Speaking in harmony:
An exploration of the potential for rhythm and song to support speech production in four young adults with Down Syndrome.

Tracy Dawn Jeffery

A thesis submitted for the degree of
Doctor of Philosophy

Department of Human Communication Sciences

University of Sheffield

July 2016

Abstract

Engagement in music and singing may be a valuable tool to support or develop speech skills in people with DS, as it is for TD populations. However, to the author's knowledge, no study has yet examined in detail the musical abilities of people with DS, or whether the conditions that support the transfer of learning from the musical domain to the speech domain (Patel, 2011; 2014) may be met. This study explores the potential value of musical learning to support speech perception and production in the population.

A set of case studies with multiple measures were used to assess the cognitive, speech, vocal and musical abilities of four young adults with DS and SLD, who participated in four sessions of a six-week programme of song-based musical activities. Their progress in learning new songs was tracked during this time. Participants had physiological, cognitive and perceptual difficulties that were common to speech and musical domains. These included a low digit span and attentional difficulties, dysfluency, and evidence of dysphonia and diplophonia in speech and song. However, their voice quality, in particular, varied according to task demands.

Individuals made progress in mastering new songs and in imitating pitch and interval-matching tasks. They were able to transfer learning from songs to the spoken domain. Differences in the performance of individuals in both domains could be primarily attributed to previous learning, hearing status, and to task effects. Aspects of music-based learning appeared to support the perception and production of speech in the four participants. The study concluded that music activities might be an appropriate means of supporting aspects of speech development in people with DS. A programme of musical activities is recommended that could support the perceptual and productive speech difficulties that are common to the wider DS population.

Key words: Down Syndrome, voice quality, speech, singing, beat entrainment

Acknowledgements

Firstly, I would like to thank my supervisors, Professor Sandra Whiteside and Dr. Stuart Cunningham, for their help and support throughout the process of this work. I am especially grateful to Sandra, who kept me motivated when circumstances worked against me, and when the magnitude of the task seemed too overwhelming. I am also grateful for her thorough approach and attention to detail, and hope that some of these attributes are at last beginning to rub off.

This work would not have been possible without the funding of The University of Sheffield, to whom I am extremely grateful. I would also like to thank Dave Hernandes and Mint Leaf Software for kindly donating the Sonneta software. This enabled a much wider body of data to be examined in a shorter period of time, and made the analysis so much richer.

I would also like to thank Professor Joy Stackhouse, whose suggestion to consider a PhD., following completion of my MSc., sowed the seed for this project; Dr. Julia Lindley-Baker, whose support for her former staff and students enabled the seed to take root; the participants of the project, without whom this project would not exist; and to all my former students who inspired the work.

Although the learning journey that this thesis represents is mine alone, it has not been a truly solo endeavour. I would like to thank my friends for still being my friends despite my prolonged reclusive status; and my former and current work colleagues, who have provided invaluable support, encouragement, and inspiration. Thanks are also due to Jack the dog, and to Frances Kelly and the Horncastle Community Choir: walking the dog and singing are the only reasons I ever left the computer in the last two years, and both provided welcome respite and therapeutic diversions. Special thanks are due to my daughter, Damaris; to my partner, Graham, who proof-read the thesis and bore the brunt of my frustrations; to my mum, dad and sister; and to my former partner, Alan: all of whom believed in me whenever I faltered, and kept reminding me of how far I had come.

Finally, I would like to thank Dr. Hilary Gardner and Professor Graham Welch for their feedback and suggestions, and for making the last, and most daunting, stage of this work such an enjoyable and memorable process.

Author's Declaration

I declare that the work presented within this thesis is my own work and has not been previously submitted for any other degree or qualification.

Table of Contents

Abstract ii

Acknowledgements iii

Table of Contents iv

List of Tables x

List of Figures xiv

List of Appendices xix

Glossary of Terms xx

List of Acronyms xxi

Ext IPA and VOQs symbols xxii

Chapter 1: Introduction 1

1.1 Introduction and motivation 1

1.2 Context and purpose of the study 3

1.3 Introduction to the thesis 4

Chapter 2: Literature review 7

2.1 Music development 7

2.1.1 Development of musical ability between birth and five years 7

2.1.2 Development of musical ability during school years 11

2.1.3 Learning and development beyond middle school years 13

2.1.4 Section summary and relevance to the study 13

2.2 Music development and its relevance for speech and voice 13

2.2.1 Musical ability and the perception and production of speech 13

2.2.2 Therapeutic effects of singing on speech and voice production 16

2.2.3 Section summary and relevance to the study 19

2.3 Down Syndrome 19

2.3.1 Physiological characteristics 19

2.3.2 Peripheral sensory impairments 20

2.3.3 Motor development 21

2.3.4 Section summary and relevance to the study 23

2.4 Speech and voice in DS 23

2.4.1 Speech characteristics and intelligibility in DS 23

2.4.2 Voice production in healthy populations and in DS 24

2.4.3 Prosody 29

2.4.4 Section summary and relevance to the study 33

2.5 Intelligence and cognition in DS 33
2.5.1 Intelligence 33
2.5.2 Cognition and language 35
2.5.3 Section summary 39
2.6 Musicality in people with learning disabilities and DS 40
2.6.1 Musicality in people with developmental difficulties 40
2.6.2 Musicality in people with DS 41
2.7 Summary and hypotheses 44
Chapter 3: Methods 47
3.1 Study design and participants 47
3.1.1 Study design 47
3.1.2 Recruitment and selection of participants 48
3.2 The battery of assessments 48
3.2.1 Speech and cognitive tests 48
3.2.2 Musical tests 52
3.2.3 Within-session and repeated assessments 56
3.3 Recording and assessment procedures 56
3.4 Lesson design 57
3.4.1 Song-based musical development 57
3.4.2 Lesson design 58
3.5 Analysis of cognitive and musical performances 61
3.5.1 Verbal and nonverbal intelligence 61
3.5.2 Motor movements 61
3.5.3 Rhythm and beat entrainment 61
3.5.4 Rhythm and melody in songs 62
3.6 Procedures for analysis of speech data 65
3.6.1 Calculation of accuracy of speech sounds 65
3.6.2 Phonation and voice quality 68
3.6.3 Calculation and representation of prosody 74
3.6.4 Use of SPSS for voice data 78
3.6.5 Intra-rater reliability 78
3.7 Preparation of the case studies 80
3.7.1 Selection of data for inclusion in the case studies 80
3.7.2 Adaptations and additional methodology for case studies 80
3.7.3 Selection of examples of voice production from individual tests 81
3.7.4 Presentation of data in Figures 81
Chapter 4: Andrew 83
4.1 Summary of history, learning and performance in core assessments 83

4.1.1 Cognitive and perceptual abilities 83
4.1.2 Motor-music abilities 85
4.1.3 Oro-motor and phonological abilities 85
4.1.4 Voice range and measures of voice quality 87
4.1.5 Learning 88
4.2 Motor-music entrainment and imitation 88
4.2.1 Beat entrainment and extraction 88
4.2.2 Rhythm entrainment 89
4.2.3 Rhythm imitation 90
4.2.4 Syllable segmentation 91
4.3 Voice and prosody in imitation tasks 91
4.3.1 Sustained vowels 91
4.3.2 Pitch and interval matching 95
4.3.3 Words of increasing syllable length 96
4.3.4 Nonword imitation 103
4.3.5 Imitation of phrases 105
4.3.6 Song Imitation: ‘Magic Book’ 107
4.4 Prosody and voice in production of speech and song from LTM 109
4.4.1 Comments 109
4.4.2 ‘Happy Birthday’ in sung and spoken conditions 111
4.4.3 ‘Twinkle Twinkle Little Star’ 113
4.5 Discussion of rhythm, voice and melody in speech and song 116
4.5.1 Temporal perception and production in music 116
4.5.2 Temporal perception and production in speech and song 117
4.5.3 Voice quality and phonation 120
4.5.4 Pitch perception and production in song and speech 123
4.6 Links between musicality and speech 124
4.6.1 Difficulties common to speech and music domains 124
4.6.2 ‘Happy Birthday’: Evidence of transfer potential 125
4.7 Implications 127
Chapter 5: Kerry 131
5.1 Summary of history, learning and performance in core assessments 131
5.1.1 Cognitive and perceptual abilities 132
5.1.2 Motor-music abilities 133
5.1.3 Oro-motor and phonological abilities 133
5.1.4 Phonation and voice quality 134
5.1.5 Learning 135
5.2 Motor-music entrainment and imitation 136

