always invite the child to respond in the here and now; they used the camera as a gateway to providing sound stimulation to the child at a future moment in time.

When a response was required, the therapist invited the child to watch and listen to salient speech sound stimuli, manipulate objects and pictures associated with the stimulus, and produce target stimuli. A characteristic feature of these invitations by the therapist was their implicit nature. Whilst explicit forms did exist, therapists showed a preference for softening them with plural pronouns and ambiguous verbs. A strong feature of the MSIVT interaction in this study was that therapists did not always pursue a response from the child when one was not forthcoming. This served to establish a non-pressurising space for the child to learn and engage in a way that was conducive to their young age.

9.4.1 Structure of MSIVT sessions and activities

The organisational structure of MSIVT sessions and activities was similar to the structure of other therapist-led interventions. Therapy sessions typically consisted of three phases: opening; stimulus routine; and closing. The ‘stimulus routine phase’ is where therapy activities took place. Activities were characterised by recurrent sequences of interaction in which the therapist demonstrated salient speech sound stimuli and invited different types of responses from the child; these are referred to as ‘stimulus routine sequences’. Figure 9.2 illustrates the relationship between these components.

Figure 9.2: Structure of MSIVT sessions and activities



The stimulus routine phase resembles the ‘remedial phase’ and ‘task phase’ in the studies by Panagos et al. (1986) and Tykkyläinen (2009), respectively. It has parallels with the phase Reed et al. (2013) call the ‘pursuit of learnables’ in vocal master classes. The stimulus routine phase was characterised by stimulus routine sequences. These were enacted in a recurrent manner, which

resembles the ‘remedial sequences’, ‘task sequences’ and ‘object introducing sequences’ in the studies by Panagos et al. (1986), Tykkyläinen (2009), and Ronkainen (2011), respectively.

Kovarsky and Duchan (1997) show that interactional features of adult-led and child-led interventions co-existing within an intervention. In the MSIVT activities, all four therapists engaged in a practice whereby they offered the child a choice, for example, a choice of game or toy that would become the basis of the activity, or a choice of toys or pictures within an activity. This is a feature of child-led intervention (Kaiser et al., 2017; Kovarsky & Duchan, 1997). In the present study, MSIVT activities were therapist-led, but this choice-offering practice created opportunities for the child to influence the content of the session. This practice aligns with existing descriptions of MSIVT, which describe activities as more therapist-led than naturalistic approaches, but that the child is actively involved (Calladine & Vance, 2019).

9.4.2 Therapist and child roles

The findings illustrate the therapist as the lead and organiser of the therapy session and the principal initiator of interaction. This is not surprising given that the aims of therapy are to provide the child with input stimulation and opportunities for output practice that will facilitate change in their speech processing system. However, therapists did not create these learning opportunities in the directive manner that Prutting et al. (1978) and Ripich et al. (1984) describe in their analyses of speech and language therapy sessions. After all, the original authors of multisensory input therapy do advocate a non-directive and nurturant style of interaction, and this does not infer an active therapist and passive child dynamic (Harding & Bryan, 2000; Pamplona et al., 2001). In the present study, whilst the therapists took the lead, within the activities they gave the child opportunities to make choices, initiate stimulus routine sequences, and engage in periods of child-led play. Prutting et al. (1978) used quantitative analysis to show that therapists in language intervention sessions primarily request responses and children primarily respond. In the present study, although therapists were the principal initiator of interaction, children did not always respond in a manner that was consistent with the design of the therapist’s turn, and the therapist did not always pursue the child for a response. Even when the therapist explicitly invited a response from the child, she did not always pursue the child for a response when it was absent. This feature of the MSIVT interaction is in contrast to Panagos et al.’s (1986) account of therapy discourse, in which they describe “when a response is inappropriate, or not forthcoming... the clinician presses for the desired response, making whatever additional requests are necessary” (p221). In the MSIVT

interaction, the lack of systematic pursuit by the therapists undermined the child as a respondent. This appeared to give the child authority to influence the interaction and the opportunity to define the parameters of their own role identity. This is more consistent with the child's role in child-led intervention (Fey, 1986; Kovarsky & Duchan, 1997). This practice was a characteristic feature of the MSIVT interaction examined in this study and may be a distinctive feature of the approach.

Reed et al. (2013), in their study of vocal master classes, give us insight into the decision dilemma that student singers and accompanists face when negotiating when to respond to their master's instruction. They distinguished between local directives, which engendered an immediate performance response by the students, and restart-relevant directives, which engendered the students' attention to a series or *cluster* of directives followed by a performance response. Local directives gave students the opportunity to respond because the master paused to create a slot for the students, whereas in *clustered* restart-relevant directives the master gave no opportunity for a response until after he had delivered the final one. In the present study, the inconsistent nature of therapists' pursuit practice, where sometimes they would pursue the child for a response and other times they would not, creates a similar dilemma for the child. The child had to negotiate when to and when not to respond. The analysis suggests therapists create this ambiguity somewhat purposively as though to give the child opportunity to actively participate and influence what will happen next.

The collaborative nature of practice by the singers and accompanists in Reed et al.'s (2013) study is similar to what Leahy (2004) observed between the therapist and child in stuttering therapy sessions. One of the features of interaction that Leahy (2004) identified in her study was the therapist's use of the pronoun *we*, which she suggests signifies inclusivity. Plural pronouns featured prominently in the MSIVT interaction. Therapists used them in all macro categories of action: *A. Demonstration to the camera; B. Invitation to attend to the stimulus; C. Invitation to participate in the stimulus routine; and D. Invitation to produce the stimulus.* The pronoun *we* was used (for 'you') in turns that elicited a child response and also (for 'we') in turns that elicited a collaborative response. This range of meanings for the same word creates ambiguity for the child, who has to correctly infer which meaning the therapist intended in order to respond accordingly. However, when the therapist uses *we* for 'you' in an explicit invitation for a response, it gives the child freedom not to respond. The analysis suggests therapists may use plural pronouns in MSIVT

as a way of minimising pressure on the child to respond and as a way of presenting therapy as an inclusive process: ‘we are doing this together’.

9.4.3 Structure and pattern of therapist-child interaction in MSIVT

The analysis found MSIVT activities were characterised by recurrent stimulus routine sequences. They featured some of the ‘boundary markers’ described by Letts (1985), which serve to regulate and maintain the structure of therapy sessions, for example, “*right pop it on top*” (Extract 9; Line 007), “*okay put him on here to dry*” (Extract 1; Line 032), and “*so what colour are you gonna do then?*” (Extract 16; Line 001). Therapists’ use of *oh* and *ah* also appeared to function as boundary markers, for example, “*oh let’s see first picture I’ve found to post*” (Extract 13; Line 001) and “*ah (child’s name) shall we wash his tail*” (Extract 1; Line 017). Letts (1985) also describes the use of ‘attention-getters’, such as ‘ready’ and ‘listen’. These words featured prominently in the MSIVT interaction, particularly in *B2. Verbal request for attention*.

The analysis revealed an unusual pattern of interaction due in part to the unsystematic nature of therapists’ pursuit practice. Some sequences of therapy talk consisted of therapist turns only with no requirement for the child to make a here and now response. This pattern of interaction does not resemble the request-response-evaluation (RRE) sequence described in previous studies of therapy discourse, such as Prutting et al. (1978) and Ripich et al. (1984). Similarly, whilst the MSIVT interaction appeared to be asymmetrical, with the therapist consuming more turns than the child, it was not to the degree that Prutting et al. (1978) and Ripich et al. (1984) describe. Hulterstam and Nettelbladt (2002) compared patterns of interaction in two different types of speech intervention, Metaphon therapy and traditional therapy. They found a less asymmetrical pattern of interaction in Metaphon therapy sessions, which reflected specific features of the Metaphon approach designed to enable the child to be a more active participant. One such feature, the use of a metalanguage to discuss and describe speech sounds, was observed recurrently in the MSIVT interaction. It is also described by Calladine and Vance (2019) in their case study of two-year-old James. They show him producing metalinguistic utterances previously produced by the therapist, e.g. “hold our nose” and “with our lips” (p347). James produced these whilst enacting nonverbal elements of the stimulus routine, such as holding up the stimulus picture for the camera. As in Metaphon therapy, this use of metalinguistic utterances appears to be a characteristic feature of MSIVT.

9.4.4 The stimulus routine: Layers and multimodality

The stimulus routine consisted of a stimulus production and/or enactment of stimulus-related behaviours by the therapist. Some features of the stimulus routine are common to the ‘speech teaching strategies’ McCartney (1989) describes. In the present study, therapists produced stimuli with saliency features, such as adaptive articulation, and described articulatory features of sounds (‘articulatory description’). Adaptive articulation resembles McCartney’s (1989) ‘augmented model’, although in her data this was most prominent in repair sequences. Articulatory description resembles McCartney’s (1989) ‘metalinguistic utterance’ and the specific feedback therapists gave to initiate repair in Gardner’s (1994) study. Adaptive articulation was a prominent feature of therapists’ stimulus productions in the MSIVT interaction, as was repetition of the stimulus in a string of multiple productions. These features reflect the underlying theoretical basis of MSIVT, such as imitation theory (Meltzoff, 1988) and mirror neuron theory (Oberman et al., 2007).

The analysis shows the stimulus routine was characteristically multimodal in nature, featuring a myriad of verbal, nonverbal and prosodic behaviours. It bears resemblance to the multimodality described by Ronkainen (2011) and Tykkyläinen (2009). Therapists in the present study produced stimuli alongside pictures and toys, which they manipulated in various ways, e.g. displayed the picture for the camera, caught a bubble on a wand and made it wobble, wiggled a toy snake, and made accompanying nonverbal finger gestures, such as cued articulation (Passy, 1993). Such enactments create associations that transform the pictures, toys and nonverbal manipulations into symbolic representations for the sound stimulus. Therapists referred to stimuli with verbal labels, typically produced with distinct intonation, and introduced them with repetitive phrases, e.g. ‘the (sound name) goes’. These behaviours add further layers of structure and multimodality to the stimulus routine. This practice by the therapists aligns with existing descriptions of the MSIVT approach, which advocate the use of visual representations to intensify the stimulus, reduce distractibility and stimulate visual processing (Calladine & Vance, 2019; Harding & Bryan, 2000). They also reflect theories underpinning the MSIVT approach, such as multimodal speech acquisition (Dodd, 1979).

9.4.5 Nature of therapist-child interaction in MSIVT

9.4.5.1 Multimodality in therapist initiations

The analysis revealed that when therapists initiated a new sequence of therapy talk with the child, their initiations were characteristically multimodal in nature. They displayed a variety of verbal, nonverbal and prosodic features, showing similarities with the interactional features of other therapy approaches. Nonverbal features include body leaning, eye gaze, manipulation of picture and toys, nodding in synchrony with stimulus production, pointing and finger gestures. Such behaviours resemble those described by Bobkoff and Panagos (1986), Gardner (1994), Panagos et al. (1988) and Ronkainen (2011). Prosodic features include prolonged vowels, rising pitch, pausing and intakes of breath before stimulus production, and emphasis on target stimuli, which resemble those described by Gardner (1994), Ronkainen (2011) and Tykkyläinen (2009). Verbal features, such as articulatory description and labels for sounds, are similar to those that Gardner (1994), Hulterstam and Nettelbladt (2002) and McCartney (1989) describe in their studies of phonology therapy. Whilst the multimodal nature of therapist initiations does not appear to distinguish MSIVT interaction from other approaches, the analysis suggests the intensity of these features was distinctive, albeit with some similarities to auditory-verbal therapy (Ronkainen, 2011). These findings explicate the input-based nature of MSIVT and reflect the common aim of MSIVT and auditory-verbal therapy to establish the child's attention in order to provide intense exposure to stimuli that will help to facilitate change in speech processing.

9.4.5.2 Nature of therapist action: Design features and consequences

The analysis found that the action projected by therapist initiations was characteristically diverse in nature. Four distinct categories were identified: *A. Demonstration to the camera*; *B. Invitation to attend to the stimulus*; *C. Invitation to participate in the stimulus routine*; and *D. Invitation to produce the stimulus*. Within these, therapist initiations fulfilled nine sub-categories of action that were distinguishable by the way they were designed; see Table 7.1 (Chapter 7, page 125). The findings suggest the specific nature of action that therapists project in MSIVT is a distinctive and characteristic feature of the approach. I demonstrate this below in my discussion of some specific sub-categories of action that emerged from the analysis.

A1. Demonstration for later is an unusual type of action seen in the data, which has not previously been reported as an interactional feature of speech sound intervention. However, it was not an

unexpected finding of the present study since it reflects Harding and Bryan (2000) and Calladine and Vance's (2019) descriptions of the video-therapy component of MSIVT. This action was characterised by a stimulus production and camera-directed gaze. It did not typically receive a response from the child, which is unsurprising given the direction of the therapist's gaze. Both Gardner (1994) and Stivers and Rossano (2010) demonstrate that child or recipient responses are strongly associated with gaze in their direction. In the MSIVT sessions, the child was free to gaze to the therapist or camera during the demonstration, but the therapist did not pursue the child's visual attention when it was absent. Calladine and Vance (2019) describe how video-therapy is a tool to support home practice by ensuring children have access to "consistent and accurate models" (p341) of speech sound stimuli and high doses of input stimulation. The absence of pursuit seen in the present study aligns with the underlying notion that the demonstration to the camera is for the child to respond to later, not in the here and now. This is similar to the clustered restart-relevant directives that Reed et al. (2013) observed in their analysis of interaction between masters and students in vocal master classes. These restart-relevant directives also engendered a delayed response. However, the spatial context is different. In MSIVT, *A1. Demonstration for later* projects a child response at a future moment in time *beyond* the 'live' therapy session, whereas the restart-relevant directions in Reed et al.'s (2013) study projected a response from students at a future moment of time *within* the class.

The analysis showed therapists switching between *A1. Demonstration for later* and the other sub-categories of action that invited a here and now response from the child within the session. This had the appearance of a television or radio broadcaster switching his/her projection back and forth between a broadcast and a 'live' audience (Goffman, 1981). In the MSIVT sessions, the child becomes a bystander participant in the 'live' frame while the therapist shifts footing to talk to the child in the 'future' frame. On the whole, this appeared to happen with relative ease, although the analysis did reveal some examples where the therapist broke her turn mid-demonstration (e.g. Extract 1 in Chapter 7; Line 004), suggesting this is a skill that requires practice. Although it was not the focus of the analysis in the present study, switching from doing therapy with the child to explaining home practice to a parent would be another example of switching footing. Adding further complexity to the interaction, the analysis also showed therapists recruiting the child in the live frame to interact with the child in the future frame. This is seen particularly in *C2. Verbal request for participation*, for example, *T: "are you going to show the telly for me (turns gaze; points to the camera)* (Extract 10 in Chapter 7; Lines 013-014). The therapists identified the camera

as 'camera', 'telly' or 'TV'. In doing so they projected to the child that the camera was the participant, which would appear to be a more tangible concept for such a young child than the alternative projection that the camera is a gateway to the child in the future.

Goffman (1981) claims: "...an utterance does not carve up the world beyond the speaker into precisely two parts, recipients and non-recipient, but rather opens up an array of structurally different possibilities, establishing the participation framework in which the speaker will be guiding his delivery" (p137). It is evident from the analysis that SLTs in MSIVT not only find themselves navigating different types of action to elicit different types of responses, they also have to navigate two different frames of interaction and a complex participation framework. Of course, it is the child as well as the therapist who needs to navigate this complex interaction.

The findings show that, in contrast to *A1. Demonstration for later*, the other sub-categories of action invited a range of within-session responses from the child: their visual and/or auditory attention to the stimulus (macro category B); their nonverbal participation in the stimulus routine (macro category C); and their production of the stimulus, either as an imitated or non-imitated production (macro category D). The nature of these actions and the practices therapists used bear resemblance to those reported in previous studies of therapy interaction involving children. For example, *B1. Demonstration for attention* is similar in nature and appearance to the attention-seeking turns therapists displayed in Ronkainen's (2011) study of auditory-verbal therapy. Some of its characteristic features include child-directed gaze, verbal label and introductory utterance, e.g. 'the bubble goes...', and with-holding a toy or picture close to the face or mouth. These features told the child something of interest was coming, helped to create anticipation, and served to elicit and maintain the child's attention. *D1. Demonstration for imitation* resembles the 'model' that features in Gardner (1994) and McCartney's (1989) studies of phonology therapy and the 'signals' that Letts (1985) describes. *D2. Incomplete demonstration* resembles what Letts (1985) refers to as a 'prompt', a type of signal that elicits completion. Letts (1985) describes how children know how to respond to signals because the therapist explains this in a prior orientation act. However, she states "more rarely, the therapist may link a stimulus with a response in an arbitrary way, explaining to the child that he should respond in that way every time she gives that stimulus" (p327), i.e. the stimulus becomes a signal for the duration of that activity. The present study did not find explicit verbal explanations by the therapist to be a characteristic feature of MSIVT interaction. However, it did find symbolic linkage in the ways therapists constructed the stimulus

routine, for example, producing a sound stimulus as they perform a specific act with a toy, making a cued articulation finger gesture alongside stimulus production, etc.

In her study of interaction during phonology therapy, Gardner (1994) found that subtle prosodic and syntactic features distinguished between a stimulus production that served as a model to elicit a child production and one that served as a redoing to accept a child production. She also found child-directed gaze was a strong feature that engendered a child target production. The present analysis uncovered similar subtleties, particularly in relation to demonstrations. Specific combinations of features represented different types of demonstrations. One of the distinguishing features was the nature of the therapist's gaze: *A1. Demonstration for later* was characterised by stimulus production and camera-directed gaze; *B1. Demonstration for attention* was characterised by stimulus production and child-directed gaze; and *D1. Demonstration for imitation* was characterised by stimulus production followed by sustained child-directed gaze (and silence).

Implicit and explicit forms of action

The analysis found that therapists displayed a preference for implicit forms of action. This was not a surprising finding given the non-directive principle of the MSIVT approach (Calladine & Vance, 2019). Implicit forms of action appear particularly suited to this approach given the young age of the children. However, explicit invitations also occurred, and this was a surprising finding. All three categories of invitation featured explicit forms: *B2. Verbal request for attention*, e.g. “got some more, look” (Extract 5; Lines 006-007); *C2. Verbal request for participation*, e.g. “are you going to show the telly for me?” (Extract 10; Line 013-014); and *D4. Verbal request for production*, e.g. “can we hear your sound then?” (Extract 16; Line 008). *B2. Verbal request for attention* and *C2. Verbal request for participation*, which both invited nonverbal responses, featured prominently, whereas *D4. Verbal request for production* occurred infrequently. This is similar to Gardner's (1994) finding that therapists in her study of phonology therapy showed a dispreference for overt requests for target word production. It also supports existing descriptions of the MSIVT approach (Calladine & Vance, 2019; Harding & Bryan, 2000). Rather than explicitly requesting production, therapists invited production from the child with modified and incomplete stimulus demonstrations.

Detailed analysis of the behaviours therapists displayed in their enactments of the stimulus routine and the symbolic associations they create provide evidence to suggest that invitations for participation implicitly invited production from the child. The layered composition and multimodal nature of the stimulus routine presented the child with a sequence of behaviours associated with

stimulus production. By inviting the child to enact one or more of these behaviours, the therapist implicitly invited the child to produce the stimulus. In their analysis of interaction during a child-led intervention session, Kovarsky and Duchan (1997) found the therapist orientated to a goal of eliciting initiations from the child yet displayed a 'hidden agenda' of eliciting initiations that were free from interactional struggle. In the MSIVT sessions, therapists orientated to the overriding goal of influencing change in the child's output by providing input stimulation. However, the analysis suggests a hidden agenda of eliciting output via implicit means. Facilitating the child's attention to the stimulus and their participation in the stimulus routine (via explicit invitation) and creating opportunities for stimulus production (via implicit invitation) are short-term goals driving therapy.

A characteristic feature of therapist initiations in the MSIVT interaction was the use of the plural pronoun 'we' in place of 'you' when making explicit requests for the child's participation, e.g. "*shall we hold it up to the camera (gaze to the child; holds picture facing child)*" (Extract 9; Lines 014-015). In this example, the therapist's nonverbal behaviours suggested 'we' meant 'you' and this is how it was interpreted by the child. The use of 'we' appeared to soften the explicitness of the therapist's request. We also saw therapists use plural pronouns when making explicit requests for the child's production, e.g. "*...what do we have to say to daddy to make sure he doesn't wake up (sustained gaze to child)*" (Extract 17; Lines 010-011). The therapist's use of 'we' appeared to signify collaboration, similar to its use in stuttering therapy (Leahy, 2004). Use of 'we' was not limited to such explicit requests; it also featured in therapists' verbal commentaries when they described their demonstrations of the stimulus routine. This creates ambiguity for the child, who needs to interpret the therapist's intention in order to respond appropriately. Similar things were seen in therapists' verb use. The verbs that therapists used in *B2. Verbal request for attention* and *D4. Verbal request for production* were typically unambiguous, e.g. 'watch' and 'say', respectively. Yet those that featured in *C2. Verbal request for participation* were characteristically ambiguous, particularly when they invited the child to interact with the camera, e.g. inviting the child to *show* or *hold* stimulus pictures 'for the telly/TV'. This requires the child to decide whether to respond to the literal meaning or, since the picture display forms part of the stimulus routine, interpret this as an implicit invitation for stimulus production. It seems plausible to suggest these linguistic preferences exhibited by the therapists are practices that reflect two specific principles of the MSIVT approach: non-directive style of interaction (based on Pamplona et al., 2001); and active involvement of the child.

The hierarchical organisation of the categories of action presented in Table 7.1 (Chapter 7, page 125) has similarities with Stivers and Rossano's (2010) scalar model of response relevance. This model depicts that the relevance of a recipient response (response mobilisation) is influenced by both the type of action (canonical or non-canonical) and the way it is delivered (turn design features). In the MSIVT interaction, different categories of action placed different levels of demand on the child for a specific type of response. *A1. Demonstration for later* is low-level because it did not invite a here and now response. Sub-categories of *B. Invitation to attend to the stimulus* and *C. Invitation to participate in the stimulus routine* are mid-level because they invited nonverbal responses. Sub-categories of *D. Invitation to produce the stimulus* are highly demanding because they invited a specific verbal response. When therapists made explicit requests for the child's attention, participation or production with design features such as sustained child-directed gaze, they made it particularly relevant for the child to respond.

9.4.5.3 Nature of child responses

The nature and range of child responses that occurred and were accepted during therapy was a characteristic feature of the MSIVT interaction. One particular action, *A1. Demonstration for later*, did not typically engender a here and now response from the child. In other studies of therapy interaction, an absent response is regarded as a problematic source of therapy talk that leads on to the therapist pursuing the child for a response (Gardner, 1994; Prutting et al., 1978). In the MSIVT interaction, even when the therapist invited a here and now response from the child, she did not consistently treat an absent or ill-fitting response as problematic.

Phase 1 found that child productions of target stimuli was a characteristic feature of MSIVT in this study; productions occurred in more than half (52%) of all activities across the five episodes of care. The detailed analysis conducted in Phase 2 revealed the interactional context in which child productions occurred in four therapist-child dyads. It found that they did not occur exclusively in response to *D. Invitation to produce the stimulus. B2. Verbal request for attention* and *C2. Verbal request for participation* also elicited stimulus productions from the child, e.g. *T: "(child's name), listen first"; C: "f:."* (Extract 3; Lines 028-029), and *T: "show the telly first"; C: "m m m m"* (Extract 10; Lines 019-021), respectively. These findings suggest that explicit requests for attention and participation also functioned as implicit requests for production. Calladine and Vance (2019) make a similar observation in their case study of two-year-old James: "...when he imitates some of the clinician's actions, such as holding up pictures and holding his nose, he imitates parts of her

commentary and does imitate the target sounds” (p345). In the present study, when making productions, the children typically produced nonverbal behaviours alongside their production. This included finger gestures that therapists had previously made alongside their demonstration of the stimulus, thus embedding it as an element of the stimulus routine. It seems reasonable to suggest that the layered composition of verbal and nonverbal elements in the stimulus routine, which in the present study appeared to work well to facilitate the children’s engagement in therapy, might influence the way a child processes speech sound stimuli during MSIVT. The multimodal nature of stimulus routines makes the therapist’s stimulus production just one of its composite elements. Whilst MSIVT aims to stimulate phonetic processing of the sound stimulus, might it instead stimulate processing of the verbal (and nonverbal) elements of the stimulus routine as a whole (A. Harding-Bell, personal communication)? This would draw some similarity with the whole-word nature of phonological processing in young children (Sutherland & Gillon, 2007). It would explain why the children in the present study sometimes enacted nonverbal elements of the stimulus routine when they had been invited to produce the stimulus, and why they made verbal responses (e.g. produced or labelled the stimulus) when they had been invited to attend to the stimulus or participate in the stimulus routine.

The analysis shows that all four children were active participants in the therapy process. Despite their young age, they demonstrated high levels of attention to the therapist and speech sound stimuli and participated in all aspects of the stimulus routines. Some of the children also initiated the stimulus routine themselves, particularly nonverbal enactments, e.g. T: *“in there (offers picture; gaze to child) shall we put mine in?”* C: *“(takes picture; reaches to post it)...(brings it to side of his face; lifts gaze to therapist)”* (Extract 11; Lines 002-004). The therapist invited the child to post the picture and instead the child drew the picture to his face, enacting the display for the camera. The present study did not systematically examine the nature of child initiations. However, the analysis suggests they may have been facilitated in some way by the layered structure and multimodal nature of the stimulus routine and distinctive nature of interaction with the camera.

9.4.6 Therapist evaluation of child responses

The study focused on therapist initiations of stimulus routine sequences and did not systematically examine the way therapists evaluated child responses. However, some interesting observations were made of therapist rejections that suggest this is worthy of future study. These relate to two specific aspects of practice: 1) therapists’ pursuit for a response when the child’s response did not

align with the design of the therapist's turn, including absent responses; and 2) therapists' pursuit for a revised response when the child's target production was inaccurate.

The analysis found therapists did not systematically evaluate child responses. This reflects existing descriptions of the approach that describe a non-corrective style of interaction (Calladine & Vance, 2019) and aligns it with child-led rather than therapist-led approaches. Kovarsky and Duchan (1997) describe how in child-centred intervention "repairs are aimed at keeping activities going and making sense of what is happening rather than at correcting performance" (p300). Rather than the nature of repair, it was the lack of it that was so striking about the MSIVT interaction. One might suggest that this could be problematic for the child because it creates ambiguity, yet in this study the children did not explicate confusion, but rather the opposite. The lack of pursuit seemed to create freedom for the child to participate on his or her own terms in a way that is consistent with child-led intervention (Kaiser et al., 2017; Kovarsky & Duchan, 1997).

McCartney (1989), in her study of phonology therapy, found that therapists gave children high levels of praise and children achieved high levels of success. This too was seen in the present study and it appeared to be facilitated by the range of possible responses available to the child. Letts (1985) found therapists showed a preference for approval rather than disapproval but she did not examine the nature of therapist feedback. In the present study, the analysis showed that when therapists made evaluations, their turns displayed similar features to the evaluations in Gardner's (1994) study involving children aged 3;11 to 5;9 with phonological disorder. Gardner (1994) found that when pursuing the child to have a re-try at a target sound or word production, the therapist used explicit phonetic and phonological description to help the child revise their production and get closer to target on their next attempt. Such metalinguistic information also contributes to the development of self-repair, which is an essential part of the therapy process (Gardner, 2006). This practice reflects the therapist's specialist knowledge in speech and language development but also the short- and long-term goals of therapy. Despite the younger age of children in the present study, and the different therapy approach, therapists' use of explicit articulatory and phonetic description in their evaluations of child responses was a characteristic feature of their repair practice. However, a distinctive feature of these evaluative utterances was the use of plural rather than singular pronouns. For example, *T: "Ooh (points to mouth) with our tongue, dee (makes cued artic gesture; sustained gaze to child)" C: "(postures linguolabial tongue placement) dee (gaze to therapist)"* (Extract 12; Lines 032-035). Notice also the multimodality in the therapist's turn,

pointing to her mouth as she provides a verbal articulatory description, and making a finger gesture as she re-demonstrates the stimulus. The child treats the therapist's turn as an invitation to produce the stimulus and effectively revises her production. This particular practice of using a plural pronoun to refer to the child rather than a singular one appeared to have a softening effect on the therapist's utterance. It is consistent with the procedural description Calladine (2009) developed and reflects the non-corrective principle of the approach (Calladine & Vance, 2019).

9.5 Phase 3: MSIVT interactions over an episode of care

Phase 3 aimed to examine therapist-child interaction during MSIVT from a longitudinal perspective in order to identify and describe change over the course of an episode of care. It showed that the onset and completion of therapy were both characterised by a prominence of implicit rather than explicit action, though there were quantitative and qualitative differences in the profile of interaction at these two stages of therapy. The findings exemplify the skill and creativity of the therapist in gradually engaging the young child with a distinct set of interactional practices, which empowered the child to be an active participator in therapy and provided the therapist with opportunities to support the child's development.

9.5.1 Evolution in action and practices

The therapist Sarah's preference for implicit forms of action at both stages of therapy reflects the non-directive principle of MSIVT, which is modelled on the interaction style described by Pamplona et al. (2001). Although Sarah did use explicit forms, they typically invited nonverbal rather than verbal responses. We did see her use *D4. Verbal request for production* in Video 7 but, by this stage in therapy, child Louise was initiating nonverbal enactments of the stimulus routine, imitating readily and repeatedly, and completing Sarah's utterances without hesitation. Sarah may have evaluated that Louise had, by then, the stimulus knowledge, output ability and confidence to respond to such explicit invitations without it compromising her participation or sense of safety and ease in the therapy environment. The timing and reliability of Louise's production responses provide evidence that she was a willing participator and there was no evidence to suggest she was not at ease. At both stages of therapy, Sarah's preference for plural pronouns in explicit requests is highly suggestive of an effort to soften their directness and further exemplifies the way MSIVT interaction can stay true to its non-directive principle.

The different profiles of therapist action in Videos 1 and 7 align with the psycholinguistic basis of MSIVT and reflect the goals of the approach. At the early stage of therapy, *A1. Demonstration for later* and *B1. Demonstration for attention* enabled Sarah to provide Louise with intense exposure to novel sound stimuli. This gave Louise opportunity to process incoming novel sounds and supported the development of new motor programs (Harding & Bryan, 2000; Stackhouse & Wells, 1997). As the MSIVT episode of care drew to an end, *D1. Demonstration for imitation* and *D4. Verbal request for production* gave Louise opportunity to retrieve, produce and refine these new motor programs. These findings show how the therapist's interactional practices can transition from the input-based MSIVT approach towards a more output-based approach if it is required. The pattern of interaction that we saw in Video 7 when Sarah and Louise engaged in repair sequences resembles that which is more characteristic of phonology therapy, e.g. Gardner (1997).

The present study used an approach similar to Flipi (2018) and Forrester (2008) to examine Sarah's initiations, and the actions they projected for Louise, from a longitudinal perspective. In Flipi (2018) and Forrester's (2008) studies, the analysis revealed a developmental trajectory in the way certain practices developed and emerged over time. In the present study, the analysis revealed differences in way Sarah delivered her initiations at the two stages of therapy, and the action that was accomplished. For example, at the early stage of therapy, Sarah's *A1. Demonstration for later* initiations appeared to reflect an initial goal of therapy to establish an environment in which there was minimal pressure on Louise to respond. She delivered these initiations recurrently with little or no pause in-between. This pattern of delivery resembles the clustered restart-relevant directives in the Reed et al. (2013) study. Louise, in response, displayed intermittent therapist- and camera-directed gaze and at times lowered her gaze. The absence of overlapping talk or unrelated action by Louise suggested she was attending to Sarah's demonstrations but she did not actively contribute to or direct the interaction. At the later stage of therapy, Sarah's *A1. Demonstration for later* initiations were embedded within sequences of interaction that featured more demanding initiations, such as *D2. Incomplete demonstration*, and they appeared to reinforce rather than demonstrate stimuli. By this stage of therapy, the timing and nature of Louise's responses were actively shaping the pattern of interaction and, as well as displaying her attention, at least one of her responses was a stimulus production. The analysis made a similar discovery in relation to *D1. Demonstration for imitation*. The one example in Video 1 did not elicit a production from Louise. Whereas, in Video 7, the timing and nature of Louise's production responses showed this was a jointly accomplished action. In contrast to Stivers and Rossano (2010), the likelihood of eliciting a

production response from Louise did not appear to be strongly related to the nature or design of Sarah's turn. The overlapping and repetitive nature of Louise's productions, as she displayed the stimulus picture for the camera in one hand and made the cued articulation gesture with her other hand, had the appearance of an action predetermined by the activity itself.

The finding that *C2. Verbal request for participation* featured at similar levels in both videos suggests it is an important action throughout therapy. Comparative analysis examined the design and consequences of Sarah's requests of Louise to display a stimulus picture for the camera. This showed that Sarah displayed a preference for the explicit verb *hold* in Video 1 and the ambiguous verb *show* in Video 7. This may represent evolution in Sarah's practice towards more implicit invitations for stimulus production. Despite the explicit verb in Video 1, Louise's participation required pursuit by Sarah on two occasions. This may reflect a young child learning the ropes of therapy. In Video 7, the timing and accuracy of Louise's participation had the characteristics of a child with experience, and in one of her responses she treated Sarah's invitation to *show* as an invitation to produce.

9.5.2 Orientation to the layered and multimodal stimulus routine

The findings show that both therapist and child orientated to the stimulus routine during therapy. At the early stage of therapy, Sarah used *A. Demonstration to the camera*, *B. Invitation to attend to the stimulus* and *C. Invitation to participate in the stimulus routine* to lay down the foundations of the routine through repeated exposure. Louise displayed attention to Sarah's enactments and participated nonverbally on request. By the end of therapy, the stimulus routine had become an interactional tool for Sarah to implicitly invite production from Louise in a non-directive manner. Louise, by this stage, was displaying her knowledge of the stimulus routine in her own initiations of nonverbal enactments and in the types of articulatory adjustments she made to her productions of target stimuli. Where Phase 2 revealed the layered structure and multimodal nature of the stimulus routine, the longitudinal analysis carried out in Phase 3 revealed its potential to facilitate and empower participation, both in terms of engagement in an activity, and in specific sequences of speech modification. As such, it exemplifies one of the ways that MSIVT provides opportunity for the young child to be an active participator in the therapy process. This may have implications for the outcomes of therapy (Weiss, 2004).

9.5.3 Insight into a therapist and young child's speech repair practices

The findings from the analysis of Video 7 reinforce those of previous studies that have shown two-year-old children engaging in phonetic repair in response to adult initiations (Gallagher, 1977; Tarplee, 1996). They also extend current knowledge because, unlike the children in Gallagher (1977) and Tarplee's (1996) studies, the children in the present study had a diagnosis of speech sound disorder. Sarah displayed a range of verbal, nonverbal and prosodic features in her evaluations when rejecting Louise's productions and inviting a repair, such as stimulus repetition with matched or mismatched prosody (as in the examples in Extract 9) and interjection with verbal and nonverbal articulatory prompt (as in Extract 10). Unlike Tarplee (1996), who found that the subtle presence of a pause before repeating a given word engendered a revision by the child in her study, the present study found a combination of verbal and nonverbal features most reliably elicited a revised production. Sarah's use of explicit articulatory description was similar to that displayed by the therapists in Gardner's (1994) study. However, by the end stage of therapy (Video 7), it did not appear to be an essential feature of the interaction. The timing of Louise's revisions suggested they were elicited by Sarah's non-specific interjections, e.g. 'ooh', *B2. Verbal request for attention*, e.g. 'watch me', and nonverbal articulatory prompt, e.g. (points to mouth). Whilst these features implicitly invited Louise to change something about her production, they did not give her specific instruction about what to change. The articulatory nature of Louise's revisions was therefore evidence that she had gained knowledge and experience from previous sessions.

It was interesting that in activities targeting the [ɖ] stimulus Sarah appeared to reject productions which the author perceived as accurate alveolar realisations for the target phoneme /d/. Sarah pursued Louise for productions with visible tongue position, i.e. to match the stimulus rather than the phoneme to which it orientated. One possible explanation is that Sarah had not perceived the accuracy of Louise's productions and was therefore seeking assurance from visible tongue position that therapy was stimulating new processing (Calladine & Vance, 2019). In general, one would expect therapy to progress a child from adaptive articulations toward more adult-like versions. McCartney (1989) suggested this was why she saw a low occurrence of augmented models in her study of late stage phonology therapy sessions. Even though Video 7 was Louise's final MSIVT session, it was her first session targeting the [ɖ] stimulus. Other productions made by Louise during this session were [ɖ d } g], suggesting she had multiple and unstable motor programs for the [ɖ] stimulus. Sarah's pursuit may therefore have reflected a goal to help Louise

establish a new motor program independently of existing processing that could later be merged with [d]. This reflects the therapist's specialist knowledge of CPSSD and speech processing. In this final session, the [d̥] stimulus was progressed from single sound level to sound sequence and real word levels. Louise made (accurate) productions at every level. It seems plausible to suggest the more gradual progression with /t/ stimuli over the course of the episode of care, and experience Sarah and Louise developed during this process, in some way enabled such a rapid progression with /d/ within a single session.

9.5.4 Camera as an interactional tool

The findings of Phase 3 reinforce that the video camera is an important interactional tool in MSIVT. At the early stage of therapy it facilitated a high level of the action *A1. Demonstration for later*. This type of action places minimal demands on the child in the live frame of therapy. The findings suggest Sarah used this to establish a non-pressurising basis for therapy from which she gradually drew Louise into the interaction as a more active participant in the here and now. The action *C2. Verbal request for participation* appeared from the outset of therapy; it enabled Sarah to actively involve Louise in activities and gave Louise opportunity to define her role in the process from the beginning. By the end of therapy the camera's presence facilitated Louise's initiations of the stimulus routine. Her displays of stimulus pictures and finger gestures to the camera reflected the nature of Sarah's prior demonstrations. By this stage, the camera had become a source of shared attention between Sarah and Louise and helped them to accomplish things together.

Existing descriptions of MSIVT describe the camera as a gateway to the child at a future moment in time (Calladine & Vance, 2019); in a future frame (Goffman, 1981). As in Phase 2, Phase 3 found that Sarah used the camera to facilitate her interaction with Louise in both the future and live frames of therapy. The longitudinal perspective on this interaction provided by Phase 3 appears to show Louise being empowered by the camera's presence to be an active participator in the live frame. The findings show Sarah and Louise working together to navigate this complex interaction.

Chapter 10: Implications and Conclusions

10.1 Introduction

The aim of this study was to examine the nature of multisensory input video-therapy (MSIVT) as an early intervention to treat cleft palate-related speech sound disorder (CPSSD) in children aged 18 months to three years. The purpose of this work is to develop understanding of MSIVT and its key components, expand the evidence base and extend and enhance current descriptions of the approach. This is necessary to support consistency in the labelling and delivery of MSIVT but also to support further development and evaluation of the approach. The findings describe how MSIVT is delivered as an episode of care within the NHS and what it looks like as an interactional process. This chapter will explicitly answer the research questions and consider the implications of the findings for clinical practice, teaching and future research, as well as acknowledge the limitations of the study.

10.2 How do therapists deliver MSIVT as an episode of care within the NHS: what is the structure of an episode of care; what speech sounds do therapists target; and what types of activities, materials and speech sound stimuli do they use?

Phase 1 of the study examined five episodes of care featuring five therapist-child dyads comprising three therapists and five children. Episodes of care consisted of up to nine monthly hospital-based therapy sessions over a period of five to 11 months. Twenty nine therapy sessions were available for analysis. A session lasted on average about 20 minutes, though this ranged from about 12 to 32 minutes. The three youngest children in the study received the shortest sessions. The findings suggest multiple speech sound targets that align with hypotheses about the nature of the child's CPSSD is a characteristic feature of MSIVT service delivery. The way therapists introduced targets across episodes of care varied but there appeared to be a distinction relating to whether therapy was 'diagnostic' or 'targeted' as well as the age of the child. For example, targets were introduced systematically in the episodes featuring the three oldest children (Louise, Ellie and Naomi), and for two of these children therapists targeted specific phonemes that were affected by specific cleft

speech patterns at the outset of therapy. There was, however, variation in how many sounds therapists targeted in each session.

The study examined 122 therapy activities that took place during 29 therapy sessions. A session typically consisted of five play-based activities and an activity lasted on average about four and a half minutes. The use of pictures in activities varied; ranging from 9% of activities in Laura-Ellie's episode of care to 68% of activities in Helen-Naomi's episode of care. The low occurrence of paper-based activities (in only 7% of activities) suggests this is not a characteristic feature of MSIVT service delivery.

The findings suggest the use of multiple speech sound stimuli produced with adaptive articulation and accompanying finger gestures are characteristic features of MSIVT. This aligns with existing descriptions of MSIVT, which refer to the use of innovative speech sound stimuli and manipulation of gestures and materials to make stimuli as salient as possible. There was variation, however, in relation to the levels that stimuli were targeted. This aspect of delivery does not fully align with existing descriptions of MSIVT or the underlying psycholinguistic basis of the approach (Calladine, & Vance, 2019; Harding & Bryan, 2000).

10.3 What are the interactional features of MSIVT, and in what ways do therapists establish a child's attention and stimulate their awareness and production of speech sounds?

Phase 2 of the study examined 120 minutes of interaction during 12 MSIVT sessions featuring four therapist-child dyads. It found MSIVT is characterised by structured sessions in which the therapist takes the lead but gives the child opportunities to influence what comes next. Activities are characterised by recurrent stimulus-routine sequences in which the therapist produces salient speech sound stimuli to the child or camera with accompanying gestures, verbal labels, articulatory descriptions and manipulations with objects and pictures. The findings suggest the layered structure and multimodal nature of the stimulus routine are characteristic features of MSIVT. Therapists use the multimodal features to make speech sound stimuli a salient part of the interaction. They draw the child's attention to the visual and acoustic properties of speech sounds to support their perception and recognition. The specific nature of articulatory descriptions, such as 'with my lips', gives the child explicit information to stimulate their awareness and production

of speech sounds. The findings suggest the layers of features within the stimulus routine and systematic manner in which therapists enact them create symbolic representations for the stimuli. The findings show therapists use these to implicitly invite the child to attend to and produce speech sounds.

A characteristic feature of the interaction was that therapists did not always invite a here and now response from the child; instead producing speech sound stimuli with camera-directed gaze. These demonstrations may stimulate the child's awareness of speech sounds in the session because the therapist, child and camera are all in close proximity, but they do not put pressure on the child to respond. Initiations fulfilling this action generated the category *A. Demonstration to the camera*. This is an unusual type of action that reflects the aims of video-therapy to stimulate the child's awareness and production of speech sounds at a future moment in time beyond the live session and support parents to implement home practice. The findings show therapists using the camera to navigate a complex framework of interaction, with actions designed to stimulate the child's awareness and production in the 'live' frame of therapy and simultaneously in a 'future' frame beyond the therapy session.

When therapists did invite a here and now response from the child, their initiation fulfilled one of three categories of action: *B. Invitation to attend to the stimulus*; *C. Invitation to participate in the stimulus routine*; and *D. Invitation to produce the stimulus*. This range of verbal and nonverbal responses that therapists invite is characteristic of the MSIVT approach and makes it distinct from other speech-based approaches. Therapists made both implicit and explicit invitations. However, they showed a preference for implicit invitations. In these, therapists made use of one or more 'response-mobilising' features that made it relevant for the child to respond, such as child-directed eye gaze. They also utilised the symbolic associations they had created with their manipulations of objects and pictures. In their explicit invitations, therapists displayed a preference for plural rather than singular pronouns and identified the camera as a recipient. The nature and range of child responses that therapists accepted is also characteristic of MSIVT; they accepted fitting, ill-fitting, and even absent responses. The nature of therapist invitations appeared to be designed so that they were purposely ambiguous since even when the therapist made an explicit request she did not always pursue the child for a response when one was absent. These distinctive features of the MSIVT interaction give the child a variety of options for how to respond and the freedom to choose whether or not to respond.

These interactional features and practices appear to work well to establish and maintain the child's visual and auditory attention to speech sound stimuli and their active involvement in therapy activities. They create opportunities, directly and indirectly, for the child to produce speech sound stimuli and all of the children in this study made use of these opportunities. The findings suggest the characteristic nature of the stimulus routine and the distinctive nature and pattern of therapist-child-camera interaction are key components of the MSIVT approach.

10.4 How does therapist-child interaction change over the course of an episode of care; (comparing first and last sessions)?

Phase 3 of the study compared the nature and pattern of interaction in two 10-minute excerpts from the first and last sessions of the episode of care featuring therapist Sarah and child Louise. The findings show differences in the interactional profile at the two stages of therapy, which reflect the goals of MSIVT. The early stage of therapy was characterised by a prominence of initiations fulfilling categories *A. Demonstration to the camera* and *B. Invitation to attend to the stimulus*. These involve the therapist making frequent and recurrent stimulus productions for the child to see and hear in the here and now of the therapy session (the live frame) or at a future moment in time (the future frame). The end of therapy was characterised by a prominence of initiations fulfilling category *D. Invitation to produce the stimulus*. These involve the therapist implicitly or explicitly inviting the child to produce the stimulus. Such responses require more levels of speech processing and therefore might be seen to be more demanding on the child. The findings show that an interactional profile of more demanding action does not need to compromise the child's participation. In Video 7, the timing and nature Louise's responses showed her actively shaping the nature and pattern of interaction: she initiated stimulus routine enactments; made stimulus productions in response to every one of Sarah's category D invitations; and revised her productions in response to Sarah's invitations for repair. Therapist initiations fulfilling category *C. Invitation to participate in the stimulus routine* featured at similar levels at both stages of therapy, suggesting this is an important feature of therapist-child interaction throughout therapy.

The findings of Phase 3 provide further support of the importance of the stimulus routine for its ability to facilitate the child's active involvement in therapy. Early in therapy the therapist's recurrent delivery of the stimulus routine and explicit invitation for the child's participation

orientate the child to its central importance within the interaction. By the end of therapy the child can use their knowledge of the stimulus routine to initiate interaction themselves.

The findings show that the video camera is more than a gateway to the child in a future frame; it is an interactional tool that supports the goals of MSIVT and the child's active involvement in the process. In the early stage of therapy it allows the therapist to provide the child with exposure to high levels of stimulus productions without making any demand for their here and now attention. Louise's initiations of picture displays in Video 7 show how it can be used as a tool to engage and empower the child to actively participate in nonverbal enactments with stimulus materials and create opportunities for stimulus production when the child is ready.

The high level of child stimulus productions and therapist evaluations seen in Video 7 gave rise to a different pattern of interaction resembling that seen in output-based therapy (Gardner, 1994). This pattern of interaction gives the child opportunity to update their stored lexical knowledge and refine their output processing. The findings suggest the evolution of interaction over the course of an episode of care is an important consideration in MSIVT for engaging the child in therapy and preparing the child for transitioning to an output-based approach later should it be indicated.

10.5 Implications for clinical practice and teaching

The findings of this study provide empirical evidence of how MSIVT is delivered as an episode of care. This evidence will inform how therapists talk about MSIVT and their own clinical practice, and will support consistent labelling of the MSIVT approach in clinical, supervision and teaching environments. This is important for facilitating meaningful conversations about the delivery of and provisions for MSIVT.

Since the study has not investigated the impact of therapy it cannot make recommendations for how MSIVT is delivered as an episode of care. However, some of the findings have implications for therapy planning. Phase 1 revealed a misalignment between the nature of target selection and stimulus design and existing descriptions of MSIM+/-VT in relation to the levels at which stimuli were targeted and how they were progressed session by session. Since the nature of speech input has implications for how it is processed by the child, it is important that targets and stimuli are considered carefully when planning therapy to ensure they align with the therapeutic goals.

Phases 2 and 3 make the distinctive interactional features of MSIVT visible. The findings expand current knowledge of MSIVT and extend and enhance existing descriptions of the approach. This is important given the unusual nature of the approach compared to other speech-based approaches. Existing descriptions refer to MSIM+/-VT as non-directive and non-corrective but more clinician-led than naturalistic approaches. The findings explicate what this means in a way that can be more easily and unambiguously communicated in clinical, supervision and teaching contexts. Again, this will support consistency in how MSIVT is discussed and how others are supported to implement it.

The study cannot make recommendations for how therapists interact with the child and camera during MSIVT because it has not evaluated the impact that interaction has on a child's outcomes. However, the findings do suggest the nature of interaction is important and therefore therapists are encouraged to pay particular attention to this aspect of therapy. For example, be mindful of their use of features like eye gaze, prosodic modification and silence because they increase the relevance (and therefore pressure) for the child to respond. Exposing such intricacies of therapy interaction will support therapists to reflect on their own clinical practice and support the practice of others and may lead to developments in the quality of therapy being provided (Horton, 2006). The study demonstrates the capacity of the video camera to facilitate the therapist's therapeutic interactions with the child within the session and delivery of stimulus demonstrations for future use beyond the session, as well as its potential to support the child's active involvement. Services are therefore encouraged to explore ways of providing video-therapy so that they can utilise this. Furthermore, therapy videos made during MSIVT are a valuable tool to support reflective practice and teaching.

The findings provide empirical evidence that MSIVT can generate the interest and participation of two-year-old children so that they can be exposed to demonstrations of speech sound stimuli and be supported to practice speech output. It is recognised that establishing and maintaining a child's attention and active involvement in therapy are an important bases for facilitating therapeutic change (Gardner, 1994; Weiss, 2004). It is therefore appropriate to suggest therapists consider the potential of this approach to support the engagement of young children (under three years of age) in speech-based therapy. However, research is needed to investigate whether the nature and level of child engagement in MSIVT are related to the child's outcomes of therapy.

10.6 Limitations of the study

This is the first study to investigate the nature of MSIVT as it is implemented in the NHS for young children with CPSSD. Phase 1 addressed the question *'How do therapists deliver MSIVT as an episode of care within the NHS: what is the structure of an episode of care; what speech sounds do therapists target; and what types of activities, materials and speech sound stimuli do they use?'* Despite the anecdotal preference for MSIM within the cleft SLT community, the study found that not all specialist SLT services combine it with video-therapy. Of the six services that do provide MSIVT, data were obtained from three, and one of these included the author's own data. The study recruited three therapists from the two external services and obtained data on five therapist-child dyads. This small sample size limits the extent to which the findings of Phase 1 can answer the research question. Where there was consistency across the five episodes of care it has been possible to identify characteristic features of MSIVT service delivery, such as therapists' use of multiple speech sound stimuli made with adaptive articulations. However, these findings would be more robust if they were supported by examination of data from the additional three services also implementing MSIVT. The finding of variation in respect of other features, such as the levels at which stimuli were targeted, is important to reveal and may reflect limitations of the existing evidence base for MSIVT, but it is not possible from such a small dataset to fully understand the nature of this variation. The findings indicate that further research is needed to explore and address the uncertainties with regards to target selection and stimulus design as well as the other features that showed variation across the episodes of care.

Phase 2 addressed the question *'What are the interactional features of MSIVT, and in what ways do therapists establish a child's attention and stimulate their awareness and production of speech sounds?'* Drawing on the detailed method of conversation analysis (CA) produced rich data on therapist-child-camera interaction in four therapist-child dyads. The analysis exposed features of interaction that would not have been made visible by other approaches. The findings effectively describe the interactional practices therapists use during MSIVT to establish a child's attention to speech sound stimuli, their participation in therapy activities, and their production of target stimuli. There are limitations, however, because the time-consuming nature of CA limited the breadth of analysis that could be performed within the scope of this study. For example, the study examined interaction in four therapist-child dyads but it did not compare interaction across dyads. The specific set of practices one therapist uses with one child in a given session or episode of care

may vary from one therapist and dyad to another. The findings show that different things are accomplished within the interaction depending on the nature of therapist initiation, so different interactional profiles will give rise to different profiles of responses. This may have consequences for the child in terms of the impact therapy has on them. Extending the current findings with an examination of similarities and differences across therapists and dyads would therefore usefully extend knowledge and description of the interactional nature of MSIVT.

Phase 2 systematically examined the practices therapists use to stimulate a child's awareness and production of speech sounds. It focused specifically on therapist initiations at the beginning of a sequence, which elicited initial productions by the child. It did not focus on therapist initiations following a mismatched production, that is, evaluations, which elicited a revision by the child. This was explored within the study, but not systematically. There is currently insufficient evidence of the influence of output practice on the outcomes of MSIVT for children with CPSSD. However, because of the theoretical possibility that output practice is important (Stackhouse & Wells, 1997), addressing this limitation of the current study would be a worthwhile line of enquiry for future research.

Phase 3 addressed the question '*How does therapist-child interaction change over the course of an episode of care; (comparing first and last sessions)?*' Unlike Phase 2, this phase of the study took into account the longitudinal nature of therapy. Supplementing the detailed method of CA with quantitative and longitudinal methods has developed understanding of MSIVT by recognising it is a process that takes place over time. The findings of Phase 3 in respect of changes in the nature and pattern of interaction in the Sarah-Louise dyad reflect the goals of MSIVT as they are stated in existing descriptions, but they are not generalizable because of the small sample size. The findings can be extended with a larger sample of therapist-child dyads. In doing so, it would be appropriate to consider including one or more videos from the middle stage of therapy to supplement the first and last sessions. Analysis of interim videos would give opportunity to track the emergence and evolution of certain practices, for example, the first time the therapist makes an explicit request for production and the nature of child initiations and stimulus productions. This would provide further insight into the evolution of MSIVT and child participation throughout the course of therapy.

Phases 2 and 3 have shown that the video camera is an interactive tool in MSIVT that facilitates a distinctive pattern of interaction and the distinct nature of action that is accomplished by therapist and child. As such, the findings of the study cannot be generalised to MSIM without video-therapy. This is an important message since approximately half of specialist SLT services in the UK report that they implement MSIM without video-therapy. Studies examining the nature (and impact) of MSIM without video-therapy should be considered in future research.

10.7 Implications for future research

The findings from this study have implications for future research to further investigate the nature of MSIVT and evaluate the approach. They will facilitate further expansion of the evidence base for MSIVT, which is urgently needed to support effective clinical practice and optimum allocation of NHS resources.

The findings of Phase 1, which describe the service delivery characteristics of five MSIVT episodes of care, will inform the development of future studies to develop and evaluate MSIVT. They give valuable insight into the structure of episodes of care as they were implemented, for example, the number, frequency and duration of sessions that therapists provided. This knowledge will help researchers to develop feasible intervention protocols that will support future implementation of MSIVT if evaluation determines it is an effective early intervention to treat CPSSD. Implementation is an important consideration when developing and evaluating complex interventions (Craig et al., 2008; MRC, 2000).

Phase 1 found variation in respect of some features of service delivery and discrepancies between clinical practice and existing descriptions of MSIVT, for example, the levels at which speech sound stimuli were targeted. This suggests uncertainties in current knowledge and reflects limitations of the existing evidence base for MSIVT. This needs to be addressed in order that these features of MSIVT can be more robustly defined. Such enquiry might involve experimental case studies using an alternating treatment study design, for example, to examine and compare the impact of single sound versus real word stimuli on the child's speech processing. It is important this is resolved to address current gaps in knowledge and update existing descriptions of MSIVT to support clinical practice and inform the development of a protocol for a trial to evaluate the approach.

The findings of Phase 1 suggest MSIVT has versatility as a targeted approach to address specific cleft speech patterns as well as a diagnostic approach to explore the nature of absent pressure consonants. Such versatility is an important requirement for any early intervention approach for children born with cleft palate because of the importance of adequate structure and function of the palate and hearing mechanisms for the development of good quality speech (Sell et al., 2017; Smyth & Wu, 2019). Early diagnosis provides opportunity for early treatment. In the present study, whether intervention was targeted or diagnostic appeared to be related to the therapists' approach to target selection and stimulus design. A worthy extension to the current study would be an exploration, using interviews or focus groups, of therapists' decision-making around these aspects of therapy planning. The objective of such a study would be a greater understanding of the rationale behind therapists' practices. This would enhance existing descriptions of MISVT and help to identify gaps in knowledge and evidence in order to inform future research.

Phases 2 and 3 used detailed methods of analyses to examine the interactional features of MSIVT and develop understanding of the approach. This was important because of the unusual nature of MSIVT compared to other speech-based approaches. The findings reveal the distinctive nature of MSIVT interaction and provide the most explicit description to date of MSIVT as an interactional process. These findings can be used as the basis of a protocol for a trial and may inform the development of suitable outcome measures. Importantly, the outcomes of an evaluation study will be better understood because of the work carried out in the present study to expose the nature of what happens during MSIVT and identify key components of the approach. Any trial of MSIVT will need to include a measure of treatment fidelity, perhaps with an approach similar to that used by Pamplona et al. (2001). They used a quantitative method similar to discourse analysis to assess the impact of training on the way parents interacted with 59 children aged three to 4;8 born with cleft palate. The findings of the present study will support treatment fidelity in a trial and meaningful application of this type of measure.

In the meantime, more research is needed to extend the current findings by comparing interaction across therapist-child dyads and examining interaction longitudinally in a larger sample of dyads. A systematic investigation of the practices therapists use to evaluate child target productions and support revisions of mismatched productions is also recommended. It was beyond the remit of the current study to examine the home delivery component of MSIVT as it was specifically interested in how therapists implement MSIVT sessions. However, since the nature of parent-led and home-

based implementation may influence the outcomes and experiences of therapy, it is important that future studies extend knowledge in this area.

10.8 Conclusions

This study has examined the nature of MSIVT as it was implemented as part of routine clinical care in the NHS. A pilot study was undertaken using data from the author's own clinical practice in one specialist SLT service. The main study collected data from five therapist-child dyads in two external SLT services. It featured three therapists and five children with CPSSD aged 1;6 to 2;11. The data consisted of 29 therapy sessions comprising 573 minutes of video recordings.

Phase 1 examined service delivery characteristics in the five external episodes of care. It found that the number of sessions provided was similar across therapist-child dyads and sessions took place, on average, once every four weeks. The duration of sessions varied with the younger children receiving shorter sessions. All episodes of care targeted multiple speech sounds at different levels and all therapists used adaptive articulations in their stimulus productions. The findings suggest these are characteristic features of MSIVT service delivery and this new knowledge will inform the development of future studies to develop and evaluate the approach. There was variation across dyads in relation to the levels at which stimuli were targeted and practice did not align with existing descriptions. This appeared to reflect a distinction between the targeted and diagnostic nature of therapy as well as limitations in the existing evidence base for MSIVT. More research is needed to resolve uncertainties in this area.

Phase 2 examined interactional characteristics in four therapist-child dyads. It found therapists use a distinct set of interactional practices to stimulate the child's awareness and production of speech sounds within the MSIVT session whilst simultaneously using the camera as a gateway to providing stimulation to the child at a future moment in time beyond the therapy session. Within the session, therapists invited a range of here and now responses from the child: their attention to speech stimuli; their participation in enactments of the stimulus routine; and their production of target stimuli. They showed a preference for implicit rather than explicit invitations, invited nonverbal as well as verbal responses, and did not systematically pursue the child for a response when one was absent or ill-fitting. The recurrent enactment of a layered and multimodal stimulus routine was central to the interaction. Despite the ambiguity they project, these sophisticated practices appear to empower the child to be actively involved and create freedom for the child to influence the

nature and pattern of interaction. The findings describe how therapists create opportunities to stimulate a young child's input and output processing within an input-based approach. They show that a therapist-led but nurturant style of interaction can give rise to child productions of target stimuli. As such, MSIVT does not face the challenges that Kaiser et al. (2017) describe of child-led naturalistic intervention. The findings demonstrate the interactional value of video-therapy in MSIVT and extend existing descriptions of the approach. They suggest MSIVT is worthy of further study as an early intervention to treat CPSSD, and perhaps other types of speech sound disorder, in two-year-old children.

Phase 3 extended the examination of interactional characteristics by profiling and comparing interaction in the first and last sessions of one episode of care. It found that, by the end of therapy, the pattern of interaction started to resemble output-based phonology therapy due to the joint contributions of both participants. The findings are limited by the small sample size, but they do show the potential of combining methods of analysis to further explicate the interactional nature of MSIVT and meaningfully extend current descriptions by taking into account the longitudinal and evolutionary nature of therapy.

This study has produced a detailed and analytic description of the characteristic features of MSIVT service delivery and interaction and in doing so has made current practice explicit. The findings will support consistent labelling of MSIVT and meaningful discussion about the approach in clinical and academic environments. They increase knowledge and understanding of the distinctive nature of MSIVT and have generated empirically-based hypotheses about its key components: the nature of target selection and stimulus design; the distinctive nature and pattern of therapist-child-camera interaction; and the layered and multimodal stimulus routine. These findings must now be extended and used to facilitate the development of high quality studies to evaluate the impact of MSIVT and investigate the influence of these components on the outcomes of therapy. The significance of this study is the basis it provides for designing these studies and for ensuring we understand and can implement the findings and evidence they produce.

References

- Andersen-Wood, L., & Smith, B. R. (2001) *Working with Pragmatics: A Practical Guide to Promoting Communicative Confidence*. Oxon: Speechmark Publishing Ltd.
- Baker, E., Croot, K., McLeod, S. & Paul, R. (2001) Psycholinguistic models of speech development and their application to clinical practice. *Journal of Speech, Language and Hearing Research*, 44:4, 685-702.
- Baker, E. & Williams, A. L. (2010) Complexity approaches to intervention. In A. L. Williams, S. McLeod & R. J. McCauley (Eds.), *Interventions for Speech Sound Disorders in Children*. (pp. 95-115). Maryland, USA: Brookes Publishing Co.
- Ball, M. J., Howard, S. J. & Miller, K. (2018) Revisions to the ExtIPA chart. *Journal of the International Phonetic Association*, 48:2, 155-164.
- Bandura, A. (1977) *Social Learning Theory*. London: Prentice-Hall.
- Barrett, J. & Extence, H. (2010) Speech and Language Therapist's predictions made at 18-month cleft palate speech assessments. How accurate are we? Poster presentation at the Craniofacial Society of Great Britain and Ireland Annual Scientific Conference; Liverpool, UK.
- Bercow, J. (2008) *The Bercow Report: A Review of Services for Children and Young People (0-19) with Speech, Language and Communication Needs*. Department for Children, Schools and Families.
- Bercow, J. (2018) *Bercow: Ten years on*. London: ICAN/Royal College of Speech and Language Therapists.
- Bessell, A., Sell, D., Whiting, P., Roulstone, S., Albery, L., Persson, M., Verhoeven, A., Burke, M. & Ness, A. (2013) Speech and language therapy interventions for children with cleft palate: A systematic review. *Cleft Palate-Craniofacial Journal*, 50:1, pp. e1-e17.
- Bishop, D. & Adams, C. (1990) A prospective study of the relationship between specific language impairment, phonological disorders and reading retardation. *Journal of Child Psychology and Psychiatry*, 31:7, 1027-1050.
- Bobkoff, K. & Panagos, J. M. (1986) The 'point' of language intervention. *Child Language Therapy and Teaching*, 2, 50-62.