5.2.1 Beat entrainment and extraction 136
5.2.2 Rhythm entrainment 137
5.2.3 Rhythm imitation 138
5.2.4 Syllable segmentation 139
5.3 Voice and prosody in imitation tasks 140
5.3.1 Sustained vowels 140
5.3.2 Pitch and interval matching 142
5.3.3 Words of increasing syllable length 144
5.3.4 Nonword imitation 149
5.3.5 Imitation of phrases 152
5.3.6 Song imitation: ‘Magic Book’ 153
5.4 Prosody and voice in production of speech and song from LTM 157
5.4.1 Spontaneous speech 157
5.4.2 ‘Happy Birthday’ in sung and spoken conditions 158
5.4.3 Reading 162
Discussion 164
5.5 Rhythm, voice and melody in speech and song 164
5.5.1 Temporal perception and production in music 164
5.5.2 Temporal perception and production in speech and song 166
5.5.3 Pitch perception and production in song and speech 167
5.5.4 Voice quality, phonation and cognitive load 168
5.6 Links between musicality and speech 171
5.6.1 Difficulties common to speech and music domains 171
5.6.2 ‘Happy Birthday’: evidence of transfer potential 172
5.7 Summary and implications 173
Chapter 6: Rachel 177
6.1. Summary of history, learning and performance in core assessments 177
6.1.1 Cognitive and perceptual abilities 177
6.1.2 Motor-music abilities 179
6.1.3 Oro-motor and phonological abilities 179
6.1.4 Voice range and measures of voice quality 180
6.1.5 Learning 180
6.2 Motor-music production 181
6.2.1 Beat entrainment 181
6.2.2 Rhythm entrainment 182
6.2.3 Syllable segmentation 182
6.2.4 Rhythm imitation 183
6.3 Voice and prosody in imitation tasks 185

6.3.1 Sustained vowels 185
6.3.2 Pitch and interval matching 187
6.3.3 Imitation of words of increasing syllable length 190
6.3.4 Nonword imitation 195
6.3.5 Phrase imitation 196
6.3.6 Song imitation: ‘Magic Book’ 197
6.4 Prosody and voice in production of speech and song from LTM 201
6.4.1 Comments 201
6.4.2 ‘Happy Birthday’ 203
Discussion 207
6.5 Rhythm, voice and melody in speech and song 207
6.5.1 Temporal perception and production in music 207
6.5.2 Temporal perception and production in speech and song 208
6.5.3 Phonation and voice quality 210
6.5.4 Pitch production and perception in song and speech 211
6.6 Links between musical and speech domains 212
6.6.1 Difficulties common to speech and music domains 212
6.6.2 ‘Happy Birthday’: evidence of transfer between song and speech 215
6.7 Summary and Implications 215
Chapter 7: Robert 219
7.1 Summary of history, learning and performance in core assessments 219
7.1.1 Cognitive and perceptual abilities 220
7.1.2 Motor-music abilities 220
7.1.3 Oro-motor and phonological abilities 221
7.1.4 Voice and phonation 223
7.1.5 Learning 224
7.2 Motor-music entrainment and imitation 224
7.2.1 Beat entrainment 224
7.2.2 Rhythm entrainment 224
7.2.3 Syllable segmentation and rhythm imitation 226
7.3 Voice and prosody in imitation tasks 227
7.3.1 Sustained vowels 227
7.3.2 Pitch and interval matching 230
7.3.3 Words of increasing syllable length 232
7.3.4 Imitation of sung phrases 237
7.4 Prosody and voice in production of speech and song from LTM 241
7.4.1 Comments to peers 241
7.4.2 ‘Happy Birthday’: transfer from song to speech 244

Discussion 247
7.5 The implications of Robert's moderate hearing loss for motor control and auditory perception 247
7.5.1 Motor movements 247
7.5.2 Speech perception 247
7.5.3 Music perception 248
7.6 Rhythm, voice and melody in speech and song 248
7.6.1 Temporal perception and production in music tasks 249
7.6.2 Temporal perception and production in speech 250
7.6.3 Phonation and voice quality 251
7.6.4 Pitch production and perception in speech and song 253
7.7 Learning and potential links between musical and speech domains 254
7.7.1 Learning of speech sounds, word rhythm and pitch 254
7.7.2 'Happy Birthday': long-term learning and transfer between song and speech 256
7.8 Summary and implications 257
Chapter 8: Summary and conclusions 261
8.1 Summary of the key findings 262
8.1.1 Key abilities (Table 8.1) 262
8.1.2 Key difficulties (Table 8.2) 264
8.1.3 Progress in song-learning (Table 8.4) 271
8.1.4 Transfer of learning between song and speech (Table 8.5) 273
8.2 General discussion 273
8.2.1 Key abilities and difficulties within the group 273
8.2.2 Key factors affecting rhythm, melodic intonation and phonation 278
8.2.3 Song learning 282
8.2.4 Transfer between domains 283
8.3 Implications for the application of findings 282
8.3.1 General implications 284
8.3.2 Implications for policy 285
8.3.3 Recommended music activities 288
8.4 Critical evaluation of the research and future direction of research 291
8.4.1 Positive aspects of the research project 292
8.4.2 Limitations of the study 294
8.4.3 Recommendations for further research 295
8.4.4 Conclusion 296
Appendices 299
References 368

List of Tables

Table 3.1	Description of prosody measures used in Prosogram.	75
Table 3.2	Results of the intra-rater reliability test scores.	79
Table 4.1	Profile: a summary of Andrew's performance in core tasks.	84
Table 4.2	Error patterns produced by Andrew in the DEAP task.	86
Table 4.3	The percentage of speech errors produced by Andrew in core speech tasks.	86
Table 4.4	The pitch range of Andrew's vocal glides and connected speech.	87
Table 4.5	Mean voice perturbation measures for Andrew in sustained vowels and vowels in words.	87
Table 4.6	The accuracy of Andrew's timing in beat entrainment and extraction tasks.	89
Table 4.7	The accuracy of Andrew's timing in three bars of rhythm entrainment, following a visual demonstration.	90
Table 4.8	Andrew's accuracy in imitating clapped rhythms in the TRaCol task (Treharne, 1999).	90
Table 4.9	The mean fundamental frequency and voice perturbation measures for Andrew's sustained vowels.	92
Table 4.10	A summary of the segmental, prosodic and vocal features of Andrew's imitation of words in the Syllables test, arranged by number of target syllables.	97
Table 4.11	Voice data for the Syllables test, arranged by the target number of syllables.	100
Table 4.12	Mean voice data for Andrew's Syllables test, arranged by the position of the vowel in the word (vertical), and according to the number of target syllables in the word (horizontal).	100
Table 4.13	Prosodic features of the demonstration and of Andrew's imitation of non words (CNREP).	104
Table 4.14	Mean voice perturbation measures of Andrew's vowels in CNREP, arranged according to the number of syllables in the target word.	104
Table 4.15	Mean voice perturbation measures of Andrew's vowels in CNREP, arranged according to the position of the syllable in the word.	105
Table 4.16	Prosodic features of the demonstration and of Andrew's imitation of phrases (Newton, 1999).	105
Table 4.17	Articulatory and musical features of Andrew's production of the first phrase of 'Magic Book'.	109
Table 4.18	The relative duration and pitch of Andrew's syllables in 'Happy Birthday', in sung and spoken conditions.	111

Table 5.1	Profile: a summary of Kerry's performance in core tasks.	132
Table 5.2	A summary of Kerry's speech errors in the DEAP task.	133
Table 5.3	The percentage of speech errors produced by Kerry in core speech tasks.	134
Table 5.4	The mean pitch and pitch range of Kerry's vocal glides and connected speech.	134
Table 5.5	Mean voice perturbation measures for Kerry in sustained vowels and vowels in words.	135
Table 5.6	The accuracy of Kerry's timing in beat entrainment and extraction tasks.	136
Table 5.7	The accuracy of Kerry's timing in two bars of rhythm entrainment.	138
Table 5.8	Kerry's accuracy in imitating clapped rhythms in the TRaCoL task (Treharne, 1999).	139
Table 5.9	The mean fundamental frequency and voice perturbation measures for Kerry's sustained vowels.	140
Table 5.10	A summary of the segmental, prosodic and vocal features of Kerry's imitation of words in the Syllables test, arranged by number of target syllables.	144
Table 5.11	Voice perturbation measures in the Syllables task according to the number of syllables produced by Kerry and the position of the vowel in the word.	144
Table 5.12	Prosodic features of the demonstration of non words (CNREP) and of Kerry's imitation.	149
Table 5.13	Mean voice perturbation measures of Kerry's vowels in CNREP, arranged according to the number of syllables in the target word.	150
Table 5.14	Phonetic and prosodic features of the demonstration and of Kerry's imitation of phrases (Newton, 1999).	152
Table 5.15	Articulatory and musical features of the demonstration of the first phrase of 'Magic Book', and of Kerry's first and final productions.	157
Table 5.16	The relative pitch and relative duration of syllables in 'Happy Birthday' in Kerry's sung and spoken conditions, taken from the same recording session.	161
Table 6.1	Profile: a summary of Rachel's performance in core tasks.	178
Table 6.2	A summary of Rachel's speech errors and error patterns in the DEAP task.	179
Table 6.3	The percentage of speech errors produced by Rachel in core speech tasks.	180
Table 6.4	The mean pitch and pitch range of Rachel's vocal glides on /la/ and connected speech.	180
Table 6.5	Mean voice perturbation measures for Rachel's sustained vowels and vowels in words.	181

Table 6.6	The accuracy of Rachel's timing in beat entrainment and extraction tasks, at 104 bpm, 132 bpm, and 160 bpm.	182
Table 6.7	The number and relative duration of claps produced by Rachel in syllable segmentation tasks.	182
Table 6.8	Rachel's accuracy in imitating clapped rhythms in the TRaCoL task (Treharne, 1999).	183
Table 6.9	The mean fundamental frequency and voice perturbation measures for Rachel's sustained vowels.	185
Table 6.10	Rachel's tonal production when 'reading' a 5-note descending scale.	190
Table 6.11	The prosodic profile of words in the Syllables test, as demonstrated to Rachel and as imitated by her.	190
Table 6.12	Mean voice perturbation measures for words of increasing syllable length.	191
Table 6.13	Voice production and speech production in words of increasing syllable length.	191
Table 6.14	Phonetic and prosodic features of the demonstration of non words (CNREP) and of Rachel's imitation.	195
Table 6.15	Prosodic profile of the demonstration of non words (CNREP) and of Rachel's imitation.	196
Table 6.16	Mean voice perturbation measures for Rachel's vowels in CNREP, arranged according to increasing target syllable length.	196
Table 6.17	Phonetic and prosodic features of the demonstration and of Rachel's imitation of phrases (Newton, 1999).	197
Table 6.18	Prosodic profile of phrase imitation (Newton, 1999).	197
Table 6.19	Articulatory and musical features of the demonstration of the first phrase of 'Magic Book', and of four of Rachel's productions.	199
Table 6.20	Prosodic profiles of Rachel's spontaneous comments and of her spoken and sung productions of 'Happy Birthday'.	201
Table 6.21	Relative timing of syllables in Rachel's productions of 'Happy Birthday' in sung and spoken conditions.	206
Table 7.1	Profile: a summary of Robert's performance in core tasks.	221
Table 7.2	Speech errors and error patterns produced by Robert in the DEAP task.	222
Table 7.3	Percentage of speech errors produced by Robert in core speech tasks.	222
Table 7.4	Mean pitch and range of Robert's vocal glides on 'la', and in connected speech.	223
Table 7.5	The mean voice perturbation measures of Robert's sustained vowels and vowels in words.	223
Table 7.6	The accuracy of Robert's timing in beat entrainment tasks.	225

Table 7.7	The number and relative duration of claps produced by Robert in syllable segmentation tasks.	226
Table 7.8	The mean fundamental frequency and voice perturbation measures for Robert's sustained vowels.	227
Table 7.9	The prosodic profile of words in the Syllables test, as demonstrated to Robert and as imitated by him.	232
Table 7.10	A summary of the segmental, prosodic and vocal features of Robert's imitation of words in the Syllables test, arranged by number of target syllables.	232
Table 7.11	Mean voice data for Robert's imitations of words in the Syllables test, arranged by number of target syllables.	233
Table 7.12	The segmental, articulatory, melodic and rhythmic features of the first phrase of 'Magic Book' as demonstrated, and as produced by Robert during his first and final productions.	239
Table 7.13	Robert's phonetic production of 'Monkey Man'.	239
Table 7.14	The relative duration and pitch of Robert's syllables in 'Happy Birthday', in sung and spoken conditions.	245
Table 8.1	Group profile: a summary of data from key tasks.	261
Table 8.2	Group profile of key physiological, perceptual and behavioural difficulties demonstrated in speech and music-related tasks.	263
Table 8.3	The relationship between auditory memory demand and performance in speech and music tasks.	268
Table 8.4	A summary of the group's progress in learning 'Magic Book'.	269
Table 8.5	Transfer of knowledge from song to speech of the words to the song 'Happy Birthday'.	270

List of Figures

Figure 3.1	Screenshot of the method used in Praat for measuring timing accuracy in a beat entrainment task.	61
Figure 3.2	Notation for the opening phrase to ‘Happy Birthday’.	62
Figure 3.3	Method of calculation of duration and pitch in ‘Happy Birthday’.	63
Figure 3.4	Method used in Pages to generate a graph of pitch and rhythm of a song produced by a participant (‘Andrew’, in green) against the target pitch and rhythm of a song (in blue).	64
Figure 3.5	Screenshot showing the calculation of pitch changes in semitones, using Praat.	65
Figure 3.6	Screenshot showing the analysis of speech sounds in Praat, using broad-band spectrogram.	66
Figure 3.7	Method used for the calculation of speech errors in words of one syllable (DEAP).	66
Figure 3.8	Screenshots of the word ‘three’, in Praat and Sonneta.	68
Figure 3.9	Jitter and Shimmer.	70
Figure 3.10	Screenshots showing the method used for voice measurement of sustained vowels, in Sonneta.	71
Figure 3.11	Screenshots showing the method used for voice measurement of voiced and devoiced vowels in words, in Sonneta.	72
Figure 3.12	Screenshot of Praat showing the calculation of interharmonics present in sustained vowels.	73
Figure 3.13	Narrow-band spectrogram to reveal interharmonics in words (window length of 15 ms, frequency range is 0-1500 Hz, dynamic range is 35 dB).	74
Figure 3.14	Screenshot of Praat showing the measurement of articulatory features for ‘Magic Book’.	77
Figure 4.1	Image showing Andrew’s rhythmic production in the Rhythm Entrainment task – *tee-tee tee-tee ta ta*.	89
Figure 4.2	Annotated waveform showing Andrew's production of a rhythmic sequence from visual notation.	91
Figure 4.3	Waveform and annotated spectrogram showing Andrew’s sustained vowel, /u/3.	93
Figure 4.4	Waveform and annotated spectrogram showing Andrew’s sustained vowel /a/1 at 97 Hz.	94
Figure 4.5	Graph showing Andrew’s accuracy when matching melodic intervals during her first and final assessments.	95
Figure 4.6	Graph showing Andrew’s accuracy when matching melodic intervals during his first and final assessments.	95