Bowden, M., Harland, K. & Sommerlad, B. (1997) How early should speech & language skills be assessed: An audit of 50 consecutive assessments undertaken at 18 months of age. Presentation at the Craniofacial Society of Great Britain and Ireland Annual Scientific Conference; Essex, UK.

Bowen, C. & Cupples, L. (1999) Parents and children together (PACT): A collaborative approach to phonological therapy. *International Journal of Language & Communication Disorders*, 34:1, 35-83.

Braun, V. & Clarke, V. (2006) Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3:2, 77-101.

Britton, L. (2001a) Early intervention: The case in favour. Presentation at the Special Interest Group Study Day for Cleft Palate and Craniofacial Anomalies, Liverpool.

Britton, L. (2001b) Babble & first words workshop. Presentation at the Special Interest Group Study Day for Cleft Palate and Craniofacial Anomalies, Dublin.

Britton, L. (2008) Input modelling. Presentation at the Trent Region Cleft SLT Network Study Day, Rotherham.

Britton, L., Albery, L., Bowden, M., Harding-Bell, A., Phippen, G. & Sell, D. (2014) A cross-sectional cohort study of speech in five-year-olds with cleft palate ± lip to support development of national standards. *The Cleft Palate-Craniofacial Journal*, 51:4, 431-451.

Britton, L., Calladine, S., Extence, H., Phippen, G. & Pinkstone, M. (2017) Inequities in speech and language therapy provision for children born with cleft palate. *Royal College of Speech and Language Therapists Bulletin*; October. (pp. 12-15).

Broen, P. A., Moller, K., Carlstrom, J. Doyle, S. S., Devers, M. & Keenan, K. M. (1996) Comparison of the hearing histories of children with and without cleft palate. *Cleft Palate-Craniofacial Journal*, 33:2, 127-133.

Bruce, I., Harman, N., Williamson, P., Tierney, S., Callery, P., Mohiuddin, S., Payne, K., Fenwich, E., Kirkham, J. & O'Brien, K. (2015) The management of Otitis Media with Effusion in children with cleft palate (mOMEnt): A Feasibility study and economic evaluation. *Health Technology Assessment*, 19:68.

Bryan, A. & Howard, D. (1992) Frozen phonology thawed: The analysis and remediation of a developmental disorder of real word phonology. *European Journal of Disorders of Communication*,

27, 343-365.

Calladine, S. (2009) Multi-sensory input modelling therapy intervention for young children with cleft palate. Unpublished M.Sc. University of Sheffield.

Calladine, S. (2010) Teaching Toddlers about Talking. Presentation at the Craniofacial Society of Great Britain and Ireland Annual Scientific Conference; Liverpool, UK.

Calladine, S. (2011) *Teaching Toddlers about Talking: A Resource CD for Speech and Language Therapists*. Nottingham University Hospitals NHS Trust.

Calladine, S. (2016) Psycholinguistic-based therapy for young children born with cleft palate. Presentation at the North West North Wales Regional Link Therapist Meeting; Warrington, UK.

Calladine, S. & Vance, M. (2019) A psycholinguistic approach to therapy with very young children born with cleft palate. In A. Harding-Bell (Ed.), *Case Studies in Cleft Palate Speech: Data Analysis and Principled Intervention*. Chichester: J&R Press.

Chapman, K. L. (1991) Vocalizations of toddlers with cleft lip and palate. *Cleft Palate-Craniofacial Journal*, 28:2, 172-178.

Chapman, K. L. (1993) Phonologic processes in children born with cleft palate. *Cleft Palate-Craniofacial Journal*, 30:1, 64-72.

Chapman, K. L. & Hardin, M. A. (1992) Phonetic and phonologic skills of two-year-olds with cleft palate. *Cleft Palate-Craniofacial Journal*, 29:5, 435-443.

Chapman, K. L., Hardin-Jones, M., Schulte, J. & Halter, K. A. (2001) Vocal development of 9 month old babies with cleft palate. *Journal of Speech, Language, and Hearing Research*, 44:6, 1268-1284.

Chapman, K. L., Hardin-Jones, M. & Halter, K. A. (2003) The relationship between early speech and later speech and language performance for children with cleft lip and palate. *Clinical Linguistics and Phonetics*, 17:3, 173-197.

Chapman, K. L. & Willadsen, E. (2011) The development of speech in children with cleft palate. In S. Howard & A. Lohmander (Eds.), *Cleft Palate Speech: Assessment and Intervention*. (pp. 23-40). West Sussex: Wiley-Blackwell.

Cleft Lip and Palate Association (2016). <https://www.clapa.com/>.

Cole, J. (2009) Emergent feature structures: Harmony systems in exemplar models of phonology. *Language Sciences*, 31, 144-160.

Cole, K. N. & Dale, P. S. (1986) Direct language instruction and interactive language instruction with language delayed preschool children: A comparison study. *Journal of Speech and Hearing Research*, 29, 206-217.

Craig, P., Dieppe, P., Macintyre, S., Michie, S., Nazareth, I. & Petticrew, M. (2008) *Developing and Evaluating Complex Interventions: New Guidance*. London: MRC.

CRANE Database (2015) Annual progress report. On behalf of the Cleft Development Group.

CRANE Database (2016) Annual progress report. On behalf of the Cleft Development Group.

CRANE Database (2017) Annual progress report. On behalf of the Cleft Development Group.

CRANE Database (2018) Annual progress report. On behalf of the Cleft Development Group.

Crosbie, S., Holm, A. & Dodd, B. (2005) Intervention for children with severe speech disorder: A comparison of two approaches. *International Journal of Language and Communication Disorders*, 40:4, 467-491.

Dive, D. (2001) Input modelling therapy. Presentation at the Cleft Palate Special Interest Group Study Day for Cleft Palate and Craniofacial Anomalies, Dublin.

Dobinson, C. & Wren, Y. (2013) *Creating Practice-Based Evidence: A Guide for SLTs*. Chichester: J&R Press.

Dodd, B. (1979) Lip reading in infants: Attention to speech presented in- and out-of-synchrony. *Cognitive Psychology*, 11:4, 78-484.

Dodd, B. (1987) The acquisition of lip reading Skills by normally hearing children. In B. Dodd & R. Campbell (Eds.), *Hearing by Eye: The Psychology of Lip Reading*. London: Lawrence Erlbaum Associates Limited.

Dodd, B. (2005) (Ed.) *Differential Diagnosis and Treatment of Children with Speech Disorder* (2nd Edition). Chichester: Whurr Publishers Ltd.

Dodd, B. & Bradford, A. (2000) A comparison of three therapy methods for children with different types of developmental phonological disorder. *International Journal of Language and*

Communication Disorders, 25:2, 189-209.

Drew, P. (1981) Adults' corrections of children's mistakes: A response to Wells and Montgomery. In P. French & M. MacLure (Eds.), *Adult-Child Conversation: Studies in Structure and Process*. (pp. 244-267). London: Croom Helm.

Eastbrooks, W. (2006) Auditory-verbal therapy and practice. In W. Eastbrooks (Ed.), *Auditory-Verbal Therapy and Practice*. Washington, DC: Alexander Graham Bell Association for the Deaf and Hard of Hearing. Cited by R. Ronkainen (2011) Enhancing listening and imitation skills in children with cochlear implants: The use of multimodal resources in speech and language therapy. *Journal of Interactional Research in Communication Disorders*, 2:2, 245-269.

Elliott, R. (2001) Psychotherapy change process research: Realizing the promise. *Psychotherapy Research*, 20:2, 123-135.

Enderby, P., Pickstone, C., John, A., Fryer, K., Cantrell, A. & Papaioannou, D. (2009) *Resource Manual for Commissioning and Planning Services for SLCN: Cleft Lip/Palate and Velopharyngeal Impairment*. London: RCSLT.

Estrem, T. & Broen, P. A. (1989) Early speech production of children with cleft palate. *Journal of Speech and Hearing Research*, 32:1, 12-23.

Extence, H. (2017) Catching them early! Minimising cleft speech errors in the early years and engaging families in speech therapy. Presentation at the Clinical Excellence Network Study Day for Cleft Palate and Craniofacial Anomalies, Newcastle.

Fenson, L., Marchman, V. A., Thal, D. J., Dale, P. S., Reznick, J. S. & Bates, E. (2007) *MacArthur-Bates Communicative Development Inventories* (2nd Edition). Baltimore: Paul H Brookes.

Ferguson, C. A. & Farwell, C. B. (1975) Words and sounds in early language acquisition. *Language*, 51:2, 419-439.

Fey, M. E. (1986) *Language Intervention with Young Children*. London: Taylor & Francis Ltd.

Field, A. (2009) *Discovering Statistics Using SPSS* (3rd Edition). London: SAGE Publications Ltd.

Fliipi, A. (2018) Making knowing visible: Tracking the development of the response token *yes* in second turn position. In S. Pekarek Doehler, J. Wagner & E. González-Martínez (Eds.), *Longitudinal Studies on the Organization of Social Interaction*. (pp. 39-66). London: Springer Nature.

Flynn, T. & Lohmander, A. (2014) A longitudinal study of hearing and middle ear status in individuals with UCLP. *Otolaryngology and Neurotology*, 35:6,989-996.

Flynn, T., Möller, C., Jöhnsson, R. & Lohmander, A. (2009) The high prevalence of otitis media with effusion in children born with cleft lip and palate as compared to children without clefts. *International Journal of Pediatric Otorhinolaryngology*, 73, 1441-1446.

Forrester, M. A. (2008) The emergence of self-repair: A case study of one child during the early preschool years. *Research on Language and Social Interaction*, 41:1, 99-128.

Foulkes, P. & Docherty, G. (2006) The social life of phonetics and phonology. *Journal of Phonetics*, 34, 409-438.

Gallagher, T. M. (1977) Revision behaviours in the speech of normal children developing language. *Journal of Speech and Hearing Research*, 20:2, 303-318.

Gardner, H. (1994) Doing talk about speech: A study of speech/language therapists and phonologically disordered children working together. D.Phil. Thesis. University of York.

Gardner, H. (1997) Are your minimal pairs too neat? The dangers of phonemicisation in phonology therapy. *European Journal of Disorders of Communication*, 32, 167-175.

Gardner, H. (1998) Social and cognitive competencies in learning: Which is which? In I. Hutchby & J. Moran-Ellis (Eds.), *Children and Social Competence: Arenas of Action*. (pp. 115-133). London: Falmer.

Gardner, H. (2006) Training others in the art of therapy for speech sound disorders: An interactional approach. *Child Language Teaching and Therapy*, 22:1, 27-46.

Gardner, H. (2007) A comparison of a mother and a therapist working on child speech. In K. Richards & P. Seedhouse (Eds.), *Applying Conversation Analysis*. (pp. 56-74). Hampshire: Palgrave Macmillan.

Geronikou, E., & Rees, R. (2016) Psycholinguistic profiling reveals underlying impairments for Greek children with speech disorders. *Child Language Teaching and Therapy*, 32:1 95-110.

Gierut, J. A., Morrisette, M. L. & Ziemer, S. M. (2010) Nonwords and generalization in children with phonological disorders. *American Journal of Speech-Language Pathology*, 19:2, 167-177.

- Girolametto, L., Pearce, P. S. & Weitzman, E. (1996) Interactive focused stimulation for toddlers with expressive vocabulary delays. *Journal of Speech and Hearing Research*, 39:6, 1274-1283.
- Goffman, E. (1981) *Forms of Talk*. Philadelphia: University of Pennsylvania Press.
- Goodwin, C. & Heritage, J. (1990) Conversation analysis. *Annual Review of Anthropology*, 19:1, 283-307.
- Grunwell, P. (1981) The development of phonology: A descriptive profile. *First Language*, 2:6, 161-191.
- Grunwell, P. (1988) Phonological assessment, evaluation and explanation of speech disorders in children. *Clinical Linguistics & Phonetics*, 2:3, 221-252.
- Grunwell, P. & Sell, D. A. (2001) Speech and cleft palate/velopharyngeal anomalies. In A. C. Watson, D. A. Sell & P. Grunwell (Eds.), *Management of Cleft Lip and Palate*. (pp. 68-86). London and Philadelphia: Whurr Publishers.
- Hardin-Jones, M. A. & Chapman, K. (2008) The impact of early lexical intervention on speech and lexical development for toddlers with cleft palate: A retrospective look at outcome. *Language, Speech & Hearing Services in Schools*, 39:1, 89-96.
- Hardin-Jones, M. A. & Chapman, K. (2011) Cognitive and language issues associated with cleft palate. *Seminars in Speech and Language*, 32:2, 127-140.
- Hardin-Jones, M. A. & Chapman, K. (2014) Early lexical characteristics of toddlers with cleft lip and palate. *The Cleft Palate-Craniofacial Journal*, 51:6, 622-631.
- Hardin-Jones, M. A. & Chapman, K. (2018) The implications of nasal substitutions in the early phonology of toddlers with repaired cleft palate. *The Cleft Palate-Craniofacial Journal*, 55:9, 1258-1266.
- Harding, A. (2001) The use of multi-sensory input modelling to stimulate output processing. Presentation at the 9th International Congress on Cleft Palate and Related Craniofacial Anomalies, Gothenburg.
- Harding, A. (2004) Multiple objective therapy for communication difficulties with 22q11. Invited lecture Maxappeal Conference; Cambridge, UK.

Harding, A. & Bryan, A. (2000) The use of multi-sensory input modelling to stimulate speech output processing. A teaching and demonstration video, Addenbrooke's NHS Trust & East & North Herts NHS Trust. Available from Cleft.NET.East Box 46, Addenbrookes Hospital, Cambridge University Hospitals Foundation Trust, Cambridge, UK.

Harding, A. & Grunwell, P. (1996) Characteristics of cleft palate speech. *European Journal of Disorders of Communication*, 31:4, 331-357.

Harding, A. & Grunwell, P. (1998) Notes and discussion: active versus passive cleft-type speech characteristics. *International Journal of Language & Communication Disorders*, 33:3, 329-352.

Harding-Bell, A. & Howard, S. (2011) Phonological approaches to speech difficulties associated with cleft palate. In S. Howard & A. Lohmander (Eds.), *Cleft Palate Speech: Assessment and Intervention*. (pp. 275-291). West Sussex: Wiley-Blackwell.

Henningsson, G. & Isberg, A. (1987) Influence of palatal fistulae on speech and resonance. *Folia Phoniatica*, 39:4, 183-191.

Hodsen, B. W., Chin, L., Redmond, B., Simpson, B. (1983) Phonological evaluation and radiation of speech deviations of a child with a repaired cleft palate: A case study. *Journal of Speech and Hearing Disorders*, 48, 93-98.

Horton, S. (2006) A framework for description and analysis of therapy for language impairment in aphasia. *Aphasiology*, 20:6, 528-564.

Horton, S. & Byng, S. (2000) Examining interaction in language therapy. *International Journal of Language and Communication Disorders*, 35:3, 355-375.

Howell, J. & Dean, E. (1994) *Treating Phonological Disorders in Children. Metaphon-Theory to Practice* (2nd Edition). London: Whurr Publishers Ltd.

Hulterstam, I. & Nettelbladt, U. (2002) Clinician elicitation strategies and child participation. Comparing two methods of phonological intervention. *Logopedics Phoniatics Vocology*, 27:4, 155-168.

Hutchby, I. & Wooffitt, R. (2008) *Conversation Analysis* (2nd Edition). Cambridge: Polity.

International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (1996) *Guidelines for Good Clinical Practice*.

International Phonetic Alphabet (2018) <https://www.internationalphoneticassociation.org>.

James Lind Alliance (2012) *Cleft lip and palate: Top 10 priorities*. <http://www.jla.nihr.ac.uk/top-10-priorities/>.

Jefferson, G. (2004) Glossary of transcript symbols with an introduction. In G. A. Lerner (Ed.), *Conversation Analysis: Studies from the First Generation*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Johnson, C. J., Beitchman, J. H., Young, A., Escobar, M., Atkinson, L., Wilson, B., Brownlie, E. B., Douglas, L., Taback, N., Lam, I. & Wang, M. (1999) Fourteen-year follow-up of children with and without speech/language impairments: Speech/language stability and outcomes. *Journal of Speech, Language, and Hearing Research*, 42, 744-760.

Jones, C. E., Chapman, K. L. & Hardin-Jones, M. A. (2003) Speech development of children with cleft palate before and after palatal surgery. *The Cleft Palate-Craniofacial Journal*, 40:1, 19-31.

Kaiser, A. P. & Hampton, L. H. (2017) Enhanced milieu teaching. In R. McCauley, M. Fey & R. Gillian (Eds.), *Treatment of Language Disorders in Children* (2nd Edition). (pp. 113-145). Baltimore, MD: Brookes.

Kaiser, A. P., Scherer, N. J., Frey, J. R. & Roberts, M. Y. (2017) The effects of enhanced milieu teaching with phonological emphasis on the speech and language skills of young children with cleft palate: A pilot study. *American Journal of Speech-Language Pathology*, 26, 806-818.

Kent, R. D. (1981) Sensorimotor aspects of speech development. In R. N. Aslin, J. R. Alberts & M. R. Peterson (Eds.), *Development of Perception: Psychobiological Perspectives*. New York: Academic Press.

Kent, R. & Read, C. (2002) The acoustic characteristics of consonants. In R. Kent & C. Read (Eds.), *Acoustic Analysis of Speech*. (2nd Edition). Australia; United States: Singular/Thomson Learning.

Kern, S., Davis, B. & Zink, I. (2010) From babbling to first words in four languages: Common trends, cross language and individual differences. In J. M. Hombert & F. d'Errico (Eds.), *Becoming Eloquent*. (pp. 205-232). Cambridge, UK: John Benjamin's Publishers.

Klintö, K., Salameh, E. K., Olsson, M., Fylnn, T., Svensson, H. & Lohmander, A. (2014) Phonology in Swedish-speaking 3-year-olds born with cleft lip and palate and the relationship with consonant

production at 18 months. *International Journal of Language & Communication Disorders*, 49:2, 240-254.

Klintö, K. Falk, E., Wilhelmsson, S., Schönmeyr, B. & Becker, M. (2018) Speech in 5-year-olds with cleft palate with or without cleft lip treated with primary palatal surgery with muscle reconstruction according to Sommerlad. *The Cleft Palate-Craniofacial Journal*, 55:10, 1399-1408.

Koshik, I. (2002) Designedly incomplete utterances: A pedagogical practice for eliciting knowledge displays in error correction sequences. *Research on Language and Social Interaction*, 35:3, 277-309.

Kovarsky, D. & Duchan, J. F. (1997) The interactional dimensions of language therapy. *Language, Speech and Hearing Services in Schools*, 28:3, 297-307.

Kuhl, P. K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S., & Iverson, P. (2006) Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Developmental Science*, 9;2, F13-F21.

Lane, H. & Wren, Y. (2018) Patterns of early speech and language therapy service delivery for children born with cleft palate in the first year of life. Poster presentation at the Craniofacial Society of Great Britain and Ireland Annual Scientific Conference; Birmingham, UK.

Leahy, M. M. (2004) Therapy talk: Analyzing therapeutic discourse. *Language, Speech and Hearing Services in School*, 35, 70-81.

Leahy, M. M. & Walsh, I. P. (2010) Paying attention to therapy discourse: Identifying therapy processes and practice in talk about talk. *Seminars in Speech and Language*, 31:2, 98-110.

Letts, C. (1985) Linguistic interaction in the clinic – how do therapists do therapy? *Child Language Teaching and Therapy*, 1, 321-331.

Lead SLT Forum (2016) National (UK) Standards for Speech for Children born with Cleft Palate (+/- Cleft Lip/Alveolus). Developed October 2009 and updated November 2016.

Linell, O., Gustavsson, L. & Juvonen, P. (1988) Interactional dominance in dyadic communication: A Presentation of initiative-response analysis. *Linguistics*, 26, 415-442.

Lohmander, A., Persson, C. & Owman-Moll, P. (2002) Unrepaired clefts in the hard palate: Speech deficits at the ages of 5 and 7 years and their relationship to size of the cleft. *Scandinavian Journal of Plastic and Reconstructive Surgery and Hand Surgery*, 36:6, 332-339.

Lohmander, A., Olsson, M. & Flynn, T. (2011) Early consonant production in Swedish infants with and without unilateral cleft lip and palate and two-stage palatal repair. *The Cleft Palate-Craniofacial Journal*, 48:3, 271-285.

McCartney, E. (1989) Constructive communication failure: The response of speech disordered children to requests for clarification. *British Journal of Disorders of Communication*, 16:3, 147-157.

McCauley, R. J., Strand, E., Lof, G. L., Schooling, T. & Frymark, T. (2009) Evidence-based systematic review: Effects of nonspeech oral motor exercises on speech. *American Journal of speech-Language Pathology*, 18, 343-360.

McCormack, J., Harrison, L. J., McLeod, S. & McAllister, L. (2011) A Nationally representative study of the association between communication impairments at 4-5 years and children's life activities at 7-9 years. *Journal of Speech, Language, and Hearing Research*, 54, 1328-1348.

McCune, L. & Vihman, M. M. (2001) Early phonetic and lexical development: A productivity approach. *Journal of Speech, Language, and Hearing Research*, 44:3, 670-684.

Medical Research Council (2000) A Framework for the Development and Evaluation of RCTs for Complex Interventions to Improve Health. London: MRC.

Meltzoff, A. (1988) Infant imitation and memory: Nine-month-olds in immediate and deferred tests. *Child Development*, 59:1, 217-225.

Morgan, L. & Wren, Y. E. (2018) A systematic review of the literature on early vocalizations and babbling patterns in young children. *Communication Disorders Quarterly*, 40:1, 3-14.

Ness, A. Wills, A. K., Mahmoud, O., Hall, A., Sell, D., Smallridge, J., Southby, L., Stokes, D., Toms, S., Waylen, A., Wren, Y., Sandy, J. R. (2017) Centre-level variation in treatment and outcomes and predictors of outcomes in 5-year-old children with non-syndromic unilateral cleft lip treated within a centralized service: The Cleft Care UK study. Part 6: Summary and implications. *Orthodontics and Craniofacial Research*, 20:2, 48-51.

NHS England (2013) NHS Standard Contract for Cleft Lip and/or Palate Services including Non-Cleft Velopharyngeal Dysfunction (VPD) (All Ages).

Norris, J. A. & Hoffman, P. R. (1990) Comparison of adult-initiated vs. child-initiated interaction styles with handicapped prelanguage children. *Language, Speech, and Hearing Services in Schools*,

21, 28-36.

Oberman, L. M., Pineda, J. A., Ramachandran, V. S. (2007) The human mirror neuron system: A link between action observation and social skills. *Social Cognitive Affective Neuroscience*, 2:1, 62-66.

Oller, D. K. (2000) *The Emergence of the Speech Capacity*. Mahwah, NJ: Lawrence Erlbaum Associates.

Pamplona, M. C., Ysunza, A. & Urióstegui, C. (1996) Linguistic interaction: The active role of parents in speech therapy for cleft palate patients. *International Journal of Pediatric Otorhinolaryngology*, 37, 17-27.

Pamplona, M.C, Ysunza, A. & Jiménez-Murat, Y. (2001) Mothers of children with cleft palate undergoing speech intervention change communicative interaction. *International Journal of Pediatric Otorhinolaryngology*, 59:3, 173-179.

Pamplona, M. C. & Ysunza, A. (2000) Active participation of mothers during speech therapy improved language development of children with cleft palate. *Scandinavian Journal of Plastic and Reconstructive Surgery and Hand Surgery*, 34, 231-236.

Panagos, J. M., Bobkoff, K. & Scott, C. M. (1986) Discourse analysis of language intervention. *Child Language Teaching and Therapy*, 2:2, 211-229.

Panagos, J. M., Bobkoff Katz, K., Kovarsky, D. & Prelock, P. A. (1988) The nonverbal component of clinical lessons. *Child Language Teaching and Therapy*, 4:3, 278-296.

Pascoe, M., Stackhouse, J. & Wells, B. (2006) *Persisting Speech Difficulties in Children: Children's Speech and Literacy Difficulties, Book 3*. Chichester: Whurr Publishers Limited.

Passy, J. (1993) *Cued Articulation* (2nd Edition). Bodmin, UK: STASS Publications.

Peterson-Falzone, S. J., Trost-Cardamone, J. E., Karnell, M. P. & Hardin-Jones, M. (2006) *The Clinician's Guide to Treating Cleft Palate Speech*. Missouri: Mosby, Inc.

Piggott, H. (2014) Speech sound therapy. In G. Phippen (Ed.), *Speech Therapy in Cleft Palate and Velopharyngeal Dysfunction*. (pp. 173-201). The Spires Cleft Centre SLT Team. Chichester: J&R Press.

Piggott, H. & Jeyes, F. (2014) Active nasal fricative (ANF). In G. Phippen (Ed.), *Speech Therapy in*

Cleft Palate and Velopharyngeal Dysfunction. (pp. 231-240). The Spires Cleft Centre SLT Team. Chichester: J&R Press.

Prutting, C. A., Bagshaw, N., Goldstein, H., Juskowitz, S. & Umen, I. (1978) Clinician-child discourse: Some preliminary questions. *Journal of Speech and Hearing Disorders*, 43, 123-139.

Purdy, S. C., Kim, J-H & Harding-Bell, A. (2019) Hearing loss and the impact on speech in children with cleft palate. In A. Harding-Bell (Ed.), *Case Studies in Cleft Palate Speech: Data Analysis and Principled Intervention*. Chichester: J&R Press.

Reed, B. S., Reed, D. & Haddon, E. (2013) NOW or NOT NOW: Coordinating restarts in the pursuit of learnables in vocal master classes. *Research on Language and Social Interaction*, 46:1, 22-46.

Rees, R. (2001) Principles of psycholinguistic intervention. In J. Stackhouse & B. Wells (Eds.), *Children's Speech and Literacy Difficulties 2: Identification and Intervention*. (pp. 41-65). London and Philadelphia: Whurr Publishers.

Regan, J. B. & Versaci, A. (1977) A home program for improving voice and speech quality of infants with repaired cleft palate. *Rhode Island Medical Journal*, 60:8, 384-5

Ripich, D. N., Hambrecht, G., Panagos, J. M. & Prelock, P. A. (1984) An analysis of articulation and language discourse patterns. *Journal of Childhood Communication Disorders*, 7:2, 17-26.

Ronkainen, R. (2011) Enhancing listening and imitation skills in children with cochlear implants: The use of multimodal resources in speech and language therapy. *Journal of Interactional Research in Communication Disorders*, 2:2, 245-269.

Russell, J. & Grunwell, P. (1993) Speech development in children with cleft lip and palate. In P. Grunwell (Ed.), *Analysing Cleft Palate Speech*. (pp. 19-47). London: Whurr Publishers Ltd.

Russell, J. & Harding, A. (2001) Speech development and early intervention. In A. C. H. Watson, D. A. Sell & P. Grunwell (Eds.), *Management of Cleft Lip and Palate*. (pp. 191-209). London and Philadelphia: Whurr Publishers.

Russell, J. & Albery, L. (2005) *Practical Intervention for Cleft Palate Speech*. London: Routledge.

Sandy, J., Williams, A., Mildinhall, S., Murphy, T., Bearn, D., Shaw, B., Sell, D., Devlin, B. & Murray, J. (1998) The Clinical Standards Advisory Group (CSAG) Cleft Lip and Palate Study. *British Journal of Orthodontics*, 25:1, 21-30.

Schegloff, E. A. (2007) *Sequence Organization in Interaction: Volume 1*. New York, USA: Cambridge University Press.

Schegloff, E. A. & Sacks, H. (1973) Opening up closings. *Semiotica*, 8, 289-327.

Scherer, N. J. (1999) The speech and language status of toddlers with cleft lip and/or palate following early vocabulary intervention. *American Journal of Speech-Language Pathology*, 9:1, 81-93.

Scherer, N. J. & Kaiser, A. P. (2007) Early intervention for children with cleft palate. *Infants & Young Children*, 20:4, 355-366.

Scherer, N. J. (1999) The speech and language status of toddlers with cleft lip and/or palate following early vocabulary intervention. *American Journal of Speech-Language Pathology*, 9:1, 81-93.

Scherer, N. J., Williams, A. L. & Proctor-Williams, K. (2008a) Early and later vocalization skills in children with and without cleft palate. *International Journal of Pediatric Otorhinolaryngology*, 72:6, 827-840.

Scherer, N. J., D'Antonio, L. & McGahey, H. (2008b) Early intervention for speech impairment in children with cleft palate. *The Cleft Palate-Craniofacial Journal*, 45:1, 18-31.

Scherer, N. J. & Kaiser, A. P. (2010) Enhanced milieu teaching with phonological emphasis for children with cleft lip and palate. In A. L. Williams, S. McLeod & R. J. McCauley (Eds.), *Interventions for Speech Sound Disorders in Children*. (pp. 427-452). Maryland, USA: Brookes Publishing Co.

Scherer, N. J. & Louw, B. (2011) Early communication assessment and intervention. In S. Howard & A. Lohmander (Eds.), *Cleft Palate Speech*. (pp. 259-274). West Sussex: Wiley-Blackwell.

Schreibman, L., Dawson, G., Stahmer, A. C., Landa, R., Rogers, S. J., McGee, G. G., Kasari, C., Ingersoll, B., Kaiser, A. P., Bruinsma, Y., McNerney, E., Wetherby, A. & Halladay, A. (2015) Naturalistic developmental behavioural Interventions: Empirically validated treatments for autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 45:8, 2411-2428.

Sell, D. A., Harding-Bell, A. & Grunwell, P. (1994) A screening assessment of cleft palate speech (Great Ormond Street Speech Assessment). *European Journal of Disorders of Communication*, 29:1, 1-15.

- Sell, D. A., Harding-Bell, A. & Grunwell, P. (1999) GOS.SP.ASS.'98: An assessment for speech disorders associated with cleft palate and/or velopharyngeal dysfunction (revised). *International Journal of Language & Communication Disorders*, 34:1, 17-33.
- Sell, D. & Harding-Bell, A. (2012) Cleft palate and velopharyngeal anomalies. In M. Kersner & J. A. Wright (Eds.), *Speech and Language Therapy: The Decision-Making Process when Working with Children* (2nd Edition). (pp. 162-172). Oxon: Routledge.
- Sell, D., Southby, L., Wren, Y., Wills, A. K., Hall, A., Mahmoud, O., Waylen, A., Sandy, J. R. & Ness, A. R. (2017) Centre-level variation in speech outcome and interventions, and factors associated with poor speech outcomes in 5-year-old children with non-syndromic unilateral cleft lip and palate: The Cleft Care UK study. Part 4. *Orthodontics and Craniofacial Research*, 20:2, 27-39.
- Sheridan, M. D. (2008) *From Birth to Five Years: Children's Developmental Progress* (3rd Edition). London and New York: Routledge.
- Shriberg, L. D., Austin, D., Lewis, B. A., McSweeney, J. L. & Wilson, D. L. (1997) The Percentage of Consonants Correct (PCC) Metric: Extensions and Reliability Data. *Journal of Speech, Language, and Hearing Research*, 40:4, 708-722.
- Shriberg, L. D., Lewis, B. L., Tomblin, J. B., McSweeney, J. L., Karlsson, H. B. & Scheer, A. R. (2005) Toward diagnostic and phenotype markers for genetically transmitted speech delay. *Journal of Speech, Language, and Hearing Research*, 48, 834-852.
- Sidnell, J. (2010) *Conversation Analysis: An Introduction*. West Sussex: Wiley-Blackwell.
- Simmons-Mackie, N., Damico, J. S. & Damico, H. L. (1999) A qualitative study of feedback in aphasia treatment. *American Journal of Speech-Language Pathology*, 8, 218-230.
- Sinclair, J. M. & Coulthard, R. M. (1975) *Towards an Analysis of Discourse*. Oxford: Oxford University Press.
- Smyth, A. & Wu, J. (2019) Cleft palate outcomes and prognostic impact of palatal fistula on subsequent velopharyngeal function – A retrospective cohort study. *The Cleft Palate-Craniofacial Journal*, 1-5. <https://doi.org/10.1177%2F1055665619829388>.
- Speake J. & Harding-Bell, A. (2019) Principles of intervention. In A. Harding-Bell (Ed.), *Case Studies in Cleft Palate Speech: Data Analysis and Principled Intervention*. Chichester: J&R Press.

Stackhouse, J. & Wells, B. (1993) Psycholinguistic assessment of developmental speech disorders. *International Journal of Language & Communication Disorders*, 28:4, 331-348.

Stackhouse, J. & Wells, B. (1997) *Children's Speech and Literacy Difficulties: A Psycholinguistic Framework*. London: Whurr Publishers Ltd.

Stackhouse, J. & Wells, B. (Eds.) (2001) *Children's Speech and Literacy Difficulties 2: Identification and Intervention*. London: Whurr Publishers Ltd.

Stackhouse, J., Vance, M., Pascoe, M. & Wells, B. (2007) *Compendium of Auditory and Speech Tasks: Children's Speech and Literacy Difficulties 4 with CD-ROM*. Chichester: John Wiley & Sons.

Stark, R. E. (1980) Stages of speech development in the first year of life. In G. H. Yeni-Komshian, C. A. Ferguson & J. Kavanagh (Eds.), *Child Phonology Volume 1: Production*. New York: Academic Press.

Stivers, T. & Rossano, F. (2010) Mobilizing response. *Research on Language and Social Interaction*, 43:1, 3-31.

Stoel-Gammon, C. & Williams, A. L. (2013) Early phonological development: Creating an assessment test. *Clinical Linguistics & Phonetics*, 27:4, 278-286.

Sutherland, D. & Gillon, G. T. (2007) Development of phonological representations and phonological awareness in children with speech impairment. *International Journal of Language & Communication Disorders*, 42:2, 229-250.

Sundman, H., Flynn, T., Tengroth, B. & Lohmander, A. (2016) ABR thresholds in infants born with CLP and OME and infants with OME. *International Journal of Pediatric Otorhinolaryngology*, 81, 21-25.

Tarplee, C. (1996) Working on young children's utterances: Prosodic aspects of repetition during picture labelling. In E. Couper-Kuhlen & M. Selting (Eds.), *Prosody in Conversation*. (pp. 406-435). Cambridge: Cambridge University Press.

Thomas-Stonell, N., Oddson, B., Robertson, B. & Rosenbaum, P. (2009) Predicted and observed outcomes in preschool children following speech and language treatment: Parent and clinician perspectives. *Journal of Communication Disorders*, 42:1, 29-42.

Thomas-Stonell, N., Washington, K., Oddson, B., Robertson, B. & Rosenbaum, P. (2013) Measuring communicative participation using the FOCUS[®]: Focus on the Outcomes of Communication Under Six. *Child: Care, Health and Development*, 39:4, 474-480.

Treslove, S. (2014) Principles of early intervention. In G. Phippen (Ed.), *Speech Therapy in Cleft Palate and Velopharyngeal Dysfunction*. (pp. 131-171). The Spires Cleft Centre SLT Team. Chichester: J&R Press.

Trost, J. E. (1981) Articulatory additions to the classical description of the speech of persons with cleft palate. *Cleft Palate Journal*, 18:3, 474-480.

Tykkyläinen, T. (2009) Task-setting at home and in speech and language therapy. *Child Language Teaching and Therapy*, 25:3, 319-340.

Van Riper, C. (1977) Action Therapy (videotapes). Memphis: Speech Foundation of America. Cited by Leahy, M. M. & Walsh, I. P. (2010) Paying attention to therapy discourse: Identifying therapy processes and practice in talk about talk. *Seminars in Speech and Language*, 31:2, 98-110.

Vance, M., Stackhouse, J. & Wells, B. (1998) Examen psycholinguistique des troubles du langage oral et du langage écrit chez l'enfant. Conference Proceedings of Entretiens D'orthophonie, Paris. 103-110.

Vance, M., Stackhouse, J. & Wells, B. (2005) Speech-production skills in children aged 3-7 years. *International Journal of Language & Communication Disorders*, 40:1, 29-48.

Vihman, M. & Harding-Bell, A. (2019) Infant vocalizations, babble and early speech. In A. Harding-Bell (Ed.), *Case Studies in Cleft Palate Speech: Data Analysis and Principled Intervention*. Chichester: J&R Press.

Viswanathan, N., Vidler, M. & Richard, B. (2008) Hearing thresholds in newborns with a cleft palate assessed by auditory brain stem response. *The Cleft Palate-Craniofacial Journal*, 45:2, 187-192.

Vivanti, G., Hocking, D. R., Fanning, P. & Dissanayake, C. (2016) Verbal labels increase the salience of novel objectives for pre-schoolers with typical development and Williams syndrome, but not in autism. *Journal of Neurodevelopmental Disorders*, 8:46, 1-10.

Wagner, J., Pekarek Doehler, S. & González-Martínez, E. (2018) Longitudinal research on the organization of social interaction: Current developments and methodological challenges. In S.

Pekarek Doehler, J. Wagner & E. González-Martínez (Eds.), *Longitudinal Studies on the Organization of Social Interaction*. (pp. 2-35). London: Springer Nature.

Waldron, J., Roberts, A., Latham, H. & Gibb, C. (2019) Can an early years speech pathway improve long term speech outcomes? Poster presentation at the Craniofacial Society of Great Britain and Ireland Annual Scientific Conference; London, UK.

Weismer, S. E., Venker, C. E. & Roberston, S. (2016) Focused stimulation approach to language intervention. In R. J. McCauley, M. E. Fey & R. B. Gillam (Eds.) *Treatment of Language Disorders in Children*. (2nd Edition), (pp. 146-177). Baltimore, Maryland: Brookes Publishing Co.

Weiss, A. L. (2004) The child as agent for change in therapy for phonological disorders. *Child Language Teaching and Therapy*, 20:3, 221-244.

Werker, J. F., Yeung, H. H. & Yoshida, K. A. (2012) How do infants become experts at native-speech perception? *Current Directions in Psychological Science*, 21:4, 221-226.

Widdowson, S. (2015) Multi-sensory input modelling with video-therapy in children with early signs of speech disorder. Unpublished M.Sc. University of Sheffield.

Wilkinson, R., Gower, M., Beeke, S. & Maxim, J. (2007) Adapting to conversation as a language-impaired speaker: Changes in aphasic turn construction over time. *Communication & Medicine*, 4:1, 79-97.

Willadsen, E. (2012) Influence of timing of hard palate repair in a two-stage procedure on early speech development in Danish children with cleft palate. *The Cleft Palate-Craniofacial Journal*, 49:5, 574-595.

Willadsen, E. (2013) Lexical selectivity in Danish toddlers with cleft palate. *The Cleft Palate-Craniofacial Journal*, 50:4, 456-465.

Willadsen, E. & Albrechtsen, H. (2006) Phonetic descriptions of babbling in Danish toddlers born with and without unilateral cleft lip and palate. *The Cleft Palate-Craniofacial Journal*, 43:2, 190-200.

Willadsen, E., Boers, M., Schöps, A., Kisling-Møller, M., Bøgh Nielson, J., Dahl Jørgensen, L., Andersen, M., Bolund, S. & Sjøgaard Andersen, H. (2018) Influence of timing of delayed hard palate closure on articulation skills in 3-year-old Danish children with unilateral cleft lip and palate. *International Journal of Language & Communication Disorders*, 53:1, 130-143.

Williams, L. A., McLeod, S. & McCauley, R. J. (Eds.) (2010) *Interventions for Speech Sound Disorders in Children*. Baltimore: Paul H. Brookes.

Williams, C. & van Eeden, S. (2019) Predicting the risk of speech problems. *Royal College of Speech and Language Therapists Bulletin*; June. (pp. 20-22).

Wiseman, H. and Rice, L., N. (1989) Sequential analyses of therapist-child interaction during change events: A task-focused approach. *Journal of Consulting and Clinical Psychology*, 57:2, 281-286.

Wren, Y., Harding, S., Goldbart, J. & Roulstone, S. (2018) A systematic review and classification of interventions for speech-sound disorder in preschool children. *International Journal of Language & Communication Disorders*, 53:3, 446-467.

Appendices

Appendix A: Pilot Study: Confirmation of ethics exemption

Re: Advice about ethics and consent

Derek Norfolk [Derek.Norfolk@leedsth.nhs.uk]

Sent: 15 October 2012 13:33

To: Calladine Samantha (LEEDS PCT)

Cc: Anne Gowing [Anne.Gowing@leedsth.nhs.uk]

Dear Sam,

Thank you for sending me details of your proposal.

Phase 1 of your proposal appears to fit squarely within the NHS definition of "Service Evaluation". I presume you are a member of the "direct care team" and I note that the LTH videos were obtained as part of normal clinical practice with full consent. Therefore, you do not require either Research Ethics or R&D approval for this part of your project.

It is, of course, essential to ensure that you meet all the requirements for patient confidentiality, data protection and information governance. The best way to do this is to pseudoanonymise the data you collect.

I hope this is helpful. Do you have a copy of the protocol for Phase 1 that I can keep with my records? I would also be happy to advise about the second two phases, when appropriate. I only work part-time in R&D and will next be in the office on Wednesday afternoon. If you still wish to speak to me, you can phone me on my mobile (07850518950) after 13.30 but I think your proposal is pretty clear cut.

Best wishes.

Derek Norfolk

>>> "Calladine Samantha (LEEDS PCT)" <scalladine@nhs.net> 10/13/12 14:10 PM >>>
Dear Derek

I spoke with one of your colleagues in R&D yesterday about a research proposal that I have prepared as part of my PhD, and he suggested I contact you for further advice. I would be keen to discuss my query with you by phone or in person, but thought it would help to send you some contextual information on email first.

I am a speech and language therapist and work part-time for the Yorkshire Regional Cleft Lip & Palate Service, hosted by LTHT, based at LGI. I also work part-time for the Trent Regional Cleft Service based in Nottingham. Alongside my work, I am studying for a PhD at the University of Sheffield. My PhD research will be sponsored by Nottingham University Hospitals NHS Trust (NUH), host for the Trent service.

The proposal for my PhD thesis includes three phases. The primary aims of the research are to develop and evaluate a method of early therapy intervention to treat speech sound difficulties in two year old children with repaired cleft palate. Briefly, phases one and two will use qualitative methods to define the key components of the intervention, and phase three will use qualitative and quantitative methods to evaluate the intervention.

The proposed methods and timeline for these phases are summarised below:

Phase One: Retrospective analysis of pre-existing data: October 2012 - January 2013 (pending ethics)

Phase Two: Qualitative interviews of parents and therapists: January 2013 - April 2013

Phase Three: Controlled cohort study using pre-post design: April 2013 - March 2015

For phase three, I will be recruiting participants from the Trent region and one other regional centre in the UK, but for phases one and two, I would like to recruit participants from the Yorkshire region in addition to Trent and other regions in the UK.

I am aware that NHS ethics will be required for phase three and possibly phase two, but my query at this stage is whether or not NHS ethics is needed for phase one. It is this that I would like advice on.

The plan for phase one is as follows:

To undertake retrospective analysis of pre-existing video recordings of treatment sessions obtained i) in an earlier pilot study in 2009 (property of NUH), and ii) in routine clinical practice (property of LTHT). I have a catalogue of videos that I have collated over the last two years from usual treatment sessions with children who are patients of LTHT. It is routine practice for us to make recordings and analyse the videos for clinical, training and service development purposes. We obtain written consent from parents to do this.

My question is, would I be able to analyse a sample of these videos in phase one of my research within the remit of a service evaluation, and/or on the grounds of the consent I already have, (even though I will be writing this up as part of my PhD thesis), or is NHS ethics required? Also, would I need to apply for sponsorship from LTHT R&D in addition to NUH?

I would be very grateful if you would be happy to discuss this with me by phone, and if so, if you would let me know a convenient time and number on which to call you.

With many thanks,
Sam

Samantha Calladine
Specialist SLT

Yorkshire Regional Cleft Lip & Palate Team
Room 44
F Floor
Martin Wing
Leeds General Infirmary
LS1 3EX

Telephone: 0113 39 23786 / 28558 / 07984 640734
(Children and Family Services NHS Leeds Community Healthcare Trust)
Website address: www.leedscommunityhealthcare.nhs.uk/cs1t

Appendix B: Pilot Study: Participant Information Sheet

Participant Information Sheet

(Version 1.2a Leeds)

Title of study: Service evaluation to inform the development of a therapy method to treat speech sound difficulties in two year old children with repaired cleft palate.

We would like to invite you and your child to take part in this service evaluation. Before you decide, please read the following information to help you understand why the evaluation is being done and what it would involve for you. You will have the opportunity to ask questions before you decide whether or not you want to take part.

Who is organising the evaluation?

The evaluation is being organised by Samantha Calladine, Speech and Language Therapist, with the support of:

- Yorkshire Regional Cleft Service and Leeds Teaching Hospitals NHS Trust
- Trent Regional Cleft Service and Nottingham University Hospitals NHS Trust
- University of Sheffield

The evaluation will form part of a research doctorate that Samantha is doing at the University of Sheffield under the supervision of Professor Joy Stackhouse and Dr Thomas Muskett.

What is the purpose of the study?

We are undertaking a service evaluation of speech and language therapy for young children with repaired cleft palate and related speech sound difficulties.

As part of this, we want to learn more about what happens during therapy, for example, the interaction that takes place between the therapist and the child, and how the child responds to different activities.

We hope that the findings will help us identify important aspects of therapy that might help make it effective. We then hope to develop an approach that incorporates these aspects and later evaluate it.

Why have I been invited?

You are being invited because your child had speech and language therapy when s/he was two years old and the therapist made videos of some or all of these therapy sessions as part of your child's routine care. These therapy videos are currently stored at the Yorkshire Regional Cleft Centre at Leeds General Infirmary.

Do I have to take part?

No, it is up to you whether or not you take part. If you decide to take part, you will be asked to sign a consent form. You will be given this information sheet to keep. You would be free to withdraw at any point and would not need to give a reason.

What will the study involve?

To help identify important aspects of therapy, Samantha will select one or more of your child's therapy videos, and conduct an analysis of a section of the video(s).

The analysis will include looking at the following:

- The way the therapist interacted with your child
- The way your child interacted with the therapist
- How your child responded to different things the therapist said and did
- If your child made changes to his/her speech sound production, what these were, when they occurred, what appeared to influence the change, and what happened as a result

It is anticipated that this study will last two months, finishing by 31st December 2012.

What will I need to do if I take part?

You will not need to do anything. If you decide to take part, you would be agreeing to give Samantha permission to look at and analyse therapy videos of your child that have already been collected as part of his/her routine care.

What are the possible advantages and disadvantages of taking part?

There are no specific advantages to you or your child if you take part in this study. Similarly, we do not anticipate you will experience any disadvantages or risks.

What if there are any problems?

If you have any concerns with the conduct of this study and/or the use of your child's data, please discuss them with Samantha in the first instance and she will do her best to resolve them. If you still have concerns about the conduct of the study, you can contact Professor Joy Stackhouse or Dr Thomas Muskett, Research Supervisors, or Professor Shelagh Brumfitt, Head of Department, at the University of Sheffield. If your concerns persist after this, you will be given the contact details for the university registrar. If you have persisting concerns or questions about the use of your child's data, you can contact the Information Governance department at Leeds General Infirmary. Should you wish to make a formal complaint, you can do this through the NHS Complaints service by contacting Leeds General Infirmary.

Will my taking part in this study be kept confidential?

Your child's therapy videos will be accessed in line with ethical and legal practice guidelines and handled in confidence.

During the course of the evaluation, your child's videos will be viewed at two locations only:

1. At the Yorkshire Regional Cleft Centre in Leeds, where Samantha will carry out the initial analysis, and
2. At the Department of Human Communication Sciences at the University of Sheffield, where Samantha may look at the videos with authorised persons involved in the evaluation. These individuals will have a duty of confidentiality to you as a participant in this study.

Your child's videos and information generated from the analysis will be kept strictly confidential and stored in a secure office at Yorkshire Regional Cleft Centre. Electronic information will be kept on a password protected computer. When information is taken out of the hospital, your child's name and address will be removed, and the information-carrying device will be encrypted.

What will happen to the results of this service evaluation?

The results will be used to inform the development of a therapy method to treat speech sound difficulties in two year old children with repaired cleft palate. This therapy method will then be evaluated in a bigger study. The results of this and the bigger study will be used by Samantha as part of her assessed work for her doctorate qualification.

The results are also likely to be shared with speech and language therapists in the UK who work with children with cleft-related speech sound difficulties in order to inform future clinical practice and potential further research. They might also be shared with therapists who work with children who have speech sound difficulties not related to cleft palate. They might also be shared with other professionals, for example, cleft surgeons, who work with children with cleft palate.

Neither you nor your child will be identified in any report or publication.

Who has reviewed the study?

This service evaluation has been authorised by Leeds Teaching Hospitals NHS Trust and Nottingham University Hospitals NHS Trust. The Research & Innovation departments at these Trusts have confirmed that formal review by an NHS research ethics committee is not required.

Contact details

Samantha Calladine

Yorkshire Regional Cleft Centre
Leeds General Infirmary
Leeds, LS1 3EX
0113 392 3876
0798 464 0734

scalladine@nhs.net
s.calladine@sheffield.ac.uk

Professor Joy Stackhouse Dr Thomas Muskett

The Department of Human Communication Sciences
University of Sheffield
Claremont Crescent
Sheffield, S10 2TA

0114 222 2401
0114 222 2443

j.stackhouse@sheffield.ac.uk
t.muskett@sheffield.ac.uk

Appendix C: Pilot Study: Parent/Carer Consent Form

Consent Form

(Version 1.2a Leeds)

Participant Identification Number for this study:

Title of Study: Service evaluation to inform the development of a therapy method to treat speech sound difficulties in two year old children with repaired cleft palate.

Name of Researcher: Samantha Calladine

Please initial box

1. I confirm that I have read and understand the information sheet version 1.2a Leeds dated 23rd October 2012 for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily.
2. I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason, without my medical care or legal rights being affected.
3. I understand that my child's therapy videos and data collected during the study will be accessed and looked at by the researcher and may also be looked at by named individuals from the University of Sheffield, from regulatory authorities or from the NHS Trust, where it is relevant to my taking part in this study. I give permission for these individuals to have access to this information. I understand that my child's personal details will be kept confidential.
4. I give my permission for the researcher and authorised individuals to store and publish information generated from analysing my child's therapy videos.
5. I give my permission for the researcher to use anonymised videos of my child in presentations and teaching.
6. I agree to take part in the above study.

Name of participant

Date

Signature

Name of person taking consent

Date

Signature

Samantha Calladine

Consent Form, Version 1.2a Leeds

23rd October 2012

Appendix D: IRAS (Integrated Research Ethics System) application form

NHS REC Form

Reference:
15/WS/0196

IRAS Version 5.0.0

Welcome to the Integrated Research Application System

IRAS Project Filter

The integrated dataset required for your project will be created from the answers you give to the following questions. The system will generate only those questions and sections which (a) apply to your study type and (b) are required by the bodies reviewing your study. Please ensure you answer all the questions before proceeding with your applications.

Please complete the questions in order. If you change the response to a question, please select 'Save' and review all the questions as your change may have affected subsequent questions.

Please enter a short title for this project (maximum 70 characters)

Multisensory input therapy for young children born with cleft palate

1. Is your project research?

Yes No

2. Select one category from the list below:

- Clinical trial of an investigational medicinal product
- Clinical investigation or other study of a medical device
- Combined trial of an investigational medicinal product and an investigational medical device
- Other clinical trial to study a novel intervention or randomised clinical trial to compare interventions in clinical practice
- Basic science study involving procedures with human participants
- Study administering questionnaires/interviews for quantitative analysis, or using mixed quantitative/qualitative methodology
- Study involving qualitative methods only
- Study limited to working with human tissue samples (or other human biological samples) and data (specific project only)
- Study limited to working with data (specific project only)
- Research tissue bank
- Research database

If your work does not fit any of these categories, select the option below:

Other study

2a. Please answer the following question(s):

- a) Does the study involve the use of any ionising radiation? Yes No
- b) Will you be taking new human tissue samples (or other human biological samples)? Yes No
- c) Will you be using existing human tissue samples (or other human biological samples)? Yes No

3. In which countries of the UK will the research sites be located? (Tick all that apply)

- England
- Scotland
- Wales
- Northern Ireland

Date: 25/08/2015

1

160513/837114/1/297

3a. In which country of the UK will the lead NHS R&D office be located:

- England
 Scotland
 Wales
 Northern Ireland
 This study does not involve the NHS

4. Which review bodies are you applying to?

- HRA Approval
 NHS/HSC Research and Development offices
 Social Care Research Ethics Committee
 Research Ethics Committee
 Confidentiality Advisory Group (CAG)
 National Offender Management Service (NOMS) (Prisons & Probation)

For NHS/HSC R&D offices, the CI must create Site-Specific Information Forms for each site, in addition to the study-wide forms, and transfer them to the PIs or local collaborators.

5. Will any research sites in this study be NHS organisations?

- Yes No

5a. Are all the research costs and infrastructure costs for this study provided by an NIHR Biomedical Research Centre, NIHR Biomedical Research Unit, NIHR Collaboration for Leadership in Health Research and Care (CLAHRC) or NIHR Research Centre for Patient Safety & Service Quality in all study sites?

- Yes No

If yes and you have selected HRA Approval in question 4 above, your study will be processed through HRA Approval.

If yes, and you have not selected HRA Approval in question 4 above, NHS permission for your study will be processed through the NIHR Coordinated System for gaining NHS Permission (NIHR CSP).

5b. Do you wish to make an application for the study to be considered for NIHR Clinical Research Network (CRN) support and inclusion in the NIHR Clinical Research Network (CRN) Portfolio? Please see information button for further details.

- Yes No

If yes, you must complete a NIHR Clinical Research Network (CRN) Portfolio Application Form immediately after completing this project filter and before submitting other applications. If you have selected HRA Approval in question 4 above your study will be processed through HRA Approval. If not, NHS permission for your study will be processed through the NIHR Coordinated System for gaining NHS Permission (NIHR CSP).

6. Do you plan to include any participants who are children?

- Yes No

7. Do you plan at any stage of the project to undertake intrusive research involving adults lacking capacity to consent for themselves?

- Yes No

Answer Yes if you plan to recruit living participants aged 16 or over who lack capacity, or to retain them in the study following loss of capacity. Intrusive research means any research with the living requiring consent in law. This includes use of identifiable tissue samples or personal information, except where application is being made to the Confidentiality Advisory

Group to set aside the common law duty of confidentiality in England and Wales. Please consult the guidance notes for further information on the legal frameworks for research involving adults lacking capacity in the UK.

8. Do you plan to include any participants who are prisoners or young offenders in the custody of HM Prison Service or who are offenders supervised by the probation service in England or Wales?

Yes No

9. Is the study or any part of it being undertaken as an educational project?

Yes No

Please describe briefly the involvement of the student(s):

This study is being designed and undertaken by a PhD student under the research supervision of an academic institution.

9a. Is the project being undertaken in part fulfilment of a PhD or other doctorate?

Yes No

10. Will this research be financially supported by the United States Department of Health and Human Services or any of its divisions, agencies or programs?

Yes No

11. Will identifiable patient data be accessed outside the care team without prior consent at any stage of the project (including identification of potential participants)?

Yes No

Integrated Research Application System
Application Form for Research administering questionnaires/interviews for quantitative analysis or mixed methodology study



Application to NHS/HSC Research Ethics Committee

The Chief Investigator should complete this form. Guidance on the questions is available wherever you see this symbol displayed. We recommend reading the guidance first. The complete guidance and a glossary are available by selecting [Help](#).

Please define any terms or acronyms that might not be familiar to lay reviewers of the application.

Short title and version number: (maximum 70 characters - this will be inserted as header on all forms)
 Multisensory input therapy for young children born with cleft palate

Please complete these details after you have booked the REC application for review.

REC Name:
 West of Scotland REC5

REC Reference Number:
 15/WS/0196

Submission date:
 25/08/2015

PART A: Core study information

1. ADMINISTRATIVE DETAILS

A1. Full title of the research:

Multisensory input therapy for young children with cleft related speech patterns

A2-1. Educational projects

Name and contact details of student(s):

Student 1

	Title	Forename/Initials	Surname
	Miss	Samantha	Calladine
Address	Apartment 1 Sharrow Point 326 Cemetery Road Sheffield		
Post Code	S11 8FT		
E-mail	s.calladine@sheffield.ac.uk		
Telephone	07951350434		
Fax			

Give details of the educational course or degree for which this research is being undertaken:

Name and level of course/ degree:
Ph.D. Multisensory input therapy for young children with cleft related speech patterns.

Name of educational establishment:
University of Sheffield

Name and contact details of academic supervisor(s):

Academic supervisor 1

Title Forename/Initials Surname
Dr Hilary Gardner

Address Department of Human Communication Sciences
362 Mushroom Lane
Sheffield

Post Code S10 2TS

E-mail h.gardner@sheffield.ac.uk

Telephone 01142222456

Fax 01142222439

Academic supervisor 2

Title Forename/Initials Surname
Dr Blanca Schaefer

Address Department of Human Communication Sciences
362 Mushroom Lane
Sheffield

Post Code S10 2TS

E-mail blanca.schaefer@sheffield.ac.uk

Telephone 01142222423

Fax 01142222439

Please state which academic supervisor(s) has responsibility for which student(s):
Please click "Save now" before completing this table. This will ensure that all of the student and academic supervisor details are shown correctly.

Student(s)	Academic supervisor(s)
Student 1 Miss Samantha Calladine	<input checked="" type="checkbox"/> Dr Hilary Gardner <input checked="" type="checkbox"/> Dr Blanca Schaefer

A copy of a current CV for the student and the academic supervisor (maximum 2 pages of A4) must be submitted with the application.

A2-2. Who will act as Chief Investigator for this study?

- Student
- Academic supervisor
- Other

A3-1. Chief Investigator:

	Title Forename/Initials Surname
	Miss Samantha Calladine
Post	Regional Lead Specialist Speech and Language Therapist
Qualifications	2002 B.Sc. Honours. Speech Pathology and Therapy. First class. 2009 M.Sc. Speech and Cleft Palate Studies. Distinction.
Employer	Leeds Teaching Hospitals NHS Trust
Work Address	Northern and Yorkshire Regional Cleft Lip and Palate Service (Leeds) Room 44 Paediatric Offices, F Floor Martin Wing Leeds General Infirmary
Post Code	LS1 3EX
Work E-mail	scalladine@nhs.net
* Personal E-mail	s.calladine@sheffield.ac.uk
Work Telephone	01133923786
* Personal Telephone/Mobile	07951350434
Fax	01133925116

** This information is optional. It will not be placed in the public domain or disclosed to any other third party without prior consent.
A copy of a current CV (maximum 2 pages of A4) for the Chief Investigator must be submitted with the application.*

A4. Who is the contact on behalf of the sponsor for all correspondence relating to applications for this project?
This contact will receive copies of all correspondence from REC and HRA/R&D reviewers that is sent to the CI.

	Title Forename/Initials Surname
	Ms Anne Gowing
Address	Research and Innovation 34 Hyde Terrace Leeds
Post Code	LS2 9LN
E-mail	anne.gowing@nhs.net
Telephone	01133920161
Fax	

A5-1. Research reference numbers. Please give any relevant references for your study:

Applicant's/organisation's own reference number, e.g. R & D (if available):	SP15/142
Sponsor's/protocol number:	SP15/142
Protocol Version:	1.0
Protocol Date:	23/07/2015
Funder's reference number:	NA
Project website:	NA

Additional reference number(s):

Ref.Number	Description	Reference Number

Registration of research studies is encouraged wherever possible. You may be able to register your study through your NHS organisation or a register run by a medical research charity, or publish your protocol through an open access publisher. If you have registered your study please give details in the "Additional reference number(s)" section.

A5-2. Is this application linked to a previous study or another current application?

Yes No

Please give brief details and reference numbers.

2. OVERVIEW OF THE RESEARCH

To provide all the information required by review bodies and research information systems, we ask a number of specific questions. This section invites you to give an overview using language comprehensible to lay reviewers and members of the public. Please read the guidance notes for advice on this section.

A6-1. Summary of the study. Please provide a brief summary of the research (maximum 300 words) using language easily understood by lay reviewers and members of the public. Where the research is reviewed by a REC within the UK Health Departments' Research Ethics Service, this summary will be published on the Health Research Authority (HRA) website following the ethical review. Please refer to the question specific guidance for this question.

A cleft palate affects early speech development. Before children have their palate operation they are unable to make sounds like 'b, d' and have a tendency to make more sounds at the back of their mouth and in their throat. For some children, such cleft related speech patterns disappear after the operation, but not for all.

Cleft related speech patterns can be identified from about 18-24 months of age, when speech and language therapy usually begins. Multisensory Input Therapy (MSIT) is one of the approaches used with children up until about three years of age. Speech sounds are modelled by the Speech and Language Therapist (SLT) to increase the child's awareness of different sounds. The aim is to change the child's internal store of sounds and for them to develop more accurate speech patterns. Speech production is not a focus, which is why it is particularly suitable for young children, but the therapist will give feedback if the child produces speech. Videos of therapy are often made for the child to watch at home with their family as a way of increasing opportunities for speech practice.

Little is known about how cleft related speech patterns change in children this young when they are receiving therapy. The current study will explore this by analysing videos of children receiving MSIT. Analysis of therapist behaviours will help us understand how therapy is delivered. Analysis of child speech productions will help us to understand how their patterns change. Consideration will be made of how the therapist and child interact with each other and how speech modelling and feedback influence the child's speech productions. The findings will develop the way SLTs do therapy with young children born with cleft palate.

A6-2. Summary of main issues. Please summarise the main ethical, legal, or management issues arising from your study and say how you have addressed them.

Not all studies raise significant issues. Some studies may have straightforward ethical or other issues that can be identified and managed routinely. Others may present significant issues requiring further consideration by a REC, R&D office or other review body (as appropriate to the issue). Studies that present a minimal risk to participants may raise complex organisational or legal issues. You should try to consider all the types of issues that the different reviewers may need to consider.

MANAGEMENT

1. Multi-site study. Speech and Language Therapists (SLTs) recruited to the study will not all be employed by the same NHS trust. The Lead Specialist SLT will be the primary contact for the regional SLT network. She will not be required to take on the role of principal investigator. The Lead Specialist SLT will provide information about the study to potential primary participants (SLTs in the network). The Chief Investigator (CI; Samantha Calladine) will undertake recruitment and collect the data.

2. Primary purpose of the study is educational. It will be self-funded by the student (Samantha Calladine). Limited funds are available. As the study will involve multiple site visits, this will restrict the number of SLT networks who will be contacted about the study.

3. Student and CI (Samantha Calladine) is an NHS employee. The employer organisation (Leeds Teaching Hospitals NHS Trust; LTH) will sponsor the study and manage the research in close collaboration with the academic institution

(University of Sheffield).

4.No funding available for participating trusts. The CI will take on the responsibilities that would otherwise have been delegated to principal investigators. The CI will need a Letter of Access (and informed consent from participants) to access and collect data.