Figure 4.7	Annotated spectrogram showing Andrew singing a musical scale on ‘la’ from a visual score.	96
Figure 4.8	Annotated spectrogram showing the stimulus (above) and Andrew's imitation of words of two-syllables.	98
Figure 4.9	Annotated spectrogram showing the stimulus (above) and Andrew's imitation of words of three syllables.	99
Figure 4.10	Waveform and annotated spectrogram showing Andrew’s imitation of ‘jab’.	101
Figure 4.11	Waveform and annotated spectrogram showing Andrew’s imitation of ‘jabber’.	102
Figure 4.12	Waveform and annotated spectrogram showing Andrew’s imitation of ‘jabbering’.	103
Figure 4.13	Annotated spectrogram showing Andrew’s imitation of ‘perplisteronk’ (CNREP).	106
Figure 4.14	Annotated spectrogram showing Andrew’s imitation of ‘contramponist’ (CNREP).	106
Figure 4.15	Waveforms and annotated spectrograms showing Andrew’s imitation of ‘My uncle is a farmer’ (Newton, 1997).	107
Figure 4.16	Phonetic transcription and syllabic stress of Andrew’s first and final production of ‘Magic Book’, in week one (prompted) and week five (solo).	107
Figure 4.17	Graph showing the development in Andrew’s recall of the words and melody of ‘Magic Book’.	108
Figure 4.18	Phonetic transcriptions and syllabic stress of Andrew’s spontaneous comments, taken from session recordings.	109
Figure 4.19	Waveform and annotated spectrogram showing Andrew’s comment to Kerry (Comment 1).	110
Figure 4.20	Phonetic transcription and syllabic stress of Andrew’s sung and spoken productions of ‘Happy Birthday’.	112
Figure 4.21	Waveform and annotated spectrogram showing phrase 4 of Andrew’s spoken production of the song ‘Happy Birthday'.	113
Figure 4.22	Waveforms and spectrograms showing detail of Andrew’s spoken ‘Happy Birthday’.	114
Figure 4.23	Phonetic transcriptions and syllabic stress of Andrew’s production of ‘Twinkle Twinkle Little Star’.	115
Figure 4.24	Graph showing Andrew’s accuracy in terms of melodic shape when singing ‘Twinkle Twinkle Little Star’ to an accompaniment.	115
Figure 5.1	Image showing Kerry’s rhythmic production during the Beat Entrainment task at 132 bpm.	137
Figure 5.2	Image showing Kerry’s rhythmic production in the Rhythm Entrainment task – *ta tee-tee ta ta*.	138
Figure 5.3	Image showing the duration and timing of Kerry’s claps, relative to her syllables when singing ‘Magic Book’.	139

Figure 5.4	Waveform and annotated spectrogram showing Kerry's /a/ vowel at G4 (/a/1).	141
Figure 5.5	Detail of the waveform from Kerry's /a/ vowel at G4 (/a/1), section 3.	142
Figure 5.6	Graph showing Kerry's accuracy when matching pitches during initial and final assessments.	143
Figure 5.7	Graph showing Kerry's accuracy when matching melodic intervals during her first and final assessments.	143
Figure 5.8	Waveform and annotated spectrogram showing the stimulus (above) and Kerry's imitation of three words of two-syllables.	145
Figure 5.9	Waveform and annotated spectrogram showing Kerry's imitation of 'love'.	146
Figure 5.10	Waveform and annotated spectrogram showing Kerry's imitation of 'loving'.	147
Figure 5.11	Waveform and annotated spectrogram showing Kerry's imitation of 'lovingly'.	148
Figure 5.12	Annotated spectrogram showing Kerry's imitation of 'doppelate'.	151
Figure 5.13	Waveforms and annotated spectrograms showing Kerry's imitation of 'Clare ate all her lunch'.	153
Figure 5.14	Phonetic transcription and syllabic stress of Kerry's first, third and final productions of 'Magic Book'.	155
Figure 5.15	Graph showing the development in Kerry's recall of the words and melody of 'Magic Book', over a five-week period.	155
Figure 5.16	Annotated spectrograms showing Kerry's first production of 'Magic Book'.	156
Figure 5.17	Waveform and annotated spectrogram showing Kerry's comment to a peer: 'really good'.	158
Figure 5.18	Graph showing the melodic shape of Kerry's productions of 'Happy Birthday' (sung).	159
Figure 5.19	Phonetic transcription and syllabic stress of Kerry's sung and spoken productions of 'Happy Birthday'.	159
Figure 5.20	Figure 5.20: Annotated spectrogram showing Kerry's production of 'Happy Birthday', spoken.	160
Figure 5.21	Figure 5.21: Annotated spectrogram showing Phrase 2 of Kerry's production of 'Happy Birthday', spoken.	160
Figure 5.22	Annotated spectrogram showing Phrase 3 of Kerry's production of 'Happy Birthday', spoken.	161
Figure 5.23	Annotated spectrogram showing Kerry's production of Phrase 1 of the reading task.	162
Figure 5.24	Waveforms of the words 'I' and 'like', as produced by Kerry in the reading task.	163

Figure 6.1	Annotated waveform showing Rachel's ability to drum word-based syllabic rhythms.	184
Figure 6.2	Waveform and annotated spectrogram showing Rachel's /i/ vowel at E4 (/i/1).	186
Figure 6.3	Waveform and annotated spectrogram showing Rachel's /a/ vowel at G4 (/a/4).	187
Figure 6.4	Graph showing Rachel's accuracy when matching pitches during initial and final assessments.	188
Figure 6.5	Graph showing Rachel's accuracy when matching melodic intervals during first and final assessments.	188
Figure 6.6	Annotated spectrogram showing changes in Rachel's pitch and voice quality when reading a 5-note scale from visual notation.	189
Figure 6.7	Annotated spectrograms showing the demonstration and Rachel's imitation of words of three syllables.	192
Figure 6.8	Annotated spectrograms showing Rachel's imitation of the words 'please', 'pleasing' and 'pleasingly'.	193
Figure 6.9	Waveform and annotated spectrogram showing a detail of Rachel's imitation of 'please'.	194
Figure 6.10	Graph showing the development in Rachel's recall of the words and melody of 'Magic Book', over a five-week period.	198
Figure 6.11	Phonetic transcription and syllabic stress of Rachel's first and final productions of 'Magic Book'.	198
Figure 6.12	Annotated waveform and spectrogram showing Rachel's production of 'Magic Book' (R3).	200
Figure 6.13	Phonetic transcription and syllabic stress of Rachel's comment to Kerry (Comment 1).	201
Figure 6.14	Waveform and annotated spectrogram showing Rachel's comment to Kerry (Comment 1).	202
Figure 6.15	A phonetic transcription of Comment 2, showing Rachel's comments on Robert's singing (R1-5) and the researcher's comments (T).	202
Figure 6.16	Waveforms and annotated spectrograms showing Rachel's Comment 2, part 1 (a: top) and part 2 (b: below).	203
Figure 6.17	Phonetic transcription and syllabic stress of Rachel's production of 'Happy Birthday', sung and spoken.	204
Figure 6.18	Waveform and annotated spectrograms showing details of Rachel's spoken production of 'Happy Birthday'.	205
Figure 6.19	Waveform and annotated spectrograms showing details of Rachel's sung production of 'Happy Birthday'.	206
Figure 7.1	Graph showing the target ITI and the ITI of Robert's production in part 1 of the Rhythm Entrainment task - ta tee-tee ta ta.	225
Figure 7.2	Graph showing the target ITI and the ITI of Robert's production in part 2 of the Rhythm Entrainment task - tee-tee tee-tee ta ta - with visual support.	226

Figure 7.3	Waveform and annotated spectrogram showing Robert's sustained vowel /a/2 at G4.	228
Figure 7.4	Waveform and annotated spectrogram showing Robert's sustained vowel /a/1 at G3.	229
Figure 7.5	Waveforms and annotated spectrogram showing a detail of Robert's sustained vowel /a/1.	230
Figure 7.6	Graph showing Robert's accuracy when matching pitches during initial and final assessments.	231
Figure 7.7	Graph showing Robert's accuracy when matching melodic intervals during initial and final assessments.	231
Figure 7.8	Annotated spectrograms showing Robert's imitation of words of two syllables (above) and the demonstration (below).	234
Figure 7.9	Annotated spectrograms showing Robert's imitation of words of three syllables (above) and the demonstration (below).	235
Figure 7.10	Waveform and annotated spectrogram showing Robert's imitation of 'hope'.	236
Figure 7.11	Waveform and annotated spectrogram showing Robert's imitation of 'hopeful'.	236
Figure 7.12	Waveform and annotated spectrogram showing Robert's imitation of 'hopefully'.	237
Figure 7.13	Phonetic transcription and syllabic stress of Robert's production of 'Magic Book', over a four-week period.	237
Figure 7.14	Graph showing the development of Robert's recall of the words and melody of 'Magic Book', over a four-week period.	238
Figure 7.15	Graph showing Robert's melodic production of 'Monkey Man', with verbal prompts (R3) and with Makaton prompts only (R4).	240
Figure 7.16	Waveform and image showing details of Robert's production of 'Monkey Man', phrase 4.	241
Figure 7.17	Phonetic transcription and syllabic stress of Robert's comments to peers: comments 1 and 2.	241
Figure 7.18	Waveform and annotated spectrogram showing Robert's 'Comment 1'.	242
Figure 7.19	Waveform and annotated spectrogram showing Robert's 'Comment 2'.	242
Figure 7.20	A phonetic transcription of Robert's 'Comment 3', showing the interaction between the researcher (T) and Robert (R).	243
Figure 7.21	Waveform and annotated spectrogram showing Robert's 'Comment 3'.	243
Figure 7.22	Waveform and annotated spectrogram showing a detail of Robert's 'Comment 3'.	244
Figure 7.23	Phonetic transcription and syllabic stress of Robert's sung and spoken productions of 'Happy Birthday'.	244