ETHICAL

1.Study uses participant identifiable data (therapy videos and consultation notes). This refers to primary (patients) and secondary participants (SLTs). It will be stated clearly in the participant information sheet that the CI will have access to participant identifiable information. However, this will be limited to essential information only. For example, the CI does not need to know the home address of primary participants. For secondary participants, identifiable data will include their face, name and place of work. For primary participants, identifiable data will include their face, name and date of birth. Throughout the study, in order to protect participant identities, identification numbers will be used on any documentation generated. For dissemination, fictitious names will be used. Consent to use participant identifiable video data in dissemination will be sought as an optional level of consent. Participants who do not give their consent for information to be used in this way can still take part in the study.

2.Study uses NHS treatment records (therapy videos and consultation notes). All data will be collected by the CI. Only encrypted storage devices will be used. Procedures for data collection and storage have been developed and agreed with Research and Innovation (R&I) at LTH. These comply with the Data Protection Act 1998 and local information governance policy and good practice. Any site-specific procedures will be agreed with the R&I and Information Governance (IG) teams at participating sites.

3.Primary participants are children. Informed consent will be sought from a person with parental responsibility for the child. This person (referred to hereon as parent) will be given sufficient written and verbal information to enable them to make an informed decision and the opportunity to ask questions. The study and participant information sheet will be presented by a clinician from the child's local clinical care team. Parents will also be given the contact details of a clinician who can provide impartial advice.

4.Study involves analysis of treatment records (videos and consultation notes) made during routine clinical care. Secondary participants (SLTs) may feel uncomfortable about their therapy videos being examined. The CI will provide assurance, verbally and in writing, that a non-judgemental approach will be maintained throughout the course of the study and in dissemination of the results.

5.Presentation and publication of data. Fictitious names will be used. Videos will only be used in dissemination if participants have given their consent for this specific purpose. Any videos used for this purpose will be anonymised as much as possible, e.g. names will be edited out.

LEGAL

1.NHS treatment records are legal and confidential documents. Collection, handling and storage of research data will be undertaken in accordance with the Data Protection Act 1998.

A6-3. Proportionate review of REC application *The initial project filter has identified that your study may be suitable for proportionate review by a REC sub-committee. Please consult the current guidance notes from NRES and indicate whether you wish to apply through the proportionate review service or, taking into account your answer to A6-2, you consider there are ethical issues that require consideration at a full REC meeting.*

Yes - proportionate review No - review by full REC meeting

Further comments (optional):

The study will use pre-existing data that was generated during routine clinical practice. Consent for producing videos of clinical treatment will have been obtained from parents as part of routine NHS practice. This may or may not include consent for research.

Note: This question only applies to the REC application.

3. PURPOSE AND DESIGN OF THE RESEARCH

A7. Select the appropriate methodology description for this research. *Please tick all that apply:*

<input checked="" type="checkbox"/> Case series/ case note review <input type="checkbox"/> Case control <input type="checkbox"/> Cohort observation <input type="checkbox"/> Controlled trial without randomisation <input type="checkbox"/> Cross-sectional study <input type="checkbox"/> Database analysis <input type="checkbox"/> Epidemiology <input type="checkbox"/> Feasibility/ pilot study <input type="checkbox"/> Laboratory study <input type="checkbox"/> Metanalysis <input checked="" type="checkbox"/> Qualitative research <input checked="" type="checkbox"/> Questionnaire, interview or observation study <input type="checkbox"/> Randomised controlled trial <input type="checkbox"/> Other (please specify)

A10. What is the principal research question/objective? *Please put this in language comprehensible to a lay person.*

1. How is Multisensory Input Therapy (MSIT) delivered by Speech and Language Therapists (SLTs) to young children with cleft related speech patterns?
2. How do the speech patterns of young children born with cleft palate develop in response to MSIT?

A11. What are the secondary research questions/objectives if applicable? *Please put this in language comprehensible to a lay person.*

- What behaviours do SLTs use in MSIT?
- What activities and stimuli do SLTs use in MSIT?
- How do therapists respond to children when they make matched and mismatched productions?
- What effects do therapist behaviours have on the accuracy of children's speech productions?
- What effects do activities and stimuli have on the accuracy of children's speech productions?
- What do children do in response to their own mismatched productions, with and without input from the therapist?
- How do children's speech productions change within a session and from session to session?
- What variation is there among children and therapists in all of the above?

A12. What is the scientific justification for the research? *Please put this in language comprehensible to a lay person.*

A cleft palate affects early speech development (Chapman and Willadsen, 2011). Before children have their palate operation (at around nine to 12 months of age) they are unable to produce the same range of sounds as children without cleft palate. They can make sounds like 'm' and 'n' but cannot make sounds like 'b' and 'd'. They have a tendency to make sounds at the back of the mouth, e.g. 'g', and in the throat, e.g. 'h', rather than at the front of the mouth. For many children, such cleft related speech patterns start to disappear after their palate operation, but this is not the case for all.

Cleft related speech patterns are usually diagnosed at a child's first specialist speech and language assessment at 18-24 months of age (Bowden et al., 1997; Barratt and Extence, 2010). The Speech and Language Therapist (SLT) will make predictions about the effects the cleft palate has had on the child's speech development and why these patterns have not resolved. Some children are still not able to make some sounds because the palate is not working properly. Other children have a palate that is capable of working properly but they have not learned how to use it to make new sounds.

Speech and language therapy is usually arranged from around 18-24 months of age (Lead SLT forum, personal communication, 2015). At this stage in typical speech development, children's speech sound systems are learning to adapt to cope with growing vocabularies and longer utterances. Their internal storage of words becomes more specific, allowing them to distinguish between more words (Sutherland and Gillon, 2007). They now produce a wide range of different sounds and are starting to use them in more complicated ways. However, 'rules' are still in play that simplify how children use these sounds in the words and utterances they produce.

One of the therapy approaches used is Multisensory Input Therapy (MSIT; Harding-Bell and Bryan, 2001). It is based on a speech processing model developed by Stackhouse and Wells (1997) This theoretical model describes how children process speech they hear (input), store information about speech (internal representations) and programme the speech they produce (output). Therapy based on this is designed to target specific areas of speech processing that are contributing to the child's speech patterns. Although a number of studies have presented such therapy (e.g. Pascoe et al., 2006); Waters, 2001), as yet there has been no study that has involved young children with cleft related speech patterns.

MSIT is designed to help a child develop more accurate speech patterns by changing their internal representations of speech. The therapist models speech to the child and draws their awareness to different features of sounds, e.g. where in the mouth sounds are made. The therapist does not ask the child to produce speech but through the repeated modelling activities, the children often do so. Speech and language therapists sometimes film MSIT so that parents and carers can watch videos at home with their child. This also triggers extra opportunities for speech production practice at times when the child is most relaxed and interested. In a modified MSIT approach described by Calladine (2009) the therapist gives feedback to the child for productions they make, either to give positive reinforcement for an accurate production or opportunity to change an inaccurate production.

A number of studies have been carried out to see how children change their speech production when receiving therapy. Gardner (1994) carried out analysis of adult-child interactions and found that children were more likely to change their speech production when the adult gave a specific speech model or prompt. Other studies have looked at how children respond to different types of therapist requests for clarification, e.g. Masso et al. (2013). These types of studies have helped to develop the way we do therapy with children with speech patterns. However, again such studies have not yet involved young children with cleft related patterns.

The proposed study will help to address gaps in current knowledge about cleft related speech patterns and MSIT. It will look at how speech patterns in young children change when they are receiving MSIT Videos of children receiving therapy will be collected and analysed. Analysis of therapist behaviours will help us understand more about how MSIT is delivered. Analysis of child speech productions will help us understand more about how patterns change over time and how they might be influenced by the SLT . The findings will help to develop the way we do therapy with children.

A13. Please summarise your design and methodology. *It should be clear exactly what will happen to the research participant, how many times and in what order. Please complete this section in language comprehensible to the lay person. Do not simply reproduce or refer to the protocol. Further guidance is available in the guidance notes.*

DESIGN AND MATERIALS

The study will use a case study design. It will use pre-existing data that has been obtained as part of routine clinical care. Data will consist of:

- Videos of children with cleft related speech patterns receiving MSIT and video-therapy
- Written information from children's corresponding consultation notes:
 - o Child's cleft type
 - o Child's surgical history and presence/absence of a fistula at the time of therapy
 - o Child's age at the time of therapy
 - o Type of cleft related speech patterns the child presents with, including the findings from the child's 18-24 month assessment
 - o Target sounds being treated in the session
 - o Therapist's comments about the therapy session
 - o Therapist's transcriptions of:
 - Stimuli s/he models during the session
 - Productions made by the child

Data will be collected by the CI and PhD student, Samantha Calladine.

PARTICIPANTS

Primary participants are children born with cleft palate who received MSIT and video-therapy for cleft related speech patterns between 18 months and three years of age. Secondary participants are Speech and Language Therapists (SLTs) who provided therapy. Being a specialist in cleft palate and related disorders is not a requirement, but recruited therapists will be asked to provide information about their training, experience and role. The study aims to recruit five therapists and five children.

IDENTIFICATION, RECRUITMENT AND INFORMED CONSENT

Participants will be identified and recruited from regional SLT networks whose routine therapy provision includes MSIT

and video-therapy. The student and CI, Samantha Calladine, will contact the regional Lead Specialist SLTs and provide information about the study verbally via telephone and in writing. The Lead Specialist SLTs will disseminate this information to the regional SLT network. Secondary participants will be recruited first. Speech and language therapists will identify if they meet the participant criteria and be invited to volunteer if they wish to take part. Therapists will be given time to consider whether or not they want to take part. The CI will offer to visit the site to provide therapists with further information about the study and answer any questions. Recruited therapists will identify children who meet criteria as primary participants. They will contact the parents and provide verbal and written information about the study. They will also provide contact details for an SLT not involved in the study so that parents can seek impartial advice if they wish. Parents will be given time to consider whether or not they want to take part and the opportunity to ask questions. Informed consent will be obtained from all participants and this will be documented in writing.

DATA COLLECTION AND STORAGE

The CI will visit the NHS sites where primary participants are based to collect the videos and extract relevant information from corresponding case notes. In doing this, the CI will have access to therapist and patient identifiable information. The CI will make copies of videos and scans of relevant entries from written case notes. All data will be saved to an encrypted external hard-drive (supplied by Leeds Teaching Hospitals NHS Trust; LTHT) and an encrypted laptop on-site and taken off-site, to Leeds General Infirmary. A back-up copy of the data will be saved on the LTHT network. Access will be restricted to the CI. Data will be removed from the laptop as soon as the back-up has been made. This procedure has been agreed with the Research and Innovation (R&I) team at LTHT.

DATA ANALYSIS

Data analysis will take place off-site at three different locations: Leeds General Infirmary, University of Sheffield, and the CI's place of residence. Electronic and manual files generated during data analysis will not include participant identifiable information.

A14-1. In which aspects of the research process have you actively involved, or will you involve, patients, service users, and/or their carers, or members of the public?

- Design of the research
- Management of the research
- Undertaking the research
- Analysis of results
- Dissemination of findings
- None of the above

Give details of involvement, or if none please justify the absence of involvement.

Parents and carers of children under the care of LTHT have played an active role in identifying the research topic. In routine clinical practice, they have been encouraged to give their views about early therapy, and they have asked questions about the timing of therapy and methods used to treat cleft related speech problems. These questions cannot be answered with any certainty because of the absence of research in this field. The James Lind Alliance (2012) produced a list of 12 questions following a priority setting partnership with parents of a child born with cleft lip and/or palate, adults born with cleft lip and/or palate, and professionals working in this field. Question number four asks about the optimum age to start speech therapy. We cannot begin to answer this question until we have a better understanding of how young children respond to therapy. Whilst it was a professional obligation to extend knowledge in this area that initially prioritised the topic of the proposed research, feedback from active Patient and Public Involvement (PPI) has also been considered.

For example, when asked about the use of therapy videos, one carer of a child receiving therapy quoted the following advantages:

1. "Share his work with grandparents
2. He can watch himself do sounds and see progress
3. He can start and stop it to suit
4. Shows lots of praise and different games that parents can copy"

A PPI advisory group is being established in collaboration with the Cleft Lip and Palate Association (CLAPA). This group will play an active role in providing advice and support for relevant aspects of the proposed research. Core members of this group have already been actively involved by reviewing and commenting on the research proposal. Members of this group have contributed to the development of study materials by reviewing the participant information sheets and consent forms. Their feedback has been used to ensure these documents are written in an accessible

way using jargon-free language.

The PPI advisory group will support dissemination of the results and help put the research findings into practice. It will have a key role in assisting with the development and delivery of dissemination approaches. For example, co-producing articles and newsletters for CLAPA and the Royal College of Speech and Language Therapists, and co-presenting at local and national events. This will help keep the research grounded in the basics of patient, carer and family needs and ensure wider distribution of findings. Collaboration with parents and carers will help to ensure written summaries of the research are accessible to the wider public and raise awareness of research. It will help ensure the findings that are shared are of interest to the widest possible audiences. Dissemination to patients, the public and colleagues will also include sharing the PPI experience and informing how PPI has added value. This is expected to encourage future collaborations in health service research and related activities.

4. RISKS AND ETHICAL ISSUES

RESEARCH PARTICIPANTS

A17-1. Please list the principal inclusion criteria (list the most important, max 5000 characters).

Principal inclusion criteria for primary participants (children) are:

- Born with cleft palate +/- cleft lip
- Diagnosis of cleft related speech patterns
- Received therapy for cleft related speech patterns between 18 months and three years
- The SLT who delivered the therapy meets the secondary participant inclusion criteria
- Therapy took place in the UK and since 2010
- Therapy consisted of MSIT and video-therapy
- A minimum of five therapy records are available (from five separate sessions)
- The therapist who provided the therapy consents to take part

Principal inclusion criteria for secondary participants (SLTs) are:

- Provided therapy to a child with cleft related speech patterns aged between 18 months and three years
- Therapy took place in the UK
- Therapy took place from 2010 onwards (after dissemination of the Calladine, 2009, study)
- Therapy consisted of MSIT and video-therapy
- A minimum of five therapy records are available (from five separate sessions)
- The parent of the child who received therapy consents to take part

Being a specialist in cleft palate and related disorders is not an inclusion criterion because in the UK therapy is not always provided by a specialist. To inform the interpretation of the results, recruited participants will be asked to provide the following information:

- Complete number of years qualified as an SLT
- Whether they are a regional specialist, local specialist, or a non-specialist
- Training received on MSIT and video-therapy
- Experience of using MSIT and video-therapy

A17-2. Please list the principal exclusion criteria (list the most important, max 5000 characters).

- Video records of the therapy sessions are not available
- Video records are of poor audio and/or visual quality
- Written case note entries for the therapy sessions are not available
- Information about child's cleft type, surgical and wound healing history, presence/absence of palatal fistula and speech diagnosis is not available
- Information about therapist's previous training, experience and current role is not available

RESEARCH PROCEDURES, RISKS AND BENEFITS

A18. Give details of all non-clinical intervention(s) or procedure(s) that will be received by participants as part of the research protocol. These include seeking consent, interviews, non-clinical observations and use of questionnaires.

Please complete the columns for each intervention/procedure as follows:

1. Total number of interventions/procedures to be received by each participant as part of the research protocol.

2. If this intervention/procedure would be routinely given to participants as part of their care outside the research, how many of the total would be routine?
3. Average time taken per intervention/procedure (minutes, hours or days)
4. Details of who will conduct the intervention/procedure, and where it will take place.

Intervention or procedure	1	2	3	4
SLTs in the regional SLT network will receive the participant information sheet and consent form and will be given the opportunity to ask questions.	1	0	30	The CI will write the participant documents and the lead specialist SLT will disseminate them on email. The CI will respond to any questions about the study by telephone or email.
SLTs will be given the opportunity to meet with the CI to receive further information about the study and ask questions.	1	0	60	The CI will visit the site or therapist's base to provide further information and answer questions.
SLTs wishing to take part in the study will volunteer and provide their consent.	1	0	30	The CI will telephone the SLTs who have volunteered to take part and undertake the process of informed consent.
Parents/carers of the children who received therapy will be informed about the study by telephone.	1	0	15	The SLT who provided the child's therapy will make the telephone contact with the parent.
Parents/carers will receive the participant information sheet and consent form and will be given the opportunity to ask questions.	1	0	30	The CI will write the participant documents and the SLT who provided the child's therapy will disseminate them. The recruiting SLT or CI (parent request) will respond to any questions about the study by telephone.
Parents/carers will receive advice by telephone or email (optional).	1	0	30	The parent wishing to seek independent advice will make contact with an SLT unrelated to the study.
Parents/carers will be given the opportunity to meet with the recruiting SLT or CI (their request) to receive further information and ask questions.	1	0	30	The recruiting SLT or CI will meet with the parent at the most appropriate NHS setting to provide further information and answer questions.
Parents/carers wishing to take part will provide their consent.	1	0	30	The recruiting SLT will undertake the process of informed consent. This will take place by telephone or face to face at the most appropriate NHS setting.
Recruited SLTs will complete a one page summary of their previous experience and current role.	1	0	30	The CI will write a one page summary template and disseminate it to recruited SLTs to complete at their home or NHS base.

A21. How long do you expect each participant to be in the study in total?

Parents/carers of primary participants will not have any contact with the research team after they have given informed consent, unless they request it. Secondary participants will be required to complete a one page summary of their previous experience and role (please see attached). This is expected to take less than half an hour. There will be no contact between secondary participants and the research team after this.

A22. What are the potential risks and burdens for research participants and how will you minimise them?

For all studies, describe any potential adverse effects, pain, discomfort, distress, intrusion, inconvenience or changes to lifestyle. Only describe risks or burdens that could occur as a result of participation in the research. Say what steps would be taken to minimise risks and burdens as far as possible.

The potential risks and burdens for primary and secondary participants associated with this study are minimal. This is because the study will use pre-existing data. The only potential risks and burdens that exist are associated with access to and analysis of participant identifiable data. These will be addressed effectively by providing detailed and accurate information at the time of recruitment and implementing adequate data protection procedures. The study will not involve any prospective data collection and therefore there are no risks or burdens associated with this.

Secondary participants (SLTs) might feel uncomfortable about their therapy videos being analysed by the CI. This is

because the CI is a senior specialist SLT in the field of cleft palate and related disorders and is known for her previous research on early therapy. In providing information about the study, the CI will be clear about its aims and the type of knowledge aiming to be achieved. The CI will explain that no judgement will be made of the competency of the therapist and provide reassurance that therapy videos will be analysed in a non-judgemental way. Potential participants will be given assurance that no therapist-identifiable information will be disclosed to anyone outside of the research team during or following the study unless it is specifically requested and consented for dissemination purposes.

A23. Will interviews/ questionnaires or group discussions include topics that might be sensitive, embarrassing or upsetting, or is it possible that criminal or other disclosures requiring action could occur during the study?

Yes No

A24. What is the potential for benefit to research participants?

Primary participants (the children) who are still receiving therapy will potentially benefit from this research because the findings are expected to help develop speech and language therapy practice. Secondary participants (the SLTs) will also potentially benefit because the knowledge gained is expected to inform their future therapy practice.

A26. What are the potential risks for the researchers themselves? (if any)

There are no perceived potential risks to the CI or research team.

RECRUITMENT AND INFORMED CONSENT

In this section we ask you to describe the recruitment procedures for the study. Please give separate details for different study groups where appropriate.

A27-1. How will potential participants, records or samples be identified? Who will carry this out and what resources will be used? For example, identification may involve a disease register, computerised search of GP records, or review of medical records. Indicate whether this will be done by the direct healthcare team or by researchers acting under arrangements with the responsible care organisation(s).

- Primary participants = children who received therapy
- Secondary participants = SLTs who provided therapy

It is already known which regional SLT networks in the UK offer MSIT and video-therapy as part of their routine provision. The network that hosted Calladine's (2009) study will be contacted first: Trent Regional Cleft Lip and Palate Service. If the target sample size is not achieved, a second network will be approached.

The CI will contact the regional Lead Specialist SLT with information about the study. The Lead Specialist SLT will disseminate this within the regional network. Secondary participants (SLTs) will be approached first. They will be asked to identify if they think they are eligible and wish to be considered as a potential participant. It is possible that SLTs will need to review patient records to check they may be eligible, but as they are in the local or regional direct healthcare team, access will already be in place. Any potential participants wishing to speak with the CI and ask further questions about the study will be given the opportunity to do so. This can be done via email or telephone if the participant wants to remain anonymous. Therapists who want to take part will volunteer. The CI will check eligibility and undertake recruitment.

SLTs recruited as secondary participants will identify potential primary participants. This will involve a review of patient records (by SLTs in the direct healthcare team). These therapists will check eligibility and undertake recruitment.

A27-2. Will the identification of potential participants involve reviewing or screening the identifiable personal information of patients, service users or any other person?

Yes No

Please give details below:

Identification of potential participants will involve reviewing the following:

- Patient database held by the regional SLT service
- Patient databases held by the local SLT service
- Speech and language therapy records of children under the care of the regional SLT service
- Speech and language therapy records of children under the care of the local SLT service

This will be undertaken by speech and language therapists in the direct professional network and/or healthcare team who have routine access to this information.

A27-4. Will researchers or individuals other than the direct care team have access to identifiable personal information of any potential participants?

Yes No

A27-5. Has prior consent been obtained or will it be obtained for access to identifiable personal information?

Yes No

If Yes, please give details below.

Informed consent will be obtained from all participants to allow the CI access to identifiable information. This will be documented in writing.

A28. Will any participants be recruited by publicity through posters, leaflets, adverts or websites?

Yes No

A29. How and by whom will potential participants first be approached?

Potential secondary participants will be approached first. The Lead Specialist SLT in the regional professional network and healthcare team will make the initial contact. The Lead Specialist SLT might herself be a potential participant but she will not be expected to inform the CI unless she volunteers to take part. Potential participants will be given the opportunity to discuss the study with the CI by telephone and/or email. If requested, an informal face to face meeting will be arranged. Potential secondary participants will not be asked to disclose their identity to the CI until they have volunteered to take part.

Primary participants are children. Parents of potential participants will be approached by the SLT who provided their child's therapy, and who is in the direct healthcare team.

The CI will make it clear that strictly no undue influence is to be used when approaching participants to encourage them to take part.

A30-1. Will you obtain informed consent from or on behalf of research participants?

Yes No

If you will be obtaining consent from adult participants, please give details of who will take consent and how it will be done, with details of any steps to provide information (a written information sheet, videos, or interactive material). Arrangements for adults unable to consent for themselves should be described separately in Part B Section 6, and for children in Part B Section 7.

If you plan to seek informed consent from vulnerable groups, say how you will ensure that consent is voluntary and fully informed.

The CI will obtain informed consent from primary participants. SLTs recruited as secondary participants will obtain informed consent from parents of primary participants.

The CI has completed the National Institute for Health Research (NIHR) Introduction to Good Clinical Practice (GCP) training, which included the process of informed consent.

The CI and R&I team at LTHT will do the following to ensure all SLTs involved in the consent process are competent to perform this role:

1. Encourage SLTs to complete the NIHR GCP training (but this will not be mandatory)
2. Explain the process of informed consent
3. Provide guidance on how to obtain consent, e.g. give examples of what to say
4. Provide links to relevant information and guidance on the NIHR website

If you are not obtaining consent, please explain why not.

Not applicable.

Please enclose a copy of the information sheet(s) and consent form(s).

A30-2. Will you record informed consent (or advice from consultees) in writing?

Yes No

A31. How long will you allow potential participants to decide whether or not to take part?

Participants will be given fourteen calendar days. Stamped addressed envelopes will be provided for participants wishing to return their consent form by post.

A33-1. What arrangements have been made for persons who might not adequately understand verbal explanations or written information given in English, or who have special communication needs? (e.g. translation, use of interpreters)

Secondary participants are UK SLTs who will have adequate understanding of English. Therapists recruiting parents of primary participants will be asked to indicate if special arrangements need to be made. LTHT will provide the facilities for translation of written documents, if required. Interpreters will be arranged via the local Trust and funded by LTHT.

A35. What steps would you take if a participant, who has given informed consent, loses capacity to consent during the study? Tick one option only.

- The participant and all identifiable data or tissue collected would be withdrawn from the study. Data or tissue which is not identifiable to the research team may be retained.
- The participant would be withdrawn from the study. Identifiable data or tissue already collected with consent would be retained and used in the study. No further data or tissue would be collected or any other research procedures carried out on or in relation to the participant.
- The participant would continue to be included in the study.
- Not applicable – informed consent will not be sought from any participants in this research.
- Not applicable – it is not practicable for the research team to monitor capacity and continued capacity will be assumed.

Further details:

The study will use pre-existing data and no contact will be made with participants after recruitment. If, however, it is brought to the CI's attention during or after the study that a participant has lost capacity to consent, their data WILL be withdrawn from the study.

CONFIDENTIALITY

In this section, personal data means any data relating to a participant who could potentially be identified. It includes pseudonymised data capable of being linked to a participant through a unique code number.

Storage and use of personal data during the study

A36. Will you be undertaking any of the following activities at any stage (including in the identification of potential participants)? (Tick as appropriate)

- Access to medical records by those outside the direct healthcare team
- Electronic transfer by magnetic or optical media, email or computer networks
- Sharing of personal data with other organisations
- Export of personal data outside the EEA
- Use of personal addresses, postcodes, faxes, emails or telephone numbers
- Publication of direct quotations from respondents
- Publication of data that might allow identification of individuals
- Use of audio/visual recording devices
- Storage of personal data on any of the following:
 - Manual files including X-rays
 - NHS computers
 - Home or other personal computers
 - University computers
 - Private company computers
 - Laptop computers

Further details:

Access to medical records will only be undertaken with consent and consent will be documented.

For data collection, participant identifiable data will be transferred electronically using encrypted devices provided by Leeds Teaching Hospitals NSH Trust (LTHT): an encrypted external hard-drive (DataLocker Enterprise) and an encrypted laptop. Only the CI can gain access to these devices. A back-up of the data will be saved on the Trust's secure network/server at Leeds General Infirmary. Access will be restricted to the CI only. Participant identifiable data will be removed from the laptop as soon as the server back-up has been made. These procedures comply with the Data Protection Act 1998 and local information governance practice.

The CI will have access to the work contact details of SLTs in order to undertake recruitment, but only when they have volunteered to take part. The CI will have access to the personal details of primary participants, but only after they have been recruited and consent has been obtained.

Any direct quotations made by participants will be anonymised in any publications.

During the study, participant identifiable data will be stored electronically on the DataLocker encrypted external hard-drive and the LTHT network only. Electronic and manual files generated during data analysis will be anonymised, except for video samples/edits. Electronic files will be stored on encrypted portable devices (external hard-drive and LTHT laptop). Manual files will be stored in a locked filing cabinet in the CI's personal office at Leeds General Infirmary. No data will be stored on personal computers or laptops. For analysis and write-up, files will be transported between Leeds General Infirmary, the University of Sheffield, and the CI's place of residence.

A38. How will you ensure the confidentiality of personal data? Please provide a general statement of the policy and procedures for ensuring confidentiality, e.g. anonymisation or pseudonymisation of data.

Confidentiality of personal data will be maintained throughout and following the study by:

1. Using only encrypted electronic devices and a secured server
2. Restricting access to the research team
3. Anonymising electronic and manual files using participant identification numbers*
4. Anonymising data in any publications*

*Video data files will not be anonymised for data analysis, but they will only exist in electronic form and access will be restricted, as described above. No videos or images will be used in publications or other forms of dissemination without consent from participants. If consent is obtained, videos will be anonymised as far as is possible. For example, names will be edited out and any overt identifiable information will be distorted using appropriate editing software.

A40. Who will have access to participants' personal data during the study? Where access is by individuals outside the direct care team, please justify and say whether consent will be sought.

The CI will have access to secondary participants' personal data from the time of recruitment, only after they have volunteered to be considered for participation. The CI will have access to primary participants' personal data from the time of data collection, only after parents have given consent and local Research and Development teams have provided a Letter of Access.

As this is a student project, the CI's educational supervisors will also have access to video data, which could identify participants, but only when necessary to support completion of the research. This is stated clearly in the participant information sheets and consent forms.

The method for data analysis will include an assessment of reliability. This requires the involvement of a second person, unrelated to the research, who will need access to a small sample of the video data. This person will be one of the CI's research colleagues at the University of Sheffield. No other information about the child or therapist will be provided. This is stated clearly in the participant information sheets and consent forms.

Storage and use of data after the end of the study

A43. How long will personal data be stored or accessed after the study has ended?

- Less than 3 months
 3 – 6 months
 6 – 12 months
 12 months – 3 years
 Over 3 years

If longer than 12 months, please justify:

Samantha is expected to complete her PhD in March 2018. Upon completion, personal data will be removed from the encrypted external hard-drive and encrypted laptop. One copy will be retained on the secure network/server at LTHT so it can be accessed for dissemination purposes and/or verification of the results. Please see sections A44/A45 for further details.

INCENTIVES AND PAYMENTS

A46. Will research participants receive any payments, reimbursement of expenses or any other benefits or incentives for taking part in this research?

- Yes No

A47. Will individual researchers receive any personal payment over and above normal salary, or any other benefits or incentives, for taking part in this research?

- Yes No

A48. Does the Chief Investigator or any other investigator/collaborator have any direct personal involvement (e.g. financial, share holding, personal relationship etc.) in the organisations sponsoring or funding the research that may give rise to a possible conflict of interest?

- Yes No

NOTIFICATION OF OTHER PROFESSIONALS

A49-1. Will you inform the participants' General Practitioners (and/or any other health or care professional responsible for their care) that they are taking part in the study?

Yes No

If Yes, please enclose a copy of the information sheet/letter for the GP/health professional with a version number and date.

PUBLICATION AND DISSEMINATION

A50. Will the research be registered on a public database?

Yes No

Please give details, or justify if not registering the research.
No suitable register exists.

Registration of research studies is encouraged wherever possible.
You may be able to register your study through your NHS organisation or a register run by a medical research charity, or publish your protocol through an open access publisher. If you are aware of a suitable register or other method of publication, please give details. If not, you may indicate that no suitable register exists. Please ensure that you have entered registry reference number(s) in question A5-1.

A51. How do you intend to report and disseminate the results of the study? Tick as appropriate:

- Peer reviewed scientific journals
 Internal report
 Conference presentation
 Publication on website
 Other publication
 Submission to regulatory authorities
 Access to raw data and right to publish freely by all investigators in study or by Independent Steering Committee on behalf of all investigators
 No plans to report or disseminate the results
 Other (please specify)

A53. Will you inform participants of the results?

Yes No

Please give details of how you will inform participants or justify if not doing so.
The CI will produce two written summaries of the results, one for parents/carers and one for SLTs. These will include details of planned dissemination activities so that participants can access these if they wish.

5. Scientific and Statistical Review

A54. How has the scientific quality of the research been assessed? Tick as appropriate:

- Independent external review
 Review within a company
 Review within a multi-centre research group
 Review within the Chief Investigator's institution or host organisation
 Review within the research team
 Review by educational supervisor
 Other

Justify and describe the review process and outcome. If the review has been undertaken but not seen by the

researcher, give details of the body which has undertaken the review:

The study has been reviewed by Professor Joy Stackhouse and Dr Hilary Gardner at the University of Sheffield. Please see report attached. Professor Stackhouse is Samantha's primary education supervisor. She is due to retire in September 2015 when Dr Gardner will take over this position.

For all studies except non-doctoral student research, please enclose a copy of any available scientific critique reports, together with any related correspondence.

For non-doctoral student research, please enclose a copy of the assessment from your educational supervisor/ institution.