Figure 7.24 Happy Birthday, spoken and sung. 246

Figure 8.1 A graph showing the mean perturbation measures produced by individuals in sustained vowels. 265

Figure 8.2 A graph showing the mean voice perturbation measures for whole sustained vowels and vowels in words of 1,2 and 3 target syllables. 266

List of Appendices

Appendix 1 Ethics documentation 300

Ethics consent 301

Letters and information sheets for participants 304

Letters and information sheets for parents/carers 310

Appendix 2 Sample test materials for assessment of music production and music and speech perception 323

Music production skills 324

Pitch production — visual support 326

Rhythm production and syllable segmentation — visual support 327

Test 13: Minimal pairs perception test 328

Test 26: Pitch and interval perception test 332

Test 27: Rhythm perception test 333

Test 27: Same/different pitch perception tests 335

Beat entrainment task 1 337

Rhythm entrainment task 2 338

Appendix 3 Teaching sessions, evaluations and teaching materials 339

Six-week teaching plan 340

Session plans and notes 342

Repertoire: sample teaching songs and visual aids 356

Appendix 4 Summary of research for MSc.: *An investigation into the effects of singing tuition on aspects of intelligibility in young people with speech impairments and learning difficulties.* 359

Glossary of Terms

Beat heirarchy The beat in music is not always explicit and is therefore inferred: as such the beat is inferred with reference to repeating rhythmic structures. There may be several repeating rhythmic structures within one main musical phrase, each with a dominant, or accented beat. Although one recurring beat will typically be more strongly accented (usually coinciding with the metrical accent: see **meter**), it is possible to perceive the beat at more than one level.

Digit span A commonly-used measure of verbal **short term memory**. The test requires individuals to repeat a string of numbers that are spoken at an even rate. The longest series of items they can correctly repeat, in the same order, is taken as an indicator of **short term memory**.

Diplophonia The perceived perception of two simultaneous pitches.

Entrainment In music, this refers to the ability to align limb-motor movements to an auditory temporal stimulus (e.g. a beat or rhythm).

Formants Resonant frequencies that are produced as a result of the length and shape of the vocal tract. The formants are affected by the relative size, shape and position of the articulators (jaw, tongue) and palate. (Typically, the first formant (F1) is related to the degree of opening at the front of the oral cavity; and the second formant (F2) is associated with the anterior-posterior tongue position).

Fundamental frequency (F0) The fundamental frequency (F0) is the lowest frequency at which a periodic waveform repeats. The perceptual correlate of F0 is **pitch**: typically a faster F0 is perceived as a higher pitch; however, the timbre of a sound can also affect the perceived pitch.

Harmonics-to-noise ratio (HNR) A measurement of the proportion of periodic signal generated by the vocal cords in comparison to any aperiodic signal, such as that generated by turbulence.

Interharmonics Multiple harmonics in the voice spectrum that are present between the dominant harmonics (**fundamental frequency**). These are visible on the spectrogram as horizontal bands, and are typically of lower intensity than the F0.

Isochronous In music, an isochronous (from the Greek for ‘equal’ and ‘time’) beat is one in which the beats are spaced equally.

Jitter A measure of irregularities in the frequency of the vibration of the vocal folds. Relative jitter is the average difference between successive periods, divided by the duration of the average period.

Melodic contour In singing, the melodic contour refers to the perceived ascending and descending patterning of **pitch** within a song or phrase.

Meter In music, meter is a rhythmic pattern of strong (accented) and weak (unaccented) beats, in which the accent on a regularly occurring beat leads to perceived temporal grouping.

Phonation Phonation is the act of producing a vocal sound via the quasi-periodic vibration of the vocal folds. The vocal folds are muscles at the base of the larynx that are capable of vibrating. According to myoelastic and aerodynamic theory, the vocal folds are set in motion by the application of sufficient breath pressure, which causes them to open. The intrinsic muscles of the larynx cause the vocal folds to come together (adduct), allowing sub-glottal breath pressure to build up until the vocal folds are pushed open again (the ‘Bernoulli effect). The cycle of regular opening and closing creates a periodic vibration, which determines **fundamental frequency** and contributes to perceived **pitch**. The nature and quality of phonation contribute to perceived **voice quality**.

Pitch Pitch in music and speech is a perceptual correlate of **fundamental frequency** (F0) - the rate of vibrations that of the source that produces it. In musical melodies, pitches are relative to each other and are described as ‘higher’, ‘unison’ (the same), or ‘lower’.

Glossary of Terms

Pulse	The pulse is the regularly recurring beat, or tactus, of a piece of music or song. Although beat is heirarchical (see **beat heirarchy**), the pulse is typically that which a listener taps their foot to.
Shimmer	A measure of the average short-term variations in the peak-to-peak amplitude of the consecutive voicing cycles, expressed as a percentage of the average amplitude.
Short term memory (STM)	A theoretical store that can hold a limited amount of information temporarily. It is a component of working memory. Auditory STM is commonly measured by **digit span.**
Subharmonics	A subset of interharmonics, consisting of a low-frequency harmonic that occurs between two consecutive dominant harmonics (the f**undamental frequency**). If the subharmonic approaches the intensity of the F0, this can result in diplophonia.
Tonality	In music or song, the tonality relates to the musical key, in which musical **pitch** is arranged according to an accepted heirarchy of perceived relationship and stability relative to a central note or tone (tonal centre).
Voice quality	The perceived attributes of phonation and resonance. Voice quality depends upon laryngeal anatomy, the physiology and configuration of the vocal tract, and upon learned behaviours. It is also affected by emotional state of the speaker, and by hormonal or neurological actors that can affect the function of the components.
Working memory (WM)	In working memory, information held in short term memory is manipulated by multiple components. According to Baddeley and Hitch's model (1974), different modalities of information are temporarily held and processed in sensory-specific stores (e.g. verbal-phonological, visual-spatial). In their model, working memory is assisted by the central executive, which regulates and directs attention; and by long term memory, which can help reduce the processing required in working memory.

List of Acronyms

ADHD	Attention Deficit and Hyperactivity Disorder
CA	Chronological Age
CPP	Cepstral Peak Performance
DS	Down Syndrome
DUV	Degree of Unvoiced Frames
DVB	Degree of Voice Breaks
HI	Hearing Impairment
HNR	Harmonics to Noise Ratio
MA	Mental Age
MIT	Melodic Intonation Therapy
MTD	Muscle Tension Dysphonia
SLD	Severe Learning Disability
STM	Short Term Memory
TD	Typically Developing

Ext IPA Symbols	
˞	Nasal escape
↓	Ingressive airflow
(	Initial partial voicing/devoicing
)	Final partial voicing/devoicing
.	Whispery voiced

VoQs : Voice Quality Symbols	
C	Creak
C̥	Whispery creak
V̰	Creaky voice
V!	Harsh voice
V!!	Ventricular phonation
V̰!!	Diplophonia
V̥!!	Whispery ventricular phonation
V̤	Breathy voice

Chapter 1: Introduction

James hurried through the open door and stationed himself in the middle of the small music room. The room was in disarray following the afternoon's 'music for communication' session, that had finished just moments before; it was a shambles of instruments and chairs, but it was empty of people, other than me. Muttering cheerfully to himself, James arranged a selection of musical equipment in a circle about him: around his feet he placed four flat, brightly-coloured switches that were connected via leads to a midi sound controller; to either side of him he positioned the two Soundbeams *- small, red devices that each emitted an ultrasonic beam extending 1.5 meters into the room, that were also connected to the midi controller. He picked up a battered African* djembe, *and said to himself, 'ah, that's it'. Cradling the drum in his left arm, he beat a pattern upon its skin with his right hand. As he drummed, he twisted and swayed so that his movements interrupted in turn each invisible beam at different points along their lengths, triggering tuneful notes and arpeggios of synthesised orchestral sounds. As he tapped his feet on each floor switch, he produced the sounds of a timpani and became a one-man orchestra: composer, conductor, rhythm section and symphony. For more than ten minutes James just played, smiling and engrossed, the centre of his own musical universe. Then he put down the drum and left without a word, returning to his footbath and bubbles in the room next door.*

> *…it is evident what an influence music has over the disposition of the mind, and how variously it can fascinate it: and if it can do this, most certainly it is what youth ought to be instructed in.*
>
> Aristotle (Politics, Book VII)

1.1 Introduction and motivation

At the start of this study, I was working as a music educator for young adults with learning disabilities. Of the hundred or so students with moderate-severe learning disabilities with whom I once regularly created music, James, a 22 year-old with Down Syndrome (DS) and Autism, was the most musically gifted. He responded to music with evident emotional sensitivity: he could listen to classical music with a smile, tears rolling down his cheek; he would sit for hours with his feet in a scented bath, headphones on, singing indecipherable sounds to Disney songs and pausing only to blow bubbles through a wand, and kisses to staff as they worked around him and tried to guess the song; and when we moved to a larger teaching space, he would lead peers and staff in an extended rhythmic improvisation. From his lone position at a table on the outskirts of the group, he would control with exquisite musicality the dynamics, pace and structure of the group's playing through his own drumming. It seemed evident that James had a musical intelligence that was in advance of his reported mental age of 2.5 years.

James was one of a group of students who changed my teaching practice and inspired this study. For three years prior to their arrival, my students had improvised and composed music, written

songs and developed their singing and vocal skills. Although they made music, their progress was measured against personal targets in social and emotional learning and communication, not against musical skills. However, two significant changes then occurred that left me struggling to meet the learners' needs. Firstly, the demographic changed. My groups became increasingly diverse and their needs more complex: now, the majority of my students were pre-verbal, with multiple diagnoses of learning difficulties such as autism and ADHD, alongside severe learning disabilities (SLD); many were unable to tolerate the noise of an active music session; some - like James - struggled to participate directly in the social activities, or even to enter the busy teaching space. Secondly, a change in college policy meant that my role shifted from an emphasis on music as a form of therapy to music education. So, having had no musical training myself, I enrolled on a local course entitled 'how to be brilliant at teaching music' (Brindle, 2005), a six-week training programme in the Kodaly approach. The method is entirely inclusive: although singing and music literacy are at its core, activities are visual and based on movement and games; they can even be silent. Furthermore, the Kodaly approach is developmental, and is itself based on observations of how children develop language (Tacka and Houlahan, 2008). The method allowed for my nonverbal, noise-sensitive learners to take part in active learning, and enabled me to easily measure and track their musical progress.

Over the next two years, I began to learn what my students might be capable of. My non-verbal students led singing games and demonstrated learning of pitch through gesture and visual images; most learned to compose and play rhythms based on simplified music notation; and some also learned to play colour-coded melodic instruments. As an ensemble, we composed and played music that was mostly in time and in tune. The advanced learners were able to sight-read rhythmic music and perform it to backing tracks. For the support staff who joined the groups for their first time, and who were as musically illiterate as I had been a year or two earlier, it seemed like magic.

However, I never learned what my students were truly capable of. Further policy changes and funding cuts meant that music no longer had a place in the college curriculum. The course ended and I left the college without really understanding the full potential of the students, nor knowing what distinguished my musically-gifted pupils with SLD, such as James, from my higher-ability but arhythmic pupils. Neither did I know whether the 'music for communication' curriculum was ever actually justified in its ambition to support and develop the non-verbal or verbal communication skills of my students. This study was devised in order to address some of this ignorance. Specifically, it sought to understand the musical abilities of one small group of young adults who have one type of learning disability in order to learn how and why they differ in their musicality, and what difficulties they face in terms of musical development. It also aimed to learn to what extent such activities might also support them in their verbal communication.

1.2 Context and purpose of the study

The notion that music is a form of communication is most pertinent to this study. The close link and overlap between music and communication is an ancient and distinguished notion. The Athenian musicologist, Damon (5th Century BCE), explained that rhythms arose from movements of the body in various states of emotion, that were 'filtered' by individual character types; he said that musical instruments emulate the many rhythms and cadences of speech, and that this gave rise to music (West, 1992). Two millenia later, archaeologist Steven Mithen (2006) argued the converse: that music - in the form of song - evolved as a form of emotional contact, a substitute for physical contact that retains and nurtures the emotional bond. Thus music is communication; furthermore, it is from music, Mithen argues, that speech evolved. The links between music and spoken language, in particular, has been the subject of research interest for many years (Patel, 2008) but researchers are just beginning to understand how musical skills relate to linguistic skills. Music, like verbal communication, requires the ability to attend to and discriminate spectral properties of sound and to perceive and track how these change over time. Several decades of research have shown that in comparison to non-musicians, musicians have better linguistic skills: they consistently demonstrate superior auditory memory, higher vocabulary scores, better executive function and are better able to perceive and learn new words. Even short periods of musical training lead to improvements to auditory memory, verbal memory, phonemic awareness, and the ability to segment words into syllables (for reviews, see Chandrasekaran et al., 2015; Schellenberg and Weiss, 2013; Schellenberg, 2016). Emerging evidence shows that music training can cause these effects in adults as well as children, across the lifespan (Moreno and Bidelman, 2014).

In the last decade, many researchers have argued that learning in one domain may influence the other (Hutka, Bidelman and Moreno, 2013), and that aspects of learning in music may transfer to speech if certain conditions are met during learning (e.g. Besson et al., 2011; Moreno and Bidelman, 2014; Patel, 2011). In support of this argument, a growing body of research has shown that the development of rhythmic or melodic abilities can enhance how rapidly and accurately the brain processes speech sounds at the neurological level (Chandrasekaran et al., 2015; Corrigall and Schellenberg, 2016; and Schellenberg and Weiss, 2013). This neural advantage transfers to the perception of phonemes as a result of non-specialist music training (François et al., 2013; Kraus et al., 2014; Tierney et al., 2015). The effect of music training on speech production is less well-researched (Fuji and Wan, 2014). However, a significant body of research in music therapy demonstrates that rhythmic and melodic activities can support speech production in people with developmental speech difficulties (Fuji and Wan, 2014; Thaut, 2008). For such populations, musical activities involving wind instruments or vocalising have been used to support articulation and respiratory control, and Kodaly-based singing and movement activities have been used to support vocal skills, word knowledge and production (Thaut, 2008). Although rigorous research

evidence is often lacking, music therapy has been successful in supporting and developing vocal and verbal communication in diverse populations with learning disabilities (Stephenson, 2006; Hooper et al., 2010).

This study aimed to discover the potential value of using music to support speech perception and production for people with DS. People with DS have well-documented speech and auditory learning difficulties, and a noted fondness for music. Potentially they have the most to gain from music's purported ability to support auditory skills. However, little is yet known about their musical abilities, or whether they face the same difficulties in perceiving or producing song as they do in speech; or of the degree of heterogeneity within the population. The purpose of this study was therefore to investigate the links between singing and musical ability and speech and vocal characteristics in four young adults with DS. Specifically, it aimed to learn whether strengths or weaknesses in the musical domain correspond to strengths or weaknesses in their speech and vocal domain, in order to consider the potential for singing and musical activities as speech and voice therapy. This study builds on an earlier investigation that investigated the effects of singing tuition on speech in subjects with moderate learning disabilities, including DS (Jeffery, MSc Dissertation: see Appendix 4). The previous study showed improvements in acoustic measures of voice production (Harmonic-to-Noise ratio) in 2 of 4 subjects with DS, suggesting that phonation had improved following six weeks of singing tuition. Although the present study examined phonation in singing and speech, it also examined the links between the wider speech and musical domains in order to understand the common areas of difficulty or ability. It also measured cognitive ability at outset, tracked the participants' learning of songs over a four-week period, and re-assessed simple pitching skills at the end of the teaching period. The case study approach generated a rich body of data and allowed for a deep analysis of how different factors interact in each domain, and in the four individuals.