A56. How have the statistical aspects of the research been reviewed? Tick as appropriate:

- Review by independent statistician commissioned by funder or sponsor
- Other review by independent statistician
- Review by company statistician
- Review by a statistician within the Chief Investigator's institution
- Review by a statistician within the research team or multi-centre group
- Review by educational supervisor
- Other review by individual with relevant statistical expertise
- No review necessary as only frequencies and associations will be assessed – details of statistical input not required

In all cases please give details below of the individual responsible for reviewing the statistical aspects. If advice has been provided in confidence, give details of the department and institution concerned.

	Title	Forename/Initials	Surname
	Dr	Hilary	Gardner
Department	Human Communication Sciences		
Institution	University of Sheffield		
Work Address	362 Mushroom Lane Sheffield		
Post Code	S10 2TS		
Telephone	01142222456		
Fax	01142222419		
Mobile			
E-mail	h.gardner@sheffield.ac.uk		

Please enclose a copy of any available comments or reports from a statistician.

A57. What is the primary outcome measure for the study?

Outcome measures do not apply to this study.

A58. What are the secondary outcome measures? (if any)

Not applicable.

A59. What is the sample size for the research? How many participants/samples/data records do you plan to study in total? If there is more than one group, please give further details below.

Total UK sample size: 10
Total international sample size (including UK): 10

Total in European Economic Area: 0

Further details:
The study aims to recruit five children (primary participants) and five SLTs (secondary participants) who provided the therapy. As the aim of the study is to explore how speech patterns develop in response to therapy, longitudinal data sets are required. Based on a pilot study involving one therapist and two children, the minimum number of therapy records required per child is five (Calladine, 2014).

A60. How was the sample size decided upon? *If a formal sample size calculation was used, indicate how this was done, giving sufficient information to justify and reproduce the calculation.*

Research in this area is at a level where case study designs involving small numbers are required. As this is the design that will be used in the current study, the chosen sample size is similar to that used in other case studies in the speech and language therapy literature.

A61. Will participants be allocated to groups at random?

Yes No

A62. Please describe the methods of analysis (statistical or other appropriate methods, e.g. for qualitative research) by which the data will be evaluated to meet the study objectives.

Appropriate methods of analysis will be adopted based on other studies in the speech and language therapy literature and psychotherapy literature. Given the exploratory nature of the study, methods will be developed and refined throughout the analysis to respond to the data as it unfolds.

Principal question 1: How is MSIT delivered by SLTs to young children with cleft related speech patterns?
METHOD OF ANALYSIS: Analysis of the therapy process via therapist behaviour in the videos.

Principal question 2: How do the speech patterns of young children born with cleft palate develop in response to MSIT?
METHOD OF ANALYSIS: Analysis of child speech behaviour in the videos, interpreted within an interactional framework and in relation to Stackhouse and Wells' (1997) speech processing model.

A random sample of videos will be analysed by one of the CI's research colleagues in the Department of Human Communication Sciences in order to provide a measure of reliability. Appropriate statistical tests will be used, e.g. Kappa statistic.

6. MANAGEMENT OF THE RESEARCH

A63. Other key investigators/collaborators. *Please include all grant co-applicants, protocol co-authors and other key members of the Chief Investigator's team, including non-doctoral student researchers.*

A64. Details of research sponsor(s)

A64-1. Sponsor

Lead Sponsor

Status: NHS or HSC care organisation
 Academic
 Pharmaceutical industry
 Medical device industry
 Local Authority

Commercial status: Non-Commercial
 Commercial

- Other social care provider (including voluntary sector or private organisation)
- Other

If Other, please specify:

Contact person

Name of organisation Leeds Teaching Hospitals NHS Trust

Given name Anne

Family name Gowing

Address Research and Innovation

Town/city 34 Hyde Terrace

Post code LS2 9LN

Country UNITED KINGDOM

Telephone 01133920161

Fax

E-mail anne.gowing@nhs.net

Is the sponsor based outside the UK?

- Yes No

Under the Research Governance Framework for Health and Social Care, a sponsor outside the UK must appoint a legal representative established in the UK. Please consult the guidance notes.

A65. Has external funding for the research been secured?

- Funding secured from one or more funders
- External funding application to one or more funders in progress
- No application for external funding will be made

What type of research project is this?

- Standalone project
- Project that is part of a programme grant
- Project that is part of a Centre grant
- Project that is part of a fellowship/ personal award/ research training award
- Other

Other – please state:

A67. Has this or a similar application been previously rejected by a Research Ethics Committee in the UK or another country?

- Yes No

Please provide a copy of the unfavourable opinion letter(s). You should explain in your answer to question A6-2 how the reasons for the unfavourable opinion have been addressed in this application.

A68-1. Give details of the lead NHS R&D contact for this research:

	Title Forename/Initials Surname Ms Anne Gowing
Organisation	Leeds Teaching Hospitals NHS Trust
Address	Research and Innovation 34 Hyde Terrace Leeds
Post Code	LS2 9LN
Work Email	anne.gowing@nhs.net
Telephone	01133920161
Fax	
Mobile	

Details can be obtained from the NHS R&D Forum website: <http://www.rdforum.nhs.uk>

A69-1. How long do you expect the study to last in the UK?

Planned start date: 01/10/2015
 Planned end date: 31/03/2018
 Total duration:
 Years: 2 Months: 5 Days: 31

A71-2. Where will the research take place? (Tick as appropriate)

England
 Scotland
 Wales
 Northern Ireland
 Other countries in European Economic Area

Total UK sites in study 3

Does this trial involve countries outside the EU?
 Yes No

A72. What host organisations (NHS or other) in the UK will be responsible for the research sites? Please indicate the type of organisation by ticking the box and give approximate numbers of planned research sites:

<input type="checkbox"/> NHS organisations in England	3
<input type="checkbox"/> NHS organisations in Wales	0
<input type="checkbox"/> NHS organisations in Scotland	0
<input type="checkbox"/> HSC organisations in Northern Ireland	0
<input type="checkbox"/> GP practices in England	0
<input type="checkbox"/> GP practices in Wales	0
<input type="checkbox"/> GP practices in Scotland	0
<input type="checkbox"/> GP practices in Northern Ireland	0
<input type="checkbox"/> Social care organisations	0
<input type="checkbox"/> Phase 1 trial units	0
<input type="checkbox"/> Prison establishments	0
<input type="checkbox"/> Probation areas	0

<input type="checkbox"/> Independent hospitals	0
<input type="checkbox"/> Educational establishments	0
<input type="checkbox"/> Independent research units	0
<input type="checkbox"/> Other (give details)	0
Total UK sites in study:	3

A76. Insurance/ indemnity to meet potential legal liabilities

Note: in this question to NHS indemnity schemes include equivalent schemes provided by Health and Social Care (HSC) in Northern Ireland

A76-1. What arrangements will be made for insurance and/or indemnity to meet the potential legal liability of the sponsor(s) for harm to participants arising from the management of the research? Please tick box(es) as applicable.

Note: Where a NHS organisation has agreed to act as sponsor or co-sponsor, indemnity is provided through NHS schemes. Indicate if this applies (there is no need to provide documentary evidence). For all other sponsors, please describe the arrangements and provide evidence.

- NHS indemnity scheme will apply (NHS sponsors only)
 Other insurance or indemnity arrangements will apply (give details below)

Please enclose a copy of relevant documents.

A76-2. What arrangements will be made for insurance and/ or indemnity to meet the potential legal liability of the sponsor(s) or employer(s) for harm to participants arising from the design of the research? Please tick box(es) as applicable.

Note: Where researchers with substantive NHS employment contracts have designed the research, indemnity is provided through NHS schemes. Indicate if this applies (there is no need to provide documentary evidence). For other protocol authors (e.g. company employees, university members), please describe the arrangements and provide evidence.

- NHS indemnity scheme will apply (protocol authors with NHS contracts only)
 Other insurance or indemnity arrangements will apply (give details below)

Please enclose a copy of relevant documents.

A76-3. What arrangements will be made for insurance and/ or indemnity to meet the potential legal liability of investigators/collaborators arising from harm to participants in the conduct of the research?

Note: Where the participants are NHS patients, indemnity is provided through the NHS schemes or through professional indemnity. Indicate if this applies to the whole study (there is no need to provide documentary evidence). Where non-NHS sites are to be included in the research, including private practices, please describe the arrangements which will be made at these sites and provide evidence.

- NHS indemnity scheme or professional indemnity will apply (participants recruited at NHS sites only)
 Research includes non-NHS sites (give details of insurance/ indemnity arrangements for these sites below)

Please enclose a copy of relevant documents.

PART B: Section 7 - Children**1. Please specify the potential age range of children under 16 who will be included and give reasons for carrying out the research in this age group.**

The purpose of this study is to explore the SLT process for young children (aged 18 months to three years) with cleft related speech patterns and how their patterns develop in response to therapy. Data on children is therefore required. The study carries relatively low ethical concerns because it will use pre-existing data that was produced as part of children's routine clinical care. In the present day (at the time the study will be conducted), participants will still be classed as children because they will not yet have reached 16 years of age.

2. Indicate whether any children under 16 will be recruited as controls and give further details.

Not applicable.

3-2. Please describe the arrangements for seeking informed consent from a person with parental responsibility and/or from children able to give consent for themselves.

SLTs recruited as secondary participants will identify children who meet criteria as primary participants. Therapists will contact the person who has parental responsibility for the children and introduce the study. It will be the role of the therapists to obtain informed consent, as they are from the child's direct/local healthcare team. The CI will provide therapists with participant information sheets and consent forms. The CI will explain the process of informed consent and ensure therapists are confident and competent to perform this role. The participant information sheet includes the contact details of the Lead Specialist SLT in the child's regional healthcare team so that parents can seek further advice if they wish. If the Lead Specialist SLT is involved in the recruitment, an alternative contact will be provided.

4. If you intend to provide children under 16 with information about the research and seek their consent or agreement, please outline how this process will vary according to their age and level of understanding.

Not applicable.

Copies of written information sheet(s) for parents and children, consent/assent form(s) and any other explanatory material should be enclosed with the application.

PART C: Overview of research sites

Please enter details of the host organisations (Local Authority, NHS or other) in the UK that will be responsible for the research sites. For NHS sites, the host organisation is the Trust or Health Board. Where the research site is a primary care site, e.g. GP practice, please insert the host organisation (PCT or Health Board) in the Institution row and insert the research site (e.g. GP practice) in the Department row.

Research site		Investigator/ Collaborator/ Contact	
Institution name	Nottingham University Hospitals NHS Trust	Title	Ms
Department name	Trent Regional Cleft Lip and Palate Service	First name/ Initials	Lorraine
Street address	Hucknall Road	Surname	Britton
Town/city	Nottingham		
Post Code	NG5 1PB		
Institution name	Salisbury NHS Foundation Trust	Title	Dr
Department name	The Spires Cleft Centre	First name/ Initials	Ginette
Street address	Odstock Road	Surname	Phippen
Town/city	Salisbury		
Post Code	SP2 8BJ		
Institution name	Cambridge University Hospitals NHS FoundationTrust	Title	Mrs
Department name	East of England Cleft Lip and Palate Network	First name/ Initials	Caroline
Street address	Hills Road	Surname	Hattee
Town/city	Cambridge		
Post Code	CB2 0QQ		

PART D: Declarations**D1. Declaration by Chief Investigator**

1. The information in this form is accurate to the best of my knowledge and belief and I take full responsibility for it.
2. I undertake to abide by the ethical principles underlying the Declaration of Helsinki and good practice guidelines on the proper conduct of research.
3. If the research is approved I undertake to adhere to the study protocol, the terms of the full application as approved and any conditions set out by review bodies in giving approval.
4. I undertake to notify review bodies of substantial amendments to the protocol or the terms of the approved application, and to seek a favourable opinion from the main REC before implementing the amendment.
5. I undertake to submit annual progress reports setting out the progress of the research, as required by review bodies.
6. I am aware of my responsibility to be up to date and comply with the requirements of the law and relevant guidelines relating to security and confidentiality of patient or other personal data, including the need to register when necessary with the appropriate Data Protection Officer. I understand that I am not permitted to disclose identifiable data to third parties unless the disclosure has the consent of the data subject or, in the case of patient data in England and Wales, the disclosure is covered by the terms of an approval under Section 251 of the NHS Act 2006.
7. I understand that research records/data may be subject to inspection by review bodies for audit purposes if required.
8. I understand that any personal data in this application will be held by review bodies and their operational managers and that this will be managed according to the principles established in the Data Protection Act 1998.
9. I understand that the information contained in this application, any supporting documentation and all correspondence with review bodies or their operational managers relating to the application:
 - Will be held by the REC (where applicable) until at least 3 years after the end of the study; and by NHS R&D offices (where the research requires NHS management permission) in accordance with the NHS Code of Practice on Records Management.
 - May be disclosed to the operational managers of review bodies, or the appointing authority for the REC (where applicable), in order to check that the application has been processed correctly or to investigate any complaint.
 - May be seen by auditors appointed to undertake accreditation of RECs (where applicable).
 - Will be subject to the provisions of the Freedom of Information Acts and may be disclosed in response to requests made under the Acts except where statutory exemptions apply.
 - May be sent by email to REC members.
10. I understand that information relating to this research, including the contact details on this application, may be held on national research information systems, and that this will be managed according to the principles established in the Data Protection Act 1998.
11. Where the research is reviewed by a REC within the UK Health Departments Research Ethics Service, I understand that the summary of this study will be published on the website of the National Research Ethics Service (NRES), together with the contact point for enquiries named below. Publication will take place no earlier than 3 months after issue of the ethics committee's final opinion or the withdrawal of the application.

Contact point for publication*(Not applicable for R&D Forms)*

NRES would like to include a contact point with the published summary of the study for those wishing to seek further information. We would be grateful if you would indicate one of the contact points below.

- Chief Investigator
 Sponsor

- Study co-ordinator
- Student
- Other – please give details
- None

Access to application for training purposes *(Not applicable for R&D Forms)*

Optional – please tick as appropriate:

I would be content for members of other RECs to have access to the information in the application in confidence for training purposes. All personal identifiers and references to sponsors, funders and research units would be removed.

This section was signed electronically by Samantha Calladine on 20/08/2015 13:15.

Job Title/Post: PhD Student / Lead Specialist Speech and Language Therapist
Organisation: University of Sheffield / Leeds Teaching Hospitals NHS Trust
Email: s.calladine@sheffield.ac.uk

D2. Declaration by the sponsor's representative

If there is more than one sponsor, this declaration should be signed on behalf of the co-sponsors by a representative of the lead sponsor named at A64-1.

I confirm that:

1. This research proposal has been discussed with the Chief Investigator and agreement in principle to sponsor the research is in place.
2. An appropriate process of scientific critique has demonstrated that this research proposal is worthwhile and of high scientific quality.
3. Any necessary indemnity or insurance arrangements, as described in question A76, will be in place before this research starts. Insurance or indemnity policies will be renewed for the duration of the study where necessary.
4. Arrangements will be in place before the study starts for the research team to access resources and support to deliver the research as proposed.
5. Arrangements to allocate responsibilities for the management, monitoring and reporting of the research will be in place before the research starts.
6. The duties of sponsors set out in the Research Governance Framework for Health and Social Care will be undertaken in relation to this research.

Please note: The declarations below do not form part of the application for approval above. They will not be considered by the Research Ethics Committee.

7. Where the research is reviewed by a REC within the UK Health Departments Research Ethics Service, I understand that the summary of this study will be published on the website of the National Research Ethics Service (NRES), together with the contact point for enquiries named in this application. Publication will take place no earlier than 3 months after issue of the ethics committee's final opinion or the withdrawal of the application.
8. Specifically, for submissions to the Research Ethics Committees (RECs) I declare that any and all clinical trials approved by the HRA since 30th September 2013 (as defined on IRAS categories as clinical trials of medicines, devices, combination of medicines and devices or other clinical trials) have been registered on a publicly accessible register in compliance with the HRA registration requirements for the UK, or that any deferral granted by the HRA still applies.

This section was signed electronically by Dr Derek Norfolk on 20/08/2015 19:45.

Job Title/Post: Associate R&D Director
Organisation: Leeds Teaching Hospitals NHS Trust
Email: derek.norfolk@nhs.net

D3. Declaration for student projects by academic supervisor(s)

1. I have read and approved both the research proposal and this application. I am satisfied that the scientific content of the research is satisfactory for an educational qualification at this level.
2. I undertake to fulfil the responsibilities of the supervisor for this study as set out in the Research Governance Framework for Health and Social Care.
3. I take responsibility for ensuring that this study is conducted in accordance with the ethical principles underlying the Declaration of Helsinki and good practice guidelines on the proper conduct of research, in conjunction with clinical supervisors as appropriate.
4. I take responsibility for ensuring that the applicant is up to date and complies with the requirements of the law and relevant guidelines relating to security and confidentiality of patient and other personal data, in conjunction with clinical supervisors as appropriate.

Academic supervisor 1

This section was signed electronically by Dr Blanca Schaefer on 20/08/2015 13:30.

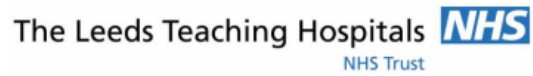
Job Title/Post: lecturer
Organisation: University of Sheffield
Email: blanca.schaefer@sheffield.ac.uk

Academic supervisor 2

This section was signed electronically by Dr Hilary Gardner on 20/08/2015 14:32.

Job Title/Post:
Organisation:
Email:

Appendix E: Research Protocol



RESEARCH PROTOCOL

TITLE OF STUDY:

Multisensory input therapy for young children with cleft related speech patterns

VERSION: 1.0

DATE: 15th August 2015

PROTOCOL NUMBER: SP15/142

CHIEF INVESTIGATOR: Samantha Calladine

This is an educational project being undertaken as part of Samantha's PhD at the
University of Sheffield

SPONSOR: Leeds Teaching Hospitals NHS Trust

CONTENTS

	Page number
1.0 Introduction	3
2.0 Design	4
3.0 Research Questions	5
3.1 Principal	5
3.2 Secondary	5
4.0 Population and sample size	6-7
4.1 Primary Participants	6
4.1.1 Inclusion criteria	6
4.2 Secondary participants	6
4.2.1 Inclusion criteria	6-7
4.3 Exclusions	7
5.0 Method and procedures	8-9
5.1 Materials	8
5.2 Data collection	8
5.3 Data analysis	8-9
6.0 Ethical considerations	10-11
6.1 Informed consent	10
6.2 Confidentiality	10-11
7.0 Publication	12
8.0 Financial arrangements	13
9.0 References	14
10.0 Contacts	15
11.0 Appendix	16
11.1 Therapist Data Collection Form Version 1.0	
11.2 Child Data Collection Form Version 1.0	
11.3 Data Collection Log Version 1.0	
11.4 Parent/Carer Information Sheet Version 1.0	
11.5 Therapist Information Sheet Version 1.0	
11.6 Parent/Carer Consent Form Version 1.0	
11.7 Therapist Consent Form Version 1.0	
11.8 Parent/Carer Consent Form Videos Version 1.0	
11.9 Therapist Consent Form Videos Version 1.0	

1.0 INTRODUCTION

A cleft palate affects speech development (Chapman and Willadsen, 2011). Before children have their palate operation they are unable to make sounds like 'b' and 'd.' Children make sounds like 'm' and 'n' and have a tendency to make sounds at the back of their mouth behind the cleft, e.g. 'g,' and in their throat, e.g. 'h.' Such cleft related speech patterns start to disappear after the palate operation. However, this does not happen for all children; for some, cleft related patterns persist and affect their later speech development.

Cleft related speech patterns are identified at about 18 months of age and therapy is usually arranged (Barratt and Extence, 2010; Lead Speech and Language Therapist (SLT) Forum, 2015). Multisensory Input Therapy (MSIT) is one of the approaches used (Harding-Bell and Bryan, 2001, Lead SLT Forum, 2015). It is based on a speech processing model developed by Stackhouse and Wells (1997). The therapist models speech stimuli to the child to help them become aware of different sounds and how they are made. The therapist does not ask the child to produce speech but creates opportunities for production practice. Feedback is given to the child about any spontaneous productions they make (Calladine, 2009). The aim of therapy is to develop more accurate speech patterns by changing the child's internal representations of speech. Speech and language therapists sometimes film therapy (video-therapy) so that parents/carers can watch videos at home to provide extra opportunities for modelling and practice.

There has been very little research into therapy for young children with cleft related speech patterns (Bessell et al., 2013). There are uncertainties about the way MSIT is delivered, how therapy interacts with children's developing speech and language systems at this age, and how children's speech patterns change during MSIT. The current study aims to address these gaps in knowledge by investigating the therapy process.

Videos of children receiving MSIT will be collected and analysed. Videos have already been collected as part of children's routine clinical care. Analysis of therapist behaviours will provide insight into how MSIT is delivered. Analysis of child speech behaviours in relation to Stackhouse and Wells' (1997) speech processing model will provide insight into how children's speech patterns develop in response to MSIT. The findings will help to develop the way we do therapy with young children born with cleft palate.

2.0 DESIGN

This is an exploratory and innovative study. Similar research has not been done before with children with cleft related speech patterns at this age. In line with the exploratory stage of this research, the study will use a case study design.

3.0 RESEARCH QUESTIONS

3.1 Principal research questions

1. How is Multisensory Input Therapy (MSIT) delivered by Speech and Language Therapists (SLTs) to young children with cleft related speech patterns?
2. How do the speech patterns of young children born with cleft palate develop in response to MSIT?

3.2 Secondary research questions

- What behaviours do SLTs use in MSIT?
- What activities and stimuli do SLTs use in MSIT?
- How do therapists respond to children when they make matched and mismatched productions?
- What effects do therapist behaviours have on the types of child behaviour?
- What effects do therapist behaviours have on (the accuracy of) children's speech productions?
- What effects do activities and stimuli have on (the accuracy of) children's speech productions?
- What do children do in reaction to their own mismatched productions, with and without input from the therapist?
- How do children's speech productions change within a session and from session to session?
- What variation is there among children and therapists in all of the above?

4.0 POPULATION AND SAMPLE SIZE

4.1 Primary participants

Primary participants are children who received therapy. The study aims to recruit five children.

4.1.1 Inclusion criteria

1. Born with cleft palate +/- cleft lip
2. Diagnosis of cleft related speech patterns
3. Received therapy for cleft related speech patterns between 18 months and three years
4. The SLT who delivered the therapy meets the secondary participant inclusion criteria
5. Therapy took place in the UK and since 2010
6. Therapy consisted of MSIT and video-therapy
7. A minimum of five therapy records are available (from five separate sessions)
8. The therapist who provided the therapy consents to take part

4.2 Secondary participants

Secondary participants are SLTs who provided therapy. The study aims to recruit five SLTs.

4.2.1 Inclusion criteria

1. Provided therapy to one of the children recruited as a primary participant
2. The child was between 18 months and three years of age and had cleft related speech patterns
3. Therapy took place in the UK
4. Therapy took place from 2010 onwards (after dissemination of the Calladine, 2009, study)
5. Therapy consisted of MSIT and video-therapy
6. A minimum of five therapy records are available (from five separate sessions)
7. The parent of the child who received therapy consents to take part

Being a specialist in cleft palate and related disorders is not an inclusion criterion because in the UK therapy is not always provided by a specialist. To inform the interpretation of the results, recruited participants will be asked to provide the following information:

- Complete number of years qualified as an SLT
- Whether they are a regional specialist, local specialist, or a non-specialist
- Training received on MSIT and video-therapy

- Experience of using MSIT and video-therapy (please see Appendix 10.1)

4.2.2 Exclusions

1. Video records of the therapy sessions are not available
2. Video records are of poor audio and/or visual quality
3. Written case note entries for the therapy sessions are not available
4. Information about child's cleft type, surgical and wound healing history, presence/absence of palatal fistula and speech diagnosis is not available
5. Information about therapist's previous training, experience and current role is not available

5.0 METHOD AND PROCEDURES

5.1 Materials

Materials used in the study will be videos of children receiving therapy and corresponding consultation notes written by the treating SLT. These materials are stored as part of children's clinical records. Videos will have been made and stored with parent/carer consent. As this is participant identifiable data, in addition to consent from participants, the Chief Investigator (CI), Samantha Calladine, will also need a Letter of Access from the treating NHS Trust.

5.2 Data collection and storage

All data will be collected by the CI, Samantha Calladine. Data will be collected in person from the sites where participants are recruited. Data will be collected electronically using encrypted devices provided by Leeds Teaching Hospitals NHS Trust (LTHT): an encrypted external hard-drive (DataLocker Enterprise) and an encrypted laptop. Only the CI has access to these devices. A back-up of the data will be saved on the Trust's secure network/server at Leeds General Infirmary. Access will be restricted to the CI only. Participant identifiable data will be removed from the laptop as soon as the server back-up has been made. These procedures comply with the Data Protection Act 1998 and local information governance practice.

Data collection will be recorded using the Child Data Collection Form and Data Collection Log (please see appendices 10.2 and 10.3 respectively).

During the study, participant identifiable data will be stored electronically on the DataLocker encrypted external hard-drive and the LTHT network only. Electronic and manual files generated during data analysis will be anonymised, except for video samples/edits. Electronic files will be stored on encrypted portable devices (external hard-drive and LTHT laptop). Manual files will be stored in a locked filing cabinet in the CI's personal office at Leeds General Infirmary. No data will be stored on personal computers or laptops. For analysis and write-up, files will be transported between Leeds General Infirmary, the University of Sheffield, and the CI's place of residence. The study will end when Samantha has successfully completed her PhD. Expected completion is March 2018. Data will be destroyed five years post completion.

5.3 Data analysis

Data will be analysed by the CI. All data related activities will take place at the following locations:

- Department of Human Communication Sciences, University of Sheffield

- University libraries, University of Sheffield
- Cleft Lip and Palate Centre, Leeds General Infirmary
- CI's place of residence

Data will be analysed using mixed methods. Appropriate methods of analysis will be adopted based on other studies in the speech and language therapy literature and psychotherapy literature, e.g. Gardner (2006) and Elliott (2010). Quantitative methods have been piloted on data consisting of two children receiving therapy from the same therapist. Given the exploratory nature of the study, methods will be developed and refined throughout the analysis to respond to the data as it unfolds.

Questions about the delivery of MSIT and the therapy process will be answered by analysis of therapist behaviours. Questions about how children's speech patterns develop in response to MSIT will be answered by analysis of child speech behaviours. Analysis will be carried out within an interactional framework and in relation to Stackhouse and Wells' (1997) speech processing model.

A random sample of videos will be analysed by one of the CI's research colleagues in the Department of Human Communication Sciences to provide a measure of reliability. (This person will not be related to the study.) Appropriate statistical tests will be used, e.g. Kappa statistic.

6.0 ETHICAL CONSIDERATIONS

All procedures outlined in this protocol have been discussed and agreed with relevant members of the Research and Innovation, Governance, and Information Technology teams at Leeds Teaching Hospitals NHS Trust to ensure they meet National Institute for Health Research Good Clinical Practice guidelines.

6.1 Informed consent

Active participant involvement in this study is minimal because it will use pre-existing data. Taking part means giving the CI, Samantha Calladine, permission to access personal identifiable information that was obtained as part of routine clinical care.

Informed consent will be obtained from all participants. Secondary participants (SLTs) will be approached first. The CI will provide information about the study and invite SLTs to volunteer if they wish to take part. Written consent will be obtained. Secondary participants will identify and recruit children who meet primary participant criteria. They will contact parents/carers to provide information about the study and written consent will be obtained from those who wish to take part. Please see Appendices 10.4-10.7 for participant information sheets and consent forms.

SLTs and parents/carers will be given the option of providing additional consent for videos to be used in dissemination and publication. Please see Appendices 10.8 and 10.9. It will be made clear to participants that this is an optional level of consent; their videos and information can still be included in the study even if they do not want them to be used in dissemination or any publications.

6.2 Confidentiality

All personal identifiable data will be handled carefully and securely to ensure confidentiality. This will be maintained throughout and following the study by:

1. Using only encrypted electronic devices and a secured server
2. Restricting access to the research team
3. Anonymising electronic and manual files using participant identification numbers*
4. Anonymising data in any publications*

*Video data files will not be anonymised for data analysis, but they will only exist in electronic form and access will be restricted, as described above. No videos or images will be used in publications or other forms of dissemination without consent from participants. If consent is obtained, videos will be anonymised as far as is possible. For example, names will be edited out and any overt identifiable

information will be distorted using appropriate editing software.

7.0 PUBLICATION

An anonymised report of this study will be written up by Samantha Calladine for her PhD thesis.

It is also intended that anonymised results will be reported and disseminated in peer reviewed journals, conference presentations and teaching events.

Participants will be given the option of providing their consent for videos to be used in dissemination.

This is an additional and optional level of consent. Please see Appendices 10.8 and 10.9.

8.0 FINANCIAL ARRANGEMENTS

The costs associated with this study are low. Please see Table 1 below.

Table 1: Research costs (from Sept 2015 to completion March 2018)

Type	Description	Cost
Equipment	DataLocker Enterprise encrypted external hard-drive	None Already purchased by LTHT in 2012 for the Calladine, 2014, pilot study
Equipment	Encrypted laptop	None The CI already has access to a personally-designated laptop provided by LTHT
Equipment	Editing software	None Available at the university
Equipment	Statistics package	None Available at the university
Travel	Maximum of two visits to each site, equating to total of 1500 miles	£300 The CI has an LTHT lease car, therefore cost = 1500 x 14p per mile
Consumables	Stationary, phone calls, printing	£250
Dissemination	Attendance at 1x national conference and 1x national study event	£600
TOTAL		<u>£1150</u>

A request has been made to the Children's Clinical Services Unit at LTHT to cover these costs, and is awaiting approval. Any additional costs will be met personally by the CI and PhD student, Samantha Calladine.

9.0 REFERENCES

Barratt, J. and Extence, H. (2010). Speech and language therapists' predictions made at 18 month cleft palate speech assessments. How accurate are we? Poster presented at The Craniofacial Society of Great Britain and Ireland Annual Scientific Conference, Liverpool.

Bessell, A., Sell, D., Whiting, P., Roulstone, S., Albery, L., Persson, M., Verhoeven, A., Burke, M., Ness, A. (2013). Speech and language therapy interventions for children with cleft palate: A systematic review. *Cleft Palate-Craniofacial Journal*, 50, 1, e1-e17.

Calladine, S. (2009). Multi-sensory input modelling therapy intervention for young children with cleft palate. Unpublished MSc dissertation, University of Sheffield.

Calladine, S. (2014). Capturing the process of therapy for working with two-year-old children with cleft related speech sound difficulties. Unpublished paper submitted as part of PhD confirmation, University of Sheffield.

Chapman, K. L. and Willadsen, E. (2011). The development of speech in children with cleft palate. Chapter 2. In Howard, S. and Lohmander, A. (Editors), 2011, *Cleft Palate Speech*. (West Sussex: Wiley-Blackwell).

Elliott, R. (2010). Psychotherapy change process research: Realizing the promise. *Psychotherapy Research*, 20, 2, 123-135.

Gardner, H. (2006). Training others in the art of therapy for speech sound disorders: An interactional approach. *Child Language Teaching and Therapy*, 22, 1, 27-46.

Harding-Bell, A. and Bryan, A. (2001). The use of multi-sensory input modelling to stimulate speech output processing. A teaching and demonstration video published by Addenbrooke's NHS Trust and East and North Herts NHS Trust.

Lead Speech and Language Therapist Forum. (2015). Personal communication.

Stackhouse, J. and Wells, B. (1997). *Children's Speech and Literacy Difficulties: A Psycholinguistic Framework*. (London: Whurr Publishers Ltd).