1.3 Introduction to the thesis

A review of the literature is given in **Chapter 2**: this chapter presents key research relevant to music development and the benefits of musical training for typically-developing populations and for those with developmental difficulties, including DS. It then outlines the key characteristics of people with DS, focussing on their ability to produce motor movement and speech, and their cognitive development. The section also provides an overview of typical voice production, and an explanation of voice perturbation measures.

Chapter 3 explains the methods of assessment, analysis and the design of the teaching programme. A case study approach was taken in order to examine the abilities of four young adults with DS in terms of their auditory perception, cognition, speech production and music production.

Mixed methods were used in order to gain a richer understanding of individual abilities, and to allow generalisation (Gray, 2013). Qualitative methods drew on observations and field recordings from teaching sessions, and visual spectral analysis of speech and voice. This allowed much of the data collection to be unobtrusive, and preserved the naturalistic flow of sessions. More formal assessments of formative learning were easily incorporated into this. Initial assessments, field recordings and reassessments also generated measurable data. Performance on standardised tests and acoustic measures of voice allowed data to be compared to comparable studies. For voice production in particular, the mixed methods approach enabled a fuller understanding of the individual difficulties, and increased the validity of the findings.

Chapters 4-7 present the four case studies. For each participant, the data related to common core tasks (as outlined in Chapter 3) are presented and discussed. Although the participants are of the same age (23 years) and approximate cognitive ability (severe learning disability), their individual stories are different, as summarised below.

Andrew (**Chapter 4)** is articulate and eager to perform pop songs and songs from musicals. He has taken part in several performing arts productions, and has a relatively good memory for words, and clear speech production. However, in common with his speech, his vocal production when singing is tense in quality: it appears to be squeezed from his larynx, rendering it unmelodic and harsh-sounding. Although he was mostly keen to take part in lessons, he preferred to lead activities. His learning and his performance in tasks was often hampered by inattention and an occasional refusal to participate in assessments or in group teaching activities.

Kerry (**Chapter 5**) is a keen drummer, with a preference for rock music. In speech, her voice is severely dysphonic and dysfluent, which can lead to frustration. Despite her poor eyesight, Kerry is a good reader and is highly motivated to learn in certain situations. For example, during assessments (Chapter 3), she spontaneously mastered the Garageband app on the iPad within minutes, creating and recording an improvised guitar and drum track. Kerry is easily distracted, especially in relation to certain activities or people. However, during teaching sessions, she remained focussed and engaged. Despite her speech difficulties, she was able to reproduce songs with a secure sense of melody and timing.

Rachel (**Chapter 6**) has engaged in dance lessons since she was a child, and is a keen member of the local performing arts group. She demonstrated a good sense of rhythm when moving to music, and a sense of melody when singing. Her speech and her singing are both produced with pronounced intonational contours, but with unclear speech sounds and low intelligibility. Although

she was co-operative in assessments and learning activities, she had a tendency to rush some tasks, and demonstrated anxiety about making mistakes.

Robert (**Chapter 7**) was the least musically-able at the start of the study. He has severe hearing loss, but does not wear a hearing aid. His speech is unclear, with a reduced repertoire of speech sounds. In speech and in song his voice is quiet, relatively high-pitched and breathy in quality, and is produced with little melodic variation. He enjoys singing and dancing, and has taken part in several performing arts productions. However, during assessments, his movements to music and in rhythmic activities were imprecise and poorly coordinated. Robert was fully co-operative in all activities and made the most progress of the group between the initial and final assessments.

The case studies highlight the individual abilities and difficulties in musical and speech domains. As a group, they had difficulties in common: these cross-case findings are presented and discussed in **Chapter 8**, and compared to the wider DS population**.** Chapter 8 then considers the extent to which rhythmic and song-based activities might support speech perception and production in DS. In light of the group findings, it recommends a programme of musical activities that might be effective in supporting aspects of speech perception and production. The chapter concludes with a discussion of the strengths and limitations of the study, and with recommendations for further development.

Chapter 2: Literature review

Introduction

The first section of this review will provide an overview of music development in Typically Developing (TD) populations, and its relevance to speech perception and voice production. The chapter will begin with an outline of typical development of musical perception and production in TD children (Section 2.1). Section 2.2 will discuss research into the links between music and other domains, focussing on speech production and on the possible uses of musical activities, including singing, to support speech perception. It will also examine the evidence that singing, and activities that are common to singing and voice therapy, can be used to support speech or voice production in people with developmental or functional voice disorders.

Section 2.3 will examine the evidence that people with developmental learning disabilities, including people with Down Syndrome (DS), are able to develop musical skills, and to benefit from musical activities in terms of communication. Section 2.4 will provide an overview of DS. It will present key physiological and medical features, including a discussion of sensory impairment and motor skill development. Next, an overview of speech and voice production in DS will be given (Section 2.5), including a discussion of phonation and acoustic measures of voice in DS and TD populations. Research into cognitive development and the learning trajectory in DS will then be presented, including a discussion of auditory processing difficulties (Section 2.6). The chapter will conclude with a summary of key points and the research hypotheses and questions (Section 2.7).

2.1 Music development

This section will review the literature regarding the typical development of perception and production of musical skills from infancy to adulthood. It will summarise the development of perception and production in the first five years, and will then explain these stages of development in more detail: from pre-birth to five years; and five years onwards, including a brief overview of learning in adulthood. The focus will be mostly on evidence of productive skills, but perceptual development will also be considered, especially in the earlier learning stages. The influence of non-maturational factors will also be considered at each stage.

2.1.1 Development of musical ability between birth and five years

Research indicates that universally, musical ability during the first five years is learned in discrete stages (Hallam, 2016), and develops broadly in line with the development of motor skills and cognition. Research by Matsuyama (2005) suggests that musical ability may be linked to, and may be a reliable indicator of nonverbal Mental Age (MA) in young children. He investigated the abilities of 92 Japanese children aged between 6 months and 5;9 years to clap back simple rhythms and to repeat simple vocal melodies. He compared these results to their scores on nonverbal

development, as measured by the locally standardised 'Kyoto Scale of Psychological Development'. The scores obtained on rhythm and melody tasks both correlated with MA scores at a statistically significant level. More recent research also argues that during early development, developing musical skills in motor-rhythm and vocal domains are strongly linked to maturation and cognitive development (Berkowska and Dalla Bella, 2009; Dalla Bella, Berkowska and Sowiński, 2015). In particular, rhythmic motor production skills depend upon increasingly refined motor movements and on the development of sensory and auditory integration (Dalla Bella, Berkowska and Sowiński, 2011; Pfordresher and Mantell, 2009). However, some aspects of ability in all domains do not reach maturity until later childhood or adulthood (Deutsch, 2013; Tsang, Friendly and Trainor, 2011). The key milestones in musical perception and production between birth and five years are discussed below.

Prenatal - infancy

Infants are born with 'adult-like' hearing (Papousek, 1996) and with most of the auditory and perceptual mechanisms that are required for musical learning (Lecanuet, 1996; Parncutt, 2016). Evidence suggests that the foetus develops auditory acuity during the second trimester of pregnancy: tests have shown that the foetus responds to musical sounds from 28-33 weeks (Lecanuet, 1996; Parncutt, 2016). There is some evidence that genetics ultimately affects musical ability (for review, see Tan, McPherson, Peretz, Berkowicz and Wilson, 2014) but pre-natal exposure to environmental sounds, including music and song may affect auditory development and functioning (Lecanuet, 1996; Parncutt, 2016; Reifinger, 2006).