10.0 CONTACTS

CHIEF INVESTIGATOR: Samantha Calladine
NHS HOST AND SPONSOR: Leeds Teaching Hospitals NHS Trust
ACADEMIC INSTITUTION: University of Sheffield

Samantha Calladine
PhD Student (University of Sheffield)
Lead Specialist Speech and Language Therapist
Regional Cleft Lip and Palate Service
Paediatric Offices
F Floor, Martin Wing
Leeds General Infirmary
Leeds, LS1 3EX
0113 392 3876
07789 650837
s.calladine@sheffield.ac.uk
scalladine@nhs.net

Ms Anne Gowing
Research Governance Manager
Leeds Teaching Hospitals NHS Trust
Research and Innovation Department
34 Hyde Terrace
Leeds
LS2 9LN
0113 392 0161
anne.gowing@nhs.net

Dr Hilary Gardner, Primary Research Supervisor
Dr Blanca Schaefer, Secondary Research Supervisor
Professor Patty Cowell, Head of Department
Department of Human Communication Sciences
University of Sheffield
362 Mushroom Lane
Sheffield
S10 2TS
Dr Gardner: 0114 222 2456
Dr Schaefer: 0114 222 2423
Professor Patty Cowell: 0114 222 2426
h.gardner@sheffield.ac.uk
blanca.schaefer@sheffield.ac.uk
p.e.cowell@sheffield.ac.uk

Appendix F: WoSRES ethics approval letter

WoSRES
West of Scotland Research Ethics Service



Miss Samantha Calladine
Apartment 1 Sharrow Point
326 Cemetery Road
Sheffield
S11 8FT

West of Scotland REC 5

Ground Floor - Tennent Building
Western Infirmary
38 Church Street
Glasgow
G11 6NT

Date 03 September 2015

Direct line 0141 211 2102
E-mail WoSREC5@ggc.scot.nhs.uk

Dear Miss Calladine

Study title: Multisensory input therapy for young children with cleft related speech patterns
REC reference: 15/WS/0196
Protocol number: SP15/142
IRAS project ID: 160513

The Proportionate Review Sub-committee of the West of Scotland REC 5 reviewed the above application on 02 September 2015.

We plan to publish your research summary wording for the above study on the HRA website, together with your contact details. Publication will be no earlier than three months from the date of this favourable opinion letter. The expectation is that this information will be published for all studies that receive an ethical opinion but should you wish to provide a substitute contact point, wish to make a request to defer, or require further information, please contact the REC Manager Mrs Sharon Macgregor, WoSREC5@ggc.scot.nhs.uk. Under very limited circumstances (e.g. for student research which has received an unfavourable opinion), it may be possible to grant an exemption to the publication of the study.

Ethical opinion

On behalf of the Committee, the sub-committee gave a favourable ethical opinion of the above research on the basis described in the application form, protocol and supporting documentation, subject to the conditions specified below.

Conditions of the favourable opinion

The favourable opinion is subject to the following conditions being met prior to the start of the study.

Management permission or approval must be obtained from each host organisation prior to the start of the study at the site concerned.

Management permission ("R&D approval") should be sought from all NHS organisations involved in the study in accordance with NHS research governance arrangements.

Guidance on applying for NHS permission for research is available in the Integrated Research Application System or at <http://www.rdforum.nhs.uk>.

Where a NHS organisation's role in the study is limited to identifying and referring potential participants to research sites ("participant identification centre"), guidance should be sought from the R&D office on the information it requires to give permission for this activity.

For non-NHS sites, site management permission should be obtained in accordance with the procedures of the relevant host organisation.

Sponsors are not required to notify the Committee of approvals from host organisations.

Registration of Clinical Trials

All clinical trials (defined as the first four categories on the IRAS filter page) must be registered on a publicly accessible database. This should be before the first participant is recruited but no later than 6 weeks after recruitment of the first participant.

There is no requirement to separately notify the REC but you should do so at the earliest opportunity e.g. when submitting an amendment. We will audit the registration details as part of the annual progress reporting process.

To ensure transparency in research, we strongly recommend that all research is registered but for non-clinical trials this is not currently mandatory.

If a sponsor wishes to request a deferral for study registration within the required timeframe, they should contact hra_studyregistration@nhs.net. The expectation is that all clinical trials will be registered, however, in exceptional circumstances non registration may be permissible with prior agreement from the HRA. Guidance on where to register is provided on the HRA website.

It is the responsibility of the sponsor to ensure that all the conditions are complied with before the start of the study or its initiation at a particular site (as applicable).

Ethical review of research sites

The favourable opinion applies to all NHS sites taking part in the study, subject to management permission being obtained from the NHS/HSC R&D office prior to the start of the study (see "Conditions of the favourable opinion").

Summary of discussion at the meeting

The PR Sub-Committee agreed that this was a well presented study with no material ethical issues.

Approved documents

The documents reviewed and approved were:

<i>Document</i>	<i>Version</i>	<i>Date</i>
Other [Data Collection Log]	1.0	28 June 2015

Other [Data Collection Form - Child]	1.0	28 June 2015
Other [Data Collection Form - Therapist]	1.0	28 June 2015
Participant consent form [Parent/Carer Consent Form]	1.0	15 August 2015
Participant consent form [Parent/Carer Consent Form - Videos]	1.0	15 August 2015
Participant consent form [Therapist Consent Form]	1.0	15 August 2015
Participant consent form [Therapist Consent Form - Videos]	1.0	15 August 2015
Participant information sheet (PIS) [Parent/Carer Information Sheet]	1.0	15 August 2015
Participant information sheet (PIS) [Therapist Information Sheet]	1.0	15 August 2015
REC Application Form [REC_Form_25082015]		25 August 2015
Referee's report or other scientific critique report [Supervisor Review]	1.0	15 August 2015
Research protocol or project proposal [Research Protocol]	1.0	15 August 2015
Summary CV for Chief Investigator (CI) [Samantha Calladine CV]	1.0	15 August 2015
Summary CV for supervisor (student research) [Hilary Gardner CV]	1.0	10 August 2015

Membership of the Proportionate Review Sub-Committee

The members of the Sub-Committee who took part in the review are listed on the attached sheet.

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

After ethical review

Reporting requirements

The attached document "After ethical review – guidance for researchers" gives detailed guidance on reporting requirements for studies with a favourable opinion, including:

- Notifying substantial amendments
- Adding new sites and investigators
- Notification of serious breaches of the protocol
- Progress and safety reports
- Notifying the end of the study

The HRA website also provides guidance on these topics, which is updated in the light of changes in reporting requirements or procedures.

User Feedback

The Health Research Authority is continually striving to provide a high quality service to all applicants and sponsors. You are invited to give your view of the service you have received and the application procedure. If you wish to make your views known please use the feedback form available on the HRA website: <http://www.hra.nhs.uk/about-the-hra/governance/quality-assurance/>

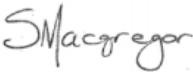
HRA Training

We are pleased to welcome researchers and R&D staff at our training days – see details at <http://www.hra.nhs.uk/hra-training/>

With the Committee's best wishes for the success of this project.

15/WS/0196	Please quote this number on all correspondence
-------------------	---

Yours sincerely



for
Dr Stewart Campbell
Chair

Enclosures: List of names and professions of members who took part in the review
"After ethical review – guidance for researchers"

Copy to: Ms Anne Gowing, Leeds Teaching Hospitals NHS Trust

West of Scotland REC 5

Attendance at PRS Sub-Committee of the REC meeting on 02 September 2015


Committee Members:

Name	Profession	Present	Notes
Dr Stewart Campbell	Consultant Physician & Gastroenterologist (CHAIR)	Yes	
Dr Gillian Harold	Consultant Radiologist	Yes	
Dr Gillian Kerr	Consultant Physician	Yes	

Also in attendance:

Name	Position (or reason for attending)
Mrs Sharon Macgregor	Co-ordinator

Appendix G: NHS sponsor R&D approval letter

The Leeds Teaching Hospitals 
NHS Trust

Anne Gowing

18/09/2015

Samantha Calladine

Leeds Teaching Hospitals NHS Trust
Lead Specialist Speech and Language Therapist
Northern and Yorkshire Regional Cleft Service
(Leeds)
Paediatric Offices, Room 44, F Floor Martin Wing
Leeds

Research & Innovation Department
34 Hyde Terrace
Leeds
LS2 9LN

Tel: 0113 392 0162
Email : leedsth-tr.lthresearch@nhs.net

www.leedsth.nhs.uk/research

Dear Ms Calladine

Re: NHS Permission at LTHT for: Multisensory input therapy for young children with cleft related speech patterns
LTHT R&I Number: SP15/142:
REC: 15/WS/0196

I confirm that *NHS Permission for research* has been granted for this project at The Leeds Teaching Hospitals NHS Trust (LTHT). NHS Permission is granted based on the information provided in the documents listed below. All amendments (including changes to the research team) must be submitted in accordance with guidance in IRAS. Any change to the status of the project must be notified to the R&I Department.

The study must be conducted in accordance with the *Research Governance Framework for Health and Social Care*, ICH GCP (if applicable), the terms of the Research Ethics Committee favourable opinion (if applicable) and NHS Trust policies and procedures (see <http://www.leedsth.nhs.uk/research/>) including the requirements for research governance and clinical trials performance management listed in appendix 1 and 2. NHS permission may be withdrawn if the above criteria are not met including the requirements for clinical trials performance

The Leeds Teaching Hospitals NHS Trust participates in the NHS risk pooling scheme administered by the NHS Litigation Authority "Clinical Negligence Scheme for NHS Trusts" for: (i) medical professional and/or medical malpractice liability; and (ii) general liability. NHS Indemnity for negligent harm is extended to researchers with an employment contract (substantive or honorary) with the Trust. The Trust only accepts liability for research activity with NHS Permission

The Trust therefore accepts liability for the above research project and extends indemnity for negligent harm. Should there be any changes to the research team please ensure that you inform the R&I Department and that s/he obtains an appropriate contract, or letter of access, with the Trust if required.

Yours sincerely



Dr D R Norfolk
Associate Director of R&I

Approved documents

The documents reviewed and approved are listed as follows:-

<i>Document</i>	<i>Version</i>	<i>Date of document</i>
NHS R&D Form		28 August 2015
REC Letter confirming favourable opinion		3 September 2015
CSU Approval		5 August 2015
Data collection log	1.0	28 June 2015
Child data collection log	1.0	28 June 2015
Therapist Data collection form	1.0	28 June 2015
Parent/carer consent form	1.0	15 August 2015
Parent/carer consent form - video	1.0	15 August 2015
Therapist consent form	1.0	15 August 2015
Therapist consent form - video	1.0	15 August 2015
Parent/Carer information sheet	1.0	15 August 2015
Therapist participant information sheet	1.0	15 August 2015
Supervisor review	1.0	15 August 2015
Protocol	1.0	15 August 2015
Samantha Calladine CV		January 2007
Hilary Gardner CV	1.0	10 August 2015

Conditions of NHS Permission for Research:

Appendix 1

Governance requirements:

Managerial approval within the Clinical Support Unit must be obtained before starting the study and healthcare staff should be suitably informed about the research their patients are taking part in and information specifically relevant to their care arising from the study should be communicated promptly.

Agreements must be in place with appropriate support departments.

Arrangements must be in place for the management of financial and other resources provided for the study, including intellectual property arising from the research.

All data and documentation associated with the study must be available for audit/monitoring by authorised Trust or external agencies.

All members of the research team, where applicable, have appropriate employment contracts or letter of agreement to carry out their work in the Trust.

Each member of the research team must be qualified by education, training and experience to discharge his/her role in the study. Students and new researchers must have adequate supervision, support and training.

The research must follow the protocol approved by the relevant research ethics committee. Any proposed amendments to or deviations from the protocol must be submitted for review (as necessary) by the Research Ethics Committee, the Research Sponsor, regulatory authority and any other appropriate body. Where the amendment has resource implications within the CSU, the Directorate research lead/clinical director and R&I should be notified.

Adverse Events in clinical trials of investigational medicinal products must be reported in accordance with the Medicines for Human Use (Clinical Trials) Regulations 2004.

Procedures should be in place to ensure collection of high quality, accurate data and the integrity and confidentiality of data during processing and storage in line with Trust Information Governance Policies and arrangements must be made for the appropriate archiving of data when the research has finished. Records must normally be kept for 15 years.

In compliance with the Health Research Authority (HRA) regulations, clinical trials (and other studies falling within the HRA definition) must be registered on a publically accessible database (such as <https://clinicaltrials.gov/>) prior to commencement. Studies sponsored by LTHT will be registered by the R&I Department.

Findings from the study should be exposed to critical review through accepted scientific and professional channels.

All members of the research team involved in seeking informed consent adheres to GCP standards. Investigators are directed to the R&I website for further information about training in consent for clinical trials.

Studies involving the use of human tissues must be performed in compliance with the code of practice of the Human Tissue Authority.

If you are not able to comply with these requirements, NHS permission to conduct the research in LTHT will be suspended.

Appendix 2

Commercially Sponsored and funded studies.

In line with Trust Standing Financial Instructions there must be a research agreement with the commercial funder signed by the R&I Department (on behalf of the Leeds Teaching Hospitals NHS Trust). Investigators do not have the authority to sign research agreements on behalf of the Trust.

NHS permission for this project to be carried out in the Trust is granted on the understanding that you:

Provide recruitment information when requested by R&I on the Clinical Trial Tracker (available on the CSU Research Hub)

Work with R&I to resolve blocks and delays on trials to ensure that LTHT meets the NIHR benchmarks.

NIHR Benchmarks for Performance in Initiating & Delivering Clinical Research

LTHT clinical trial performance is measured against 2 national benchmarks to improve the initiation and delivery of clinical trials approved by the Trust. NIHR funding to the Trust is conditional on meeting these benchmarks.

Initiation – it should take no more than 70 days from receipt of a valid research application (signed SSI form) by the R&I Department to the recruitment of (ie consenting) the 1st patient to the trial

Delivery – for all trials hosted by the Trust the agreed number of patients must be recruited within the agreed recruitment period

The Trust submits quarterly performance reports to the Department of Health setting out our performance.

For more information about the benchmarks and the work we are doing to support clinical trial management please see the R&I website.

<http://www.leedsth.nhs.uk/research/>

Appendix H: Table showing eligible sites

National profile of early therapy provision for children born with cleft palate

Network:	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	Totals
Children receive therapy between 18mths and 3yrs for cleft related speech problems	Y	Y			Y	Y	Y	Y	Y		Y	Y	Y	Y			Y	12/12
Therapy is provided by:																		5 x no responses
Regional specialist SLT	Y	Y			Y	Y	Y	Y	Y		Y	Y	Y	Y			Y	12/12
Local specialist or link SLT	N	N			Y	Y	Y	Y	Y		Y	Y	Y	Y			N	9/12
Local SLT (not a specialist in cleft)	N	N			Y	Y	Y	Y	Y		N	Y	N	Y some			Y	8/12
Therapy consists of:																		
Multisensory Input Modelling (MSIM) with videotherapy	Y	N			N	N	Y	N	Y		N	Rare but some	N	Y			Y	6/12
MSIM without videotherapy	Y	Y			Y some	Y	Y	Y	Y		Occ.	Y	N	Y			Y	11/12
Other, please specify below	Y	N			Y	Y	NR	Y	Y		Y	N	Y	Y			Y	9/12
The treating SLT/Trust stores copies of therapy videos made for videotherapy	Y	N/A			N	N/A	Y some	N/A	Y		N/A	Y	N/A	N but could get from family			Y	5 of the 6 4(-5) x external sites meet criteria

Y = yes; N = no; NA = not applicable; Occ. = occasionally

Appendix I: R&D approval letter Site 1



Nottingham University Hospitals 
NHS Trust

6th June 2016

Research & Innovation
Nottingham Health Science Partners
C Floor, South Block
Queen's Medical Centre Campus
Derby Road, Nottingham NG7 2UH

Tel: 0115 9709049
www.nuhrise.org

Ms Lorraine Britton
Regional Lead Speech & Language Therapist
Nottinghamshire Health Care NHS Trust
Trent Regional Cleft Lip and Palate Service
Nottingham City Hospital
Hucknall Road
NG5 1PB

Dear Ms Lorraine Britton

Short Title / Acronym	Multisensory input therapy for young children with cleft related speech patterns /
CSP Number	
R&I REF	16CS012
Long Title	Multisensory input therapy for young children with cleft related speech patterns
PROJECT MILESTONES	
Recruitment Target	2
Date of Valid Submission	06/06/2016
Recruitment End Date	31/12/2016
1st Patient to be Recruited by	15/08/2016

The R&I Department has reviewed the following documents and NHS permission for the above research has been granted on the basis described in the application form, protocol, and supporting documentation. The documents reviewed were:

Document	Version	Date
Other-data collection log	1.0	28 June 2015
Other- data collection form child	1.0	28 June 2015
Other- data collection form-therapist	1.0	28 June 2015
Participant Consent Form- parent/carer consent form	1.0	15 August 2015
Participant Consent Form- parent/carer consent form videos	1.0	15 August 2015
Participant Consent Form- Therapist	1.0	15 August 2015
Participant Consent Form- Therapist videos	1.0	15 August 2015
Participant Information Sheet-parent/carer	1.0	15 August 2015
Participant Information Sheet- therapist	1.0	15 August 2015
Protocol	1.0	15 August 2015

Your study now has NHS permission, on the understanding and provision that you will follow the conditions set out below.

Conditions of Approval

The Principal Investigator is responsible for:

1. Compliance with all relevant laws, regulations and codes of practice applicable to the trial including but not limited to, the UK Clinical Trials Regulations, Medicines for Human Use (Clinical Trial) Regulations 2004, principles of Good Clinical Practice, the World Medical Association Declaration of Helsinki entitled 'Ethical Principles for Medical Research Involving Human Subjects' (2013 version), the Human Rights Act 1998, the Data Protection Act 1998, the Medicines Act 1968, and the NHS Research Governance Framework for Health and Social Care (version 2 April 2005). Should any of these be revised and reissued this will apply. Copies of the up-to-date regulations are available from the R&I Office or via the R&I website <http://nuhrise.org>.
2. Submission of study amendments to the Ethics committee and MHRA in accordance with the IRAS guidelines. Amendments and information with regards to changes in study status must be sent to R&I, (this includes changes to the local study team). Within 35 days from the receipt of a valid amendment submission, the R&I department will inform you if the amendment cannot be implemented locally. If no objections are raised NHS permission is valid and the amendment may be implemented.

When submitting documents for studies adopted into the NIHR portfolio please send the information to the Clinical Research Network: East Midlands (CRN:EM) (CSP.CRNEastMidlands@NIHR.ac.uk).

When submitting documents for all other studies please use the email address rdamend@nuh.nhs.uk.

3. Ensuring all study personnel, not employed by the Nottingham University Hospitals NHS Trust hold either honorary contracts/letters of access with this Trust, before they have access to any patients or staff, their data, tissue or organs or any NUH facilities.
4. For initiating and delivering research in accordance with the Department of Health's Plan for Growth. The first patient, first visit should occur within 70 days from the receipt of a valid submission in R&I. This applies to all studies where :
 - i. The research is classed as a "clinical trial "on the IRAS filter page (first 4 categories)
5. Ensuring the research team via an identified individual, collaborates with the department of R&I and the CRN:EM in reporting recruitment data using Documas and the CRN EM Study Tracker.
6. Ensuring that for GTAC-approved studies, the NHS permission is forwarded to GTAC via the sponsor. GTAC should then issue a site authorisation letter which must be received by each site prior to recruitment commencing. A copy of this letter must be forwarded to R&I.
7. Comply with requests from NUH R&I to allow monitoring of research to comply with the Research Governance Framework and other applicable regulations.
8. Record all types of adverse events (including Suspected Unexpected Serious Adverse Drug Reaction - SUSARS) in the patient medical records and study documentation and report to the sponsor as required by the protocol.
9. Report any Serious Breach of the UK Clinical Trial regulations in connection with the trial or Serious Breach of the protocol, immediately after becoming aware of the breach to the study sponsor.
10. Reporting any changes to the study to R&I by letter or e mail. These should not be implemented until agreed with R&I.

For NUH sponsored studies only, the Chief Investigator is responsible for:

- i. All duties as detailed in the "Clinical Trial Delegation of Sponsorship responsibilities to Chief Investigator" agreement.
- ii. Contacting the sponsor for review of all amendment documentation prior to submission to the HRA and MHRA. Please note that according to HRA and MHRA regulations, all submissions of amendments need to be signed by the authorised sponsor's representative. All relevant documentation should be emailed to rdamend@nuh.nhs.uk.
- iii. Sending copies of the completed Annual Progress Reports, Development Safety Update Reports, and End of Study report required by the Ethics Committee and the MHRA (if appropriate) to the sponsor researchsponsor@nuh.nhs.uk.

- iv. Notifying NUH R&I of all SAEs by completing and sending the *"Serious Adverse Event reporting form"* to R&I (only via fax, e-mail or by hand), within 24hrs of becoming aware of the event. Further guidance can be found in the R&I Adverse Event SOP (SOP-RES-019).
- v. Reporting any Serious Breach of the UK Clinical Trial regulations in connection with the trial or Serious Breach of the protocol, immediately after becoming aware of the breach to NUH R&I as sponsor. Further guidance can be found in the R&I Non Compliance and Serious Breach Reporting SOP (SOP-RES-017).

This approval letter constitutes a favourable Site Specific Assessment (SSA) for this site.

If you have any queries regarding the milestones or points detailed in this letter, please contact the Research Project Manager responsible for managing the performance of the study at NUH. This information is available on <http://nuhrise.org>.

Please note that the R&I department maintains a database containing study related information, and personal information about individual investigators e.g. name, address, contact details etc. This information will be managed according to the principles established in the Data Protection Act.

Yours sincerely,



Dr Stephen Ryder / Dr Maria Koufali
Director of Research and Innovation / Deputy Director Research and Innovation

Appendix J: R&D approval letter Site 2

Research and Development Department

R&D ref: A093892

27 April 2016

Mrs Caroline Hattee
East of England Cleft Network, Box 46
Addenbrooke's Hospital
Hills Road, Cambridge
CB2 0QQ

Box 277
Addenbrooke's Hospital
Hills Road
Cambridge
CB2 0QQ

Direct Dial: 01223 348455
Switchboard: 01223 245151

E-mail: tracy.assari@addenbrookes.nhs.uk
r&denquiries@addenbrookes.nhs.uk
www.addenbrookes.org.uk

Dear Mrs Caroline Hattee

Re: 15/WS/0196 Multisensory input therapy for young children with cleft related speech patterns

In accordance with the Department of Health's Research Governance Framework for Health and Social Care, all research projects taking place within the Trust must receive a favourable opinion from an ethics committee and approval from the Department of Research and Development (R&D) prior to commencement.

R&D have reviewed the documentation submitted for this project, and has undertaken a **site specific assessment** based on the information provided in the SSI form, and I am pleased to inform you that we have no objection to the research proceeding within Cambridge University Hospitals NHS Foundation Trust.

Sponsor: Leeds Teaching Hospitals NHS Trust

Funder: Own Account

End date: 2/9/2016

Protocol: version 1, dated 15 August 2015

Conditions of Trust approval:

- The project must follow the agreed protocol and be conducted in accordance with all Trust Policies and Procedures especially those relating to research and data management. Any mobile devices used must also comply with Trust policies and procedures for encryption to AES 256.
- You and your research team must ensure that you understand and comply with the requirements of the NHS Confidentiality Code of Practice and the Data Protection Act 1998 and are aware of your responsibilities in relation to the Human Tissue Act 2004, Good Clinical Practice, the NHS Research Governance Framework for Health and Social Care, Second Edition April 2005 and any further legislation released during the time of this study.

- Members of the research team must have appropriate substantive or honorary contracts with the Trust prior to the study commencing. Any additional researchers who join the study at a later stage must also hold a suitable contract.
- You and your research team must provide to R&D, as soon as available, the date of first patient first visit/date of randomisation.
- Provide to R&D on a quarterly basis the recruitment data and progress of the study including copies of any/all event reports.

If the project is a clinical trial under the European Union Clinical Trials Directive the following must also be complied with:

- the EU Directive on Clinical Trials (Directive 2001/20/EC) and UK's implementation of the Directive: The Medicines for Human Use (Clinical Trials) Regulations 2004;
- the EU Directive on Principles and Guidelines for Good Clinical Practice (EU Commission Directive 2005/28/EC); and UK's implementation of the Directive: The Medicines for Human Use (Clinical Trials) Amendment Regulations 2006;

Amendments

Please ensure that you submit a copy of any amendments made to this study to the R&D Department.

Annual Report

It is obligatory that an annual report is submitted by the Chief Investigator to the research ethics committee, and we ask that a copy is sent to the R&D Department. The yearly period commences from the date of receiving a favourable opinion from the ethics committee.

Please refer to our website www.cuh.org.uk/research for all information relating to R&D including honorary contract forms, policies and procedures and data protection.

Should you require any further information please do not hesitate to contact us.

Yours sincerely



Louise Stockley
Research Governance Manager

V7 March 2012

Appendix K: Letter of Access form Site 1



Nottingham University Hospitals **NHS**
NHS Trust

Research & Innovation
Nottingham Integrated Clinical Research Centre
C Floor, South Block
QMC Campus
Derby Road
Nottingham
NG7 2UH

RSCH 693

8th June 2016

Samantha Calladine
Leeds Teaching Hospitals NHS Trust
Regional Cleft Lip and Palate Service
Room 44 Pediatric Offices
F Floor, Martin Wing
Leeds General Infirmary
LS2 9LN

Tel: 0115 970 9049

www.nuhrise.org

Dear Miss Calladine

Re: Letter of Access for Research

Study Title:	Multisensory input therapy for young children with cleft related speech patterns
Chief Investigator:	Samantha Calladine
Local Collaborator at NUH:	Lorraine Britton
R&D Ref:	16CS012
Sponsor:	Leeds Teaching Hospitals NHS Trust

Letter of access for research

As an existing NHS employee you do not require an additional honorary research contract with this NHS organisation. We are satisfied that the research activities that you will undertake in this NHS organisation are commensurate with the activities you undertake for your employer. Your employer is fully responsible for ensuring such checks as are necessary have been carried out. Your employer has confirmed in writing to this NHS organisation that the necessary pre-engagement checks are in place in accordance with the role you plan to carry out in this organisation. This letter confirms your right of access to conduct research through **Nottingham University Hospitals** for the purpose and on the terms and conditions set out below. This right of access commences on **08/06/2016** until **31/12/2016** unless terminated earlier in accordance with the clauses below.

You have a right of access to conduct such research as confirmed in writing in the letter of permission for research from this NHS organisation. Please note that you cannot start the research until the Principal Investigator for the research project has received a letter from us giving permission to conduct the project.

You are considered to be a legal visitor to **Nottingham University Hospitals** premises. You are not entitled to any form of payment or access to other benefits provided by this organisation to employees and this letter does not give rise to any other relationship between you and this NHS organisation, in particular that of an employee.

While undertaking research through **Nottingham University Hospitals** you will remain accountable to your employer **Leeds Teaching Hospitals NHS**



Trust but you are required to follow the reasonable instructions of your nominated manager **Lorraine Britton** in this NHS organisation or those given on her/his behalf in relation to the terms of this right of access.

Where any third party claim is made, whether or not legal proceedings are issued, arising out of or in connection with your right of access, you are required to co-operate fully with any investigation by this NHS organisation in connection with any such claim and to give all such assistance as may reasonably be required regarding the conduct of any legal proceedings.

You must act in accordance with **Nottingham University Hospitals** policies and procedures, which are available to you upon request, and the Research Governance Framework.

You are required to co-operate with **Nottingham University Hospitals** in discharging its duties under the Health and Safety at Work etc Act 1974 and other health and safety legislation and to take reasonable care for the health and safety of yourself and others while on **Nottingham University Hospitals** premises. Although you are not a contract holder, you must observe the same standards of care and propriety in dealing with patients, staff, visitors, equipment and premises as is expected of a contract holder and you must act appropriately, responsibly and professionally at all times.

If you have a physical or mental health condition or disability which may affect your research role and which might require special adjustments to your role, if you have not already done so, you must notify your employer and the Trust **R&I office** prior to commencing your research role at the Trust.

You are required to ensure that all information regarding patients or staff remains secure and *strictly confidential* at all times. You must ensure that you understand and comply with the requirements of the NHS Confidentiality Code of Practice (<http://www.dh.gov.uk/assetRoot/04/06/92/54/04069254.pdf>) and the Data Protection Act 1998. Furthermore you should be aware that under the Act, unauthorised disclosure of information is an offence and such disclosures may lead to prosecution.

Nottingham University Hospitals will not indemnify you against any liability incurred as a result of any breach of confidentiality or breach of the Data Protection Act 1998. Any breach of the Data Protection Act 1998 may result in legal action against you and/or your substantive employer.

You should ensure that, where you are issued with an identity or security card, a bleep number, email or library account, keys or protective clothing, these are returned upon termination of this arrangement. Please also ensure that while on the premises you wear your ID badge at all times, or are able to prove your identity if challenged. Please note that this NHS organisation accepts no responsibility for damage to or loss of personal property.

We may terminate your right to attend at any time either by giving seven days' written notice to you or immediately without any notice if you are in breach of any of the terms or conditions described in this letter or if you commit any act that we reasonably consider to amount to serious misconduct or to be disruptive and/or prejudicial to the interests and/or business of this NHS organisation or if you are convicted of any criminal offence. You must not undertake regulated activity if you are barred from such work. If you are barred from working with adults or children this letter of access is immediately terminated. Your employer will immediately withdraw you from undertaking this or any other regulated activity and you **MUST** stop undertaking any regulated activity immediately. Your substantive employer is responsible for your conduct during this research project and may in the circumstances described above instigate disciplinary action against you.

If your circumstances change in relation to your health, criminal record, professional registration or suitability to work with adults or children, or any other aspect that may impact on your suitability to conduct research, or your role in research changes, you must inform the NHS organisation that employs



you through its normal procedures. You must also inform your nominated manager in this NHS organisation.

Yours sincerely

Joanne Thornhill
Research Project Manager Team Leader

**cc: HR department of Employing Organisation
NUH HR Department**



Appendix L: Letter of Access form Site 2

Research and Development Department

Box 277
Addenbrooke's Hospital
Hills Road
Cambridge
CB2 0QQ

Miss Samantha Calladine
Clinical Lead Specialist Speech and Language Therapist
Leeds Teaching Hospitals NHS Trust
Research and Innovation Dept
34 Hyde Terrace
Leeds General Infirmary
Leeds
LS2 9LN

R&D Manager: Stephen Kelleher
stephen.kelleher@addenbrookes.nhs.uk
HR Manager: Debbie Richards
01223 274660
deborah.richards@addenbrookes.nhs.uk
HR Advisor: Gayle Lindsay
01223 348496
gayle.lindsay@addenbrookes.nhs.uk

6th June 2016

Dear Samantha

Letter of access for research – A093892 - Multisensory input therapy for young children with cleft related speech patterns

As an existing NHS employee you do not require an additional honorary research contract with this NHS organisation. We are satisfied that the research activities that you will undertake in this NHS organisation are commensurate with the activities you undertake for your employer. Your employer is fully responsible for ensuring such checks as are necessary have been carried out. Your employer has confirmed in writing to this NHS organisation that the necessary pre-engagement checks are in place in accordance with the role you plan to carry out in this organisation. This letter confirms your right of access to conduct research through Cambridge University Hospitals NHS Foundation Trust for the purpose and on the terms and conditions set out below. This right of access commences on **23rd May 2016** and ends on **22nd May 2019** unless terminated earlier in accordance with the clauses below.

You have a right of access to conduct such research as confirmed in writing in the letter of permission for research from this NHS organisation. Please note that you cannot start the research until the Principal Investigator for the research project has received a letter from us giving permission to conduct the project and you have provided the Trust's R&D department with written evidence that you have completed GCP training from an EU institution before you start your research.

The information supplied about your role in research at Cambridge University Hospitals NHS Foundation Trust has been reviewed and you do not require an honorary research contract with this NHS organisation. We are satisfied that such pre-engagement checks as we consider necessary have been carried out.

You are considered to be a legal visitor to Cambridge University Hospitals NHS Foundation Trust premises. You are not entitled to any form of payment or access to other benefits provided by this NHS organisation to employees and this letter does not give rise to any other relationship between you and this NHS organisation, in particular that of an employee.

While undertaking research through Cambridge University Hospitals NHS Foundation Trust, you will remain accountable to your place of work, **Leeds General Infirmary** but you are required to follow the reasonable instructions of **Mrs Caroline Hattee and Stephen Farrell** in this NHS organisation or those given on his behalf in relation to the terms of this right of access.

Where any third party claim is made, whether or not legal proceedings are issued, arising out of or in connection with your right of access, you are required to co-operate fully with any investigation by this NHS organisation in connection with any such claim and to give all such assistance as may reasonably be required regarding the conduct of any legal proceedings.

You must act in accordance with Cambridge University Hospitals NHS Foundation Trust policies and procedures, which are available to you upon request, and the Research Governance Framework.

You are required to co-operate with Cambridge University Hospitals NHS Foundation Trust in discharging its duties under the Health and Safety at Work etc Act 1974 and other health and safety legislation and to take reasonable care for the health and safety of yourself and others while on Cambridge University Hospitals NHS Foundation Trust premises. You must observe the same standards of care and propriety in dealing with patients, staff, visitors, equipment and premises as is expected of any other contract holder and you must act appropriately, responsibly and professionally at all times.

If you have a health condition or disability which may affect your research role and which might require reasonable special adjustments to your role, if you have not already done so, you must notify your employer and the Trust's R&D HR Office prior to commencing your research role at the Trust.

You are required to ensure that all information regarding patients or staff remains secure and *strictly confidential* at all times. Personal identifiable data must be carried securely at all times and mobile devices must be encrypted. You must ensure that you understand and comply with the requirements of the NHS Confidentiality Code of Practice (<https://www.gov.uk/government/publications/confidentiality-nhs-code-of-practice>) and the Data Protection Act 1998. Furthermore you should be aware that under the Act, unauthorised disclosure of information is an offence and such disclosures may lead to prosecution. Data controllers could also be fined for a breach of the Data Protection Act 1998. You must familiarise yourself with the Trust's Information Governance Code of Conduct.

You must keep confidential any information regarding the design, conduct or management or results of any research unless authorised in writing by the Trust to disclose it. You must acknowledge the Trust's contribution in any publication arising out of this Agreement.

Subject to any agreement with your employer to the contrary (e.g. as part of a multi-centre study), any Intellectual Property (IP) resulting from research carried out under this Agreement will be the property of the Trust and you will do all things necessary or desirable to give effect to the assignment of this IP.

You should ensure that, where you are issued with an identity or security card, a bleep number, email or library account, keys or protective clothing, these are returned upon termination of this arrangement. Please also ensure that while on the premises you wear your ID badge at all times, or are able to prove your identity if challenged. Please note that this NHS organisation accepts no responsibility for damage to or loss of personal property.

We may terminate your right to attend at any time either by giving seven days' written notice to you or immediately without any notice if you are in breach of any of the terms or conditions described in this letter or if you commit any act that we reasonably consider to amount to serious misconduct or to be disruptive and/or prejudicial to the interests and/or business of this NHS organisation or if you are convicted of any criminal offence. You must not undertake regulated activity if you are barred from such work. If you are barred from working with adults or children this letter of access is immediately terminated. Your employer will immediately withdraw you from undertaking this or any other regulated activity and you **MUST** stop undertaking any regulated activity immediately.

Your substantive employer is responsible for your conduct during this research project and may in the circumstances described above instigate disciplinary action against you.

Cambridge University Hospitals NHS Foundation Trust will not indemnify you against any liability incurred as a result of any breach of confidentiality or breach of the Data Protection Act 1998. Any breach of the Data Protection Act 1998 may result in legal action against you and/or your substantive employer.

INDUCTION AND MANDATORY TRAINING

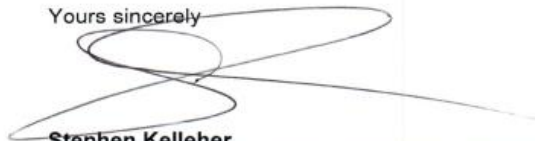
You are responsible for familiarising yourself with the Trust's policies and mandatory training courses such as Moving and Handling, Health and Safety, Fire Training etc and be aware of the responsibility to maintain a safe environment for patients, staff and visitors

Your host Manager will ensure that you receive a comprehensive Departmental Induction. She/he will also provide you with details of Corporate Induction, research specific induction and annual Mandatory Refresher Training.

If your letter of access is for more than 3 months, you must attend Corporate Induction. Where your letter of access is for more than 12 months, you must attend annual Mandatory Refresher Training.

If your current role or involvement in research changes, or any of the information provided in your Research Passport changes, you must inform your employer through their normal procedures. You must also inform your nominated manager in this NHS organisation.

Yours sincerely




Stephen Kelleher
R&D Manager, Cambridge University Hospitals NHS Foundation Trust

cc: Mrs Caroline Hattee, East of England Cleft Network, Box 46
Stephen Farrell, Clinical Director, Surgery/Paeds, Division E, Box 267
Richard Pierce, Senior Human Resources Manager, Richard.pierce.nhs.net

Enc: P6 form to confirm Letter of Access issued
ID Badge Form

Appendix M: Invitation letter to therapist participants



The Leeds Teaching Hospitals 
NHS Trust

Samantha Calladine
PhD Student

Department of Human Communication Sciences
University of Sheffield
Mushroom Lane
Sheffield
S10 2TS

Mobile: 07789 650837/07951 350434

Email: s.calladine@sheffield.ac.uk

19th June 2016

Dear Specialist Speech and Language Therapist

Re: Invitation to take part in a research study

I would like to invite you to take part in a research study. The title of the study is:

'Multisensory input therapy for young children with cleft related speech patterns'

The study has received NHS ethical approval (REFERENCE) and NAME OF NHS TRUST Research and Development approval (REFERENCE).

I am writing to you because you have been identified as potentially meeting the study's inclusion criteria. I would be grateful if you would verify this by reviewing the criteria listed in the email forwarded to you by NAME OF LEAD SLT, Lead Specialist Speech and Language Therapist.

I have enclosed with this invitation letter a participant information sheet providing details of the study and what it would involve for you if you were to take part. Please take your time to read this carefully. If you have any questions about the study and/or being involved, please contact me using the details at the top of this letter.

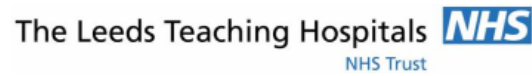
If you are interested in taking part in the study, please let me know and I will respond with further details. If we confirm that you meet the inclusion criteria I would ask you to complete a consent form. I will go through this with you on the telephone or, if you prefer, I would meet with you in person at the Cleft Lip and Palate Centre where you are based.

Thank you very much for taking the time to read this invitation and the enclosed information.

Yours faithfully

Samantha Calladine
PhD Student

Appendix N: Therapist Information Sheet



Therapist Participant Information Sheet

(Version 1.0)

Research project: Multisensory input therapy for young children with cleft related speech patterns

Researcher: Samantha Calladine

Invitation

You are being invited to take part in the above research project by giving your permission for us to use previously recorded videos of you doing Multisensory Input Therapy (MSIT).

It is important you understand why the research is being carried out and what it will involve. Please take time to read the following information carefully and discuss it with others if you wish. Please ask Samantha Calladine if anything is not clear or if you would like more information.

Who is carrying out the study?

This project is part of Samantha Calladine's PhD. It is being carried out by Samantha with the support of:

- University of Sheffield, where Samantha is studying for her PhD
- Leeds Teaching Hospitals NHS Trust, where Samantha is employed as Lead Specialist Speech and Language Therapist for the Northern and Yorkshire Regional Cleft Lip Palate Service (Leeds)

What is the purpose of the study?

A cleft palate affects babble and early speech development. Cleft palate repair helps to normalise the structure and function of children's palates so that cleft related patterns resolve. However, this does not happen for all children and their speech development can continue to be affected after surgery.

Cleft related speech patterns are usually identified at children's 18-24 month assessment. Speech and language therapy is often arranged. One type of therapy is MSIT. The Speech and Language Therapist (SLT) models speech for the child to increase their awareness of different sounds and to give them opportunities to practice new sounds. Filming this therapy (video-therapy) is sometimes used so that parents/carers can watch the therapy at home with their child.

This study aims to learn more about the impact of MSIT on children's speech patterns. We want to understand more about the types of activities and stimuli used in therapy and how these affect a child's speech. We are also interested in how the SLT and child interact with each other during therapy and how the SLTs help children to change their speech production.

Why have I been invited?

You have been invited to take part in this study because you have done MSIT with a child born with cleft palate when s/he was between 18 months and three years old, and you filmed this for video-therapy. In order for you to take part there would need to be:

- Copies of these videos in the child's health records, and
- Videos from at least five different therapy sessions with the same child

What will I need to do if I take part?

If you want to take part you will need to give your permission for Samantha Calladine to look at and analyse videos of your therapy that have already been collected. We will also ask you to do the following:

1. Invite the parent/carer of the children you saw for therapy to also give permission for Samantha to view the videos of their therapy sessions with you

This will involve contacting them by phone to explain the study and sending them an information sheet and consent form that we will give to you. Both you and each child's parents/carers will need to give permission to take part in order for us to include videos of you and the child in the study.

2. Provide brief information about your previous experience, training and your current role
3. Help us collect information about the child and his/her therapy; please see section below

Do I have to take part?

No, it is completely up to you.

What will the study involve?

Samantha will visit you at your workplace and ask for your help in finding the following information in the child's health records:

- Type of cleft the child was born with
- What cleft related operations the child has had and how old s/he was
- Whether the child has had a hearing loss and how this has been treated
- How old the child was when s/he had therapy
- Videos of the child's therapy between 18 months and three years
- Notes you made about the child's therapy sessions

Samantha will make copies of this information and save it on two devices given to her by Leeds Teaching Hospitals NHS Trust: a laptop and an external hard-drive. Both devices are encrypted and only Samantha knows the passwords. This information will be returned to Leeds General Infirmary where it will be transferred from the laptop to the hospital's secure computer network.

Samantha will watch the therapy videos and carry out analysis of the following:

- How different activities and stimuli affect the child's speech productions
- How you and the child interact with each other
- Speech sound productions the child makes during therapy
- Changes in the child's speech production during therapy and how and when these occurred

To check the reliability of the analysis a small sample of your child's videos will also be analysed by another postgraduate student in the Human Communication Sciences department at the University of Sheffield. No other information about your child would be given.

Samantha is expected to finish her PhD in March 2018. We will keep the information we collect for five years after this, and then destroy it.

What are the possible advantages and disadvantages?

There are no specific advantages to you. However, we hope the findings will help to develop therapy with children with speech production difficulties. This might benefit you in your future SLT practice.

We do not know of any disadvantages. We will not make judgements about you as a therapist or the therapy you provide. We will handle your videos and information securely to minimise any risks of loss and/or of it being seen by someone outside the research team.

Is there someone I can talk to about the study?

Yes, you can talk with Samantha, who is carrying out the study. If you want to speak with someone who is not involved in the study, you can speak with the Lead Specialist Speech and Language Therapist (*name will be inserted*) at the Regional Cleft Lip and Palate Service you liaise with. We have put all the contact details on the last page of this information sheet.

What if I have any concerns?

Please discuss these with Samantha in the first instance and she will do her best to resolve them. If you still have concerns you can contact Samantha's research supervisors (Dr Hilary Gardner and Dr Blanca Schaefer) at the University of Sheffield or Professor Patty Cowell, the Head of Department. If you still have concerns after this you will be given the contact details for the university registrar. You can also contact the Research and Innovation department at Leeds General Infirmary. We have put all the contact details on the last page of this information sheet.

Will my involvement in this study be kept confidential?

Yes. Samantha will access your information in line with ethical and legal practice guidelines and handle it in confidence. She will analyse the videos in secure places only. She may look at these with her research supervisors at the university, who also have a duty of confidentiality to you.

Your information will be kept strictly confidential and stored in the secure offices of the Northern and Yorkshire Regional Cleft Lip and Palate Service at Leeds General Infirmary. Electronic information will be stored securely on a password protected computer network and only Samantha will have access to this. When information is taken out of the hospital, it will not contain written personal details about you, such as your name or place of work, and the information-carrying device will be encrypted.

What will happen to the results of this study?

Samantha will write up an anonymised report of the study as her PhD thesis.

She will also share the anonymised results with other SLTs to help develop their therapy practice and other studies in this area. They might also be shared with other professionals, for example, cleft lip and palate surgeons, who work with children born with cleft palate. Samantha will share the results by giving talks and writing papers and other publications.

You will not be named in the thesis or any report or publication arising from it. When we present results from studies about therapy, and when we train students, we sometimes show video examples. We will ask your permission to use videos of you in this way. You can still give us permission to watch your videos in the study even if you do not want us to use them when we present the results.

Who has reviewed the study?

This study has been formally reviewed by an NHS research ethics committee. It has also been reviewed by the Research and Innovation departments at Leeds Teaching Hospitals NHS Trust and the University of Sheffield.

Contact details

Samantha Calladine
PhD Student (University of Sheffield)
Lead Specialist Speech and Language Therapist
Regional Cleft Lip and Palate Service (Leeds)
Paediatric Offices
F Floor, Martin Wing
Leeds General Infirmary
Leeds, LS1 3EX
0113 392 3876
07789 650837
s.calladine@sheffield.ac.uk
scalladine@nhs.net

NAME
Lead Specialist Speech and Language Therapist
Cleft Lip and Palate Centre
ADDRESS
ADDRESS
LANDLINE
MOBILE
EMAIL

NAME
Speech and Language Therapist
Speech and Language Therapy Department
ADDRESS
ADDRESS
LANDLINE
MOBILE
EMAIL


Dr Hilary Gardner, Research Supervisor
Dr Blanca Schaefer, Research Supervisor
Professor Patty Cowell, Head of Department
Department of Human Communication Sciences
University of Sheffield
362 Mushroom Lane
Sheffield
S10 2TS
Dr Hilary Gardner: 0114 222 2456
Dr Blanca Schaefer: 0114 222 2423
Professor Patty Cowell: 0114 222 2426
h.gardner@sheffield.ac.uk
blanca.schaefer@sheffield.ac.uk
p.e.cowell@sheffield.ac.uk

Ms Anne Gowing
Research Governance Manager
Leeds Teaching Hospitals NHS Trust
Research and Innovation Department
34 Hyde Terrace
Leeds
LS2 9LN
0113 392 0161
anne.gowing@nhs.net

Patient Advice and Liaison Service (PALS)
Leeds Teaching Hospitals NHS Trust
0113 206 7168 or 0113 206 6261
patientexperience.leedsth@nhs.net

Appendix O: Invitation letter to parent/carers of child participants



The Leeds Teaching Hospitals 
NHS Trust

Samantha Calladine
PhD Student
Department of Human Communication Sciences
University of Sheffield
Mushroom Lane
Sheffield
S10 2TS
Mobile: 07789 650837/07951 350434
Email: s.calladine@sheffield.ac.uk

25th July 2016

Dear Parent/Carer

Re: Invitation to take part in a research study

I would like to invite you to take part in a research study. The title of the study is:

'Multisensory input therapy for young children with cleft related speech patterns'

An NHS research ethics committee (REFERENCE) and the Research and Development department at NAME OF NHS TRUST (REFERENCE) have approved this study.

I am writing to you because your child has been found to meet the study's inclusion criteria. The Speech and Language Therapist who has treated your child, NAME OF SLT, also meets the criteria.

I have included an information sheet with this letter which contains details of the study. It explains what the study would involve for you and your child if you took part. Please take your time to read this carefully. If you have any questions about the study and/or being involved, please contact NAME OF SLT on TELEPHONE NUMBER. You are also welcome to contact me if you wish using the contact details at the top of this letter.

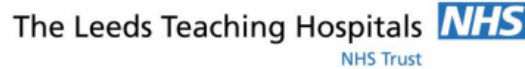
I have also included two consent forms with this invitation. If you are interested in taking part in the study, please complete these forms and return them in the envelope provided to NAME OF SLT. If you would prefer, we can go through this with you on the telephone or in person at the NAME OF THE CLEFT LIP AND PALATE UNIT in NAME OF CITY.

Thank you very much for taking the time to read this invitation and the enclosed information.

Yours faithfully

Samantha Calladine
PhD Student (University of Sheffield)
Lead Specialist Speech and Language Therapist (Leeds Teaching Hospitals NHS Trust)

Appendix: P: Parent/Carer Information Sheet



Parent/Carer Participant Information Sheet

(Version 1.0)

Research project: Multisensory input therapy for young children with cleft related speech patterns

Researcher: Samantha Calladine

Invitation

You are being invited to take part in the above research project by giving your permission for us to use previously recorded videos of your child that were made as part of her/his speech and language therapy care.

It is important you understand why the research is being carried out and what it will involve. Please take time to read the following information carefully and discuss it with others if you wish. Please ask Samantha or your Speech and Language Therapist if anything is not clear or if you would like more information.

Who is carrying out the study?

This project is part of Samantha Calladine's PhD. It is being carried out by Samantha with the support of:

- University of Sheffield, where Samantha is studying for her PhD
- Leeds Teaching Hospitals NHS Trust, where Samantha is employed as Lead Specialist Speech and Language Therapist for the Northern and Yorkshire Regional Cleft Lip and Palate Service (Leeds)

What is the purpose of the study?

A cleft palate affects what sounds children can make. Before their palate operation, children's 'b' and 'd' sounds will sound more like 'm' and 'n'. Children may also make more sounds at the back of their mouth, like 'g,' than at the front of their mouth. Such cleft related speech patterns may start to disappear after the palate operation, but may not in some children.

All children born with a cleft palate are seen for a speech and language assessment at 18-24 months of age. Speech and language therapy is usually arranged for children who have cleft related patterns in their speech. One type of therapy is Multisensory Input Therapy (MSIT). The therapist models speech for the child to increase their awareness of different sounds and to give them opportunities to practice new sounds. Filming this therapy (video-therapy) is sometimes used so that parents/carers can watch the therapy at home with their child.

Samantha is carrying out a study to learn more about how MSIT affects children's speech productions. She wants to understand more about the types of activities and stimuli used in therapy and how these affect a child's speech. She is also interested in how the therapist and child interact with each other during therapy and how to help children change their speech production.

Why have I been invited?

You have been invited to take part in this study because your child had MSIT and video-therapy when s/he was between 18 months and three years old.

What will I need to do if I take part?

You will not need to do anything other than give your permission for Samantha Calladine to look at and analyse your child's previously recorded therapy videos and to access information that has already been collected as part of your child's routine care.

Do I have to take part?

No, it is completely up to you.

What will the study involve?

Samantha will visit the speech and language therapist who carried out therapy with your child. The therapist will help Samantha find the following information in your child's health records.

- Type of cleft your child was born with
- What cleft related operations your child has had and how old s/he was
- Whether your child has had a hearing loss and how this has been treated
- Your child's speech and language development at her/his 18-24 month assessment
- How old your child was when s/he had therapy
- Videos of your child's therapy between 18 months and three years
- Notes made by the therapist about your child's therapy sessions

Samantha will make copies of this information and save it on two devices given to her by Leeds Teaching Hospitals NHS Trust: a laptop and an external hard-drive. Both devices are encrypted and only Samantha knows the passwords. This information will be returned to Leeds General Infirmary where it will be transferred from the laptop to the hospital's secure computer network.

Samantha will watch your child's therapy videos and carry out analysis of the following:

- How different activities and stimuli influence your child's speech productions
- How the therapist and your child interact with each other
- Speech productions your child makes during therapy
- Changes in your child's speech production during therapy, and how and when these occurred

To check the reliability of the analysis a small sample of your child's videos will also be analysed by one of Samantha's research colleagues in the Human Communication Sciences department at the University of Sheffield. No other information about your child would be given.

Samantha is expected to finish her PhD in March 2018. We will keep information about your child for five years after this, and then destroy it.

What are the possible advantages and disadvantages?

There are no specific advantages to you or your child. However, we hope the findings will be used to develop speech and language therapy with children with speech production difficulties. This might benefit your child if s/he has speech and language therapy in the future.

We do not know of any disadvantages. We will handle your child's videos and information securely to minimise any risks of loss and/or of it being seen by someone outside the research team.

Is there someone I can talk to about the study?

Yes, your child's Speech and Language Therapist (*name will be inserted*) can discuss and answer questions about the study. You can also talk with Samantha, who is carrying out the study, if you wish. If you want to speak with someone who is not involved in the study you can speak with the Lead Specialist Speech and Language Therapist (*name will be inserted*) at the Regional Cleft Lip and Palate Service that cares for your child. We have put all the contact details on the last page of this information sheet.

What if I have any concerns?

Please discuss these with your Speech and Language Therapist (*name will be inserted*) and/or Samantha in the first instance and they will do their best to resolve them. If you still have concerns you can contact either of Samantha's research supervisors (Dr Hilary Gardner and Dr Blanca Schaefer) at the University of Sheffield or Professor Patty Cowell, the Head of Department. If you still have concerns after this you will be given the contact details for the university registrar. You can also contact the Research and Innovation department at Leeds General Infirmary. If you wish to make a formal complaint you can do this through the Patient Advice and Liaison Service at Leeds Teaching Hospitals NHS Trust. We have put all the contact details on the last page of this information sheet.

Will my involvement in this study be kept confidential?

Yes. Samantha will access your child's information in line with ethical and legal practice guidelines and handle it in confidence. She will analyse your child's videos in secure places only. She may look at these with her supervisors at the university, who also have a duty of confidentiality to you and your child.

Your child's information will be kept strictly confidential and stored securely in the offices of the Northern and Yorkshire Regional Cleft Lip and Palate Service at Leeds General Infirmary. Electronic information will be stored securely on a password protected computer network and only Samantha will have access to this. When information is taken out of the hospital, it will not contain written personal details about your child, such as their name or date of birth, and the information-carrying device will be encrypted.

What will happen to the results of this study?

Samantha will write up an anonymised report of the study as her PhD thesis.

She will also share the anonymised results with speech and language therapists to help develop their therapy practice and other studies in this area. They might also be shared with other professionals, for

example, cleft lip and palate surgeons, who work with children born with cleft palate. Samantha will share the anonymised results by giving talks and writing papers and other publications.

You and your child will not be named in the thesis or any report or publication arising from it. When we present results from studies about speech and language therapy, and when we train students, we sometimes show video examples. We may ask your permission to use videos of your child in this way. You can still give us permission to watch your child's videos in the study even if you do not want us to use them when we present the results later.

Who has reviewed the study?

An NHS research ethics committee and the Research and Innovation departments at Leeds Teaching Hospitals NHS Trust and the University of Sheffield have given approval for this study to go ahead.

Contact details

Samantha Calladine
PhD Student (University of Sheffield)
Lead Specialist Speech and Language Therapist
Regional Cleft Lip and Palate Service (Leeds)
Paediatric Offices
F Floor, Martin Wing
Leeds General Infirmary
Leeds, LS1 3EX
0113 392 3876
07789 650837
s.calladine@sheffield.ac.uk
scalladine@nhs.net

NAME

Lead Specialist Speech and Language Therapist
Cleft Lip and Palate Centre
ADDRESS
ADDRESS
LANDLINE
MOBILE
[EMAIL](#)

NAME

Speech and Language Therapist
Speech and Language Therapy Department
ADDRESS
ADDRESS
LANDLINE
MOBILE
[EMAIL](#)

Dr Hilary Gardner, Research Supervisor
Dr Blanca Schaefer, Research Supervisor
Professor Patty Cowell, Head of Department
Department of Human Communication Sciences
University of Sheffield
362 Mushroom Lane
Sheffield
S10 2TS
Dr Hilary Gardner: 0114 222 2456
Dr Blanca Schaefer: 0114 222 2423
Professor Patty Cowell: 0114 222 2426
h.gardner@sheffield.ac.uk
blanca.schaefer@sheffield.ac.uk
p.e.cowell@sheffield.ac.uk

Ms Anne Gowing

Research Governance Manager
Leeds Teaching Hospitals NHS Trust
Research and Innovation Department
34 Hyde Terrace
Leeds
LS2 9LN
0113 392 0161
anne.gowing@nhs.net

Patient Advice and Liaison Service (PALS)

Leeds Teaching Hospitals NHS Trust
0113 206 7168 or 0113 206 6261
patientexperience.leedsth@nhs.net

Appendix S: Parent/Carer Consent Form



Parent/Carer Consent Form

(Version 1.0)

Participant identification number for this study:

Title of study: Multisensory input therapy for young children with cleft related speech patterns

Name of researcher: Samantha Calladine

Please read the statements below. Indicate whether you agree or disagree by drawing a circle round your response and writing your initials at the side. Please give a response to every statement. If you want to give your permission for your child's speech and language therapy information and videos to be included in this study you will need to agree with every statement on this page.

**Please circle as appropriate and
put your initials at the side**

- | | |
|--|----------------------|
| 1. I have read and understood the information sheet version 1.0 dated 15th August 2015 for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily. | I agree / I disagree |
| 2. I understand I do not have to agree to my child's information being used in this study. This is completely voluntary. If I give my permission and then change my mind, I can do this without it affecting my child's care in any way. | I agree / I disagree |
| 3. I give my permission for the researcher (Samantha) to access information about my child and her/his therapy, including videos. | I agree / I disagree |
| 4. I give my permission for Samantha's supervisors and a research colleague to see videos of my child receiving speech and language therapy, as described in the information sheet. | I agree / I disagree |
| 5. I give my permission for the researcher (Samantha) to store and to present/publish anonymised information produced from analysing my child's therapy information, as described in the information sheet. | I agree / I disagree |

Name of parent/carers

Date

Signature

Name of person taking consent

Date

Signature

Samantha Calladine

Parent/Carer Consent Form, Version 1.0

15th August 2015

Appendix T: Parent/Carer Consent Form: Videos



The Leeds Teaching Hospitals

Parent/Carer Consent Form:

Use of Videos in Presentations and Teaching

(Version 1.0)

Participant identification number for this study:

Title of study: Multisensory input therapy for young children with cleft related speech patterns

Name of researcher: Samantha Calladine

If you have completed the first consent form you will have given us your permission to use videos of your child in the above study. We would now like to ask for your permission to use these videos in presentations and teaching. We would remove your child's name and other obvious information that could identify your child and your family, for example, photographs, nursery logos on clothes.

This is completely optional. You do not have to give your permission for this. If you do not complete this second consent form we can still include your child's videos in the study but will not use them in presentations or teaching.

Please circle as appropriate and
put your initials at the side

- | | |
|--|----------------------|
| 1. I understand this is completely optional. If I do not give my permission for videos of my child to be used in presentations and teaching, I understand that the videos can still be used in the study, (if I have given permission for that). | I agree / I disagree |
| 2. I understand I can change my mind without having to give a reason and this will not affect my child's care in any way. | I agree / I disagree |
| 3. I understand the researcher (Samantha) will remove my child's name from the videos, and any other obvious information, without it affecting the images, e.g. photographs, nursery logos. | I agree / I disagree |
| 4. I give my permission for the researcher (Samantha) to use videos of my child in presentations and teaching. | I agree / I disagree |

Name of parent/carers

Date

Signature

Name of person taking consent

Date

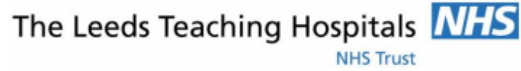
Signature

Samantha Calladine

Parent/Carer Consent Form (Videos), Version 1.0

15th August 2015

Appendix U: Child Data Collection Form



Child Data Collection Form (Version 1.0)

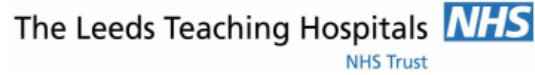
Title of study: Multisensory input therapy for young children with cleft related speech patterns

Name of researcher: Samantha Calladine

Participant identification number:	Date:
---	--------------

Characteristics:	
Cleft diagnosis	
Age at palate repair	
Secondary procedures? If yes, age and type	
Findings at 18-24 month assessment:	
Age at assessment	
Fistula present? If yes, size and location	
Receptive language skills	
Expressive language skills	
Resonance	
Airflow	
Phonetic inventory	
Phonological processes	
Cleft speech characteristics	
Multisensory input therapy information (18 months to 3 years):	
Age at first session	
Age at last session	
Number of sessions	
Number of data sets (videos + case notes) available	
Information about hearing:	
Hearing history	
Hearing management	

Appendix V: Therapist Data Collection Form



Therapist Data Collection Form (Version 1.0)

Participant identification number:

Title of study: Multisensory input therapy for young children with cleft related speech patterns

Name of researcher: Samantha Calladine

As a participant in this study we need some information from you about your previous training and experience and your current role. Please complete all sections, ticking all that apply.

1. How many years have you been qualified as a Speech and Language Therapist?

Less than 5		Between 5 and 15		More than 15	
-------------	--	------------------	--	--------------	--

2. What is your current role as a Speech and Language Therapist?

Link or Local Specialist		Regional Specialist	
--------------------------	--	---------------------	--

3. What training have you received on Multisensory Input Therapy (MSIT)?

Type	Tick if yes	Approximate dates, e.g. year(s)
Local or regional training event		
National training event, e.g. Clinical Excellence Network (CEN) study day		
Postgraduate level training		
Other, please describe here:		

4. What experience do you have of doing MSIT?

	Age of child	Approximate number of children		
		<5	5-10	>10
Children with cleft related speech patterns (with or without diagnosis of cleft palate)	<18 months			
	18 months to 3 years			
	3 years +			
Children with other types of speech patterns	<18 months			
	18 months to 3 years			
	3 years +			

Date completed:

Samantha Calladine

Therapist Data Collection Form, Version 1.0

28th June 2015

Appendix W: Transcription conventions (based on Jefferson, 2004, and Sidnell, 2010)

[
] Separate left-hand brackets, one above the other on successive lines, indicate an overlapping utterance or nonverbal action at the point where the overlap begins

(0.4) Numbers in parentheses indicate silence, marked in seconds and tenths of seconds, i.e. (0.4) is four tenths of a second

(.) A dot in parentheses indicates a brief interval of tenth of a second or less within or between utterances

(0.0) Italicised numbers in parentheses between line numbers indicate duration of talk excluded from the extract, shown in minutes and seconds. E.g. in the extract below, one minute and nine seconds of talk/action has been excluded between the end of the utterance/action in line 023 and start of the utterance/action in line 024

023 T It **fiel** ((gaze to camera)) **af:** ((gaze to the reel/string in C's hand))
(1.09)

024 T This one goes **f:** ((holds reel by mouth; gaze to camera))

oh: Colons indicate prolongation of the sound or syllable it follows; the longer the row of colons, the longer the prolongation

mine- A single dash indicates a cut-off to a word or part of a word

↑↓ Arrows indicate shifts in pitch; they are positioned immediately prior to the rise (upward arrow) or fall (downward arrow)

stress Underlining indicates emphasis

((nods)) Italicised double parentheses mark the author's descriptions of nonverbal action, e.g. ((holds picture up)) and direction of eye gaze, e.g. ((gaze to T))

() Empty parentheses indicate that the author was unable to decipher what was said

heləu Transcriptions in bold are phonetically transcribed using the IPA (2018) and ExtIPA (Ball et al., 2018)

text Grey highlighting indicates line of interest (therapist talk)

text Yellow highlighting indicates line of interest (child talk)

Abbreviations: C=child; T=therapist; M=mother; F=father; S=sister; Artic=articulation; A=activity