From birth, the majority of infants display behavioural responses to musical stimuli that indicate they are sensitive to changes in **pitch**[1] (Chang and Trehub, 1977; Trainor and Trehub, 1992) and rhythm and tempo (Chang and Trehub, 1977; Trehub and Thorpe, 1989), and that they recognise previously heard songs (Saffran, Loman and Robertson, 2000; Trainor, Wu and Tsang, 2004). From 5-6 months, infants respond to the rhythm of music (Moog, 1976 in Hargreaves, 1986; Lecanuet, 1996; Zentner and Eerola, 2010) and respond more rhythmically to music than to infant-directed speech (IDS) (Ilari, 2002; Ilari, 2014). Vocally, infants progress from producing vegetative sounds and crying at birth to 'musical babbling' at 2-4 months, which shares pitch and rhythmic qualities with song and IDS; and at 4-7 months, they produce melodic and rhythmic patterns that are influenced by the mother tongue (Hargreaves, 1986; Iverson, Patel and Ohgushi, 2004; Mang, 2001; Mang, 2006; Welch, 2016). Music learning at this age is dependent upon 'enculturation' (Trainor and Hannon, 2013; Trehub and Degé, 2016; Welch, 2016). For example, by 9 months, infants recognise tonality that is characteristic of their culture (Ilari, 2002; Schellenberg and Trehub, 1999). Between the ages of 6 and 12 months, infants lose sensitivity to non-native tones and

[1] terms in bold type are presented in the glossary (p. xx)

rhythms (Hannon and Trehub, 2005; Trainor and Trehub, 1992) which may parallel their reduced ability to discriminate non-native speech sounds (Ilari, 2002).

Several historical studies have observed that at approximately 18 months, infants begin to produce sustained vocal sounds and to create syllable-based melodic patterns (Hargreaves, 1996). At about 19 months, in a longitudinal study of nine children, Davidson, McKernon and Gardner (1981) observed that small melodic intervals appear in vocal productions, beginning with major seconds and minor thirds. At this stage, however, there is no clear distinction between the production of speech and singing (Mang, 2001; Trainor and Hannon, 2013; Welch, 2016).

Two to five years

Rhythm

The ability to synchronise motor movements to an **isochronous** beat (**entrainment**) develops slowly in infants and in line with motor maturity (Zentner and Eerola, 2010). By 2 years, just ten percent of children in a large study by Moog (1976) were able to match movements to music for short periods of time (in Hargreaves, 1986); but by 4 years, the 20 children in a study by Drake, Jones and Baruch (2000) were judged competent in tapping to a beat. However, there is considerable variation in rhythmic ability between children at this age (Drake, Jones and Baruch, 2000; Matsuyama, 2005; Reifinger, 2006). This may be explained by differences in motor development, but also by differences in early experiences. For example, in infants, perception of musical **meter** depends upon prior experience in being bounced to music (Phillips-Silver and Trainor, 2005), which is linked to the early activation of the vestibular system (Phillips-Silver, Aktipis and Bryant, 2010; Philips-Silver and Trainor, 2007; Philips-Silver and Trainor, 2008). Additionally, the presentation of the task can affect performance at this age: for example, the performance of 36 children (mean age 2.5 years) when asked to synchronise to a beat improved at a statistically significant level when they drummed with a partner, in comparison to drumming in response to visual or auditory instructions (Kirschner and Tomasello, 2010). Accuracy in beat entrainment tasks also depends upon the target tempo: in general, younger children are able to drum or clap to faster beats with greater accuracy than slow beats (Drake et al., 2000; Repp, 2005; Repp and Doggett, 2007; Thaut, 2008). This may be linked to the development of their short-term memory capacity which increases with age (Gathercole, Pickering, Ambridge and Wearing, 2004; Laws, 1998) and which can affect their ability to perceive the temporal pattern (Grahn and Schuit, 2012; Krumhansl, 2000; Repp and Doggett, 2007; Snyder, 2000, 2009).

Singing

Several studies agree that children's early singing develops in stages (Davidson et al., 1981; Dowling, 1982; Davidson and Scarlett, 1987). Davidson et al. (1981) observed that between the

ages of 23 and 28 months, songs develop rhythmically. Moog (1976, in Hargreaves, 1986) and Hargreaves (1986) both observed that the majority of the two year-olds they studied produced spontaneous songs. These consisted of repeated musical phrases with distinct, repeating **melodic contour**s and rhythms. Between repetitions, the outline of these songs remained similar, but intervals and **tonality** varied (Hargreaves, 1986). Dowling (1982) reported that at the age of approximately 30 months, infants begin to imitate songs they hear, repeating words and phrases, but without reproducing the melody or rhythm of the original. In a cross-sectional study, Davidson and Scarlett (1987) observed the singing abilities of 70 children aged between 18 months and 5 years and observed that their participants began to acquire lyrics, then tempo and dynamics before mastering the rhythmic and melodic detail of songs. By 3 years, the majority of their children could clap the pulse (or 'beat') of rhythm and sing the melodic shape, and by 4 they were beginning to master intervals, but without consistency between productions. By 5 years, studies agree that most children are able to reproduce familiar songs with all rhythmic and melodic elements in place (Hargreaves, 1986; Sloboda, 1985). However, some children may sing in a style that more closely resembles speech, unless teaching is able to correct performance errors (Welch, 1994).

Despite a universal learning trajectory in music development (Hallam, 2016), children demonstrate a wide range of singing abilities when starting school (Welch, 2016). Several studies have recommended that singing development is characterised by phases, rather than ages (Davidson 1981, in Hargreaves, 1986; Rutkowski, 1997; Welch, 1998; Welch, Sargeant and White, 1998). Davidson et al. (1987) identified four main phases in children of 4 and 5 years. Firstly, the global outline is learned, including words, phrases, **pulse** and pace. In the second stage, children can clap the word rhythms in time to the pulse and sing key phrases with approximate melodic contour. In the third stage, matching contour improves and intervals begin to stabilise. In the fourth stage, all aspects begin to stabilise, children can find the pulse from listening to the word rhythm, and they begin to perform with a tonal centre (i.e. keep within key). Alternative models have been proposed by Rutkowski (1997) and Welch (1998). Rutkowski (1997) identified nine phases in singing development and Welch identified four main phases that characterise pitching ability in young singers (1998). These models chart the development from chanting to singing within speaking range, to refinement of tuning and increasing accuracy of melodic intervals, and expansion of vocal range. These models propose that development in singing accuracy depends more upon experience and motor maturity than upon age. Experience is necessary to support cognitive understanding: for example, Davidson (1985) suggests that singing development depends on acquiring a schematic processing of melodic contour, in which individual intervals are gradually mastered and tonality is gradually achieved. This is comparable to models of **pitch** processing that were proposed by Dowling (1978, 1982) and Deutsch (Deutsch and Feroe, 1981). These models propose that pitch is hierarchical (Deutsch and Feroe, 1981) and that the concept of tonality (a

higher-level hierarchy) supports learning of melodic intervals (a lower-level hierarchy). Therefore, melodic intervals are more easily learned when a representation of tonality or key has been developed (Dowling, 1982). The ability to master precise intervals is in turn linked to the development of memory (Deutsch, 2013), which increases with age (for a review, see Gathercole et al., 2004).

2.1.2 Development of musical ability during school years

Early development in music is strongly linked to the maturation of motor and cognitive systems, as well as to exposure to and the opportunity to practice (Bispham, 2006; Corrigall and Schellenberg, 2016; Hargreaves, 1986; Reifinger, 2006; Lamont, 2009). The diversity of these factors between children increasingly complicates the study of development beyond 5 years (Deutsch, 2013; Lamont, 2009; Demorest and Pfordresher, 2015). This is further complicated by pedagogical issues and by the individual's own perception of their abilities (Welch, 2016). Furthermore, opportunities to learn and develop musical skills in music at Primary and Secondary levels have historically been affected by the constraints of the National Curriculum, and by issues relating to school ethos and pedagogy (Swanwick, 2004). These types of factors can impact further upon development of musical ability within individuals (Welch, 2009). Despite these confounding factors, research indicates that rhythmic music-motor skills and singing skills continue to develop broadly in line with motor maturity and cognitive development.

Rhythmic development

A key study by Drake et al. (2000) examined rhythmic ability and focal attention in children between the ages of 4 and 10 years. A cohort of 180 children and adults was split into groups of 20 according to four age levels and prior musical training. Five groups of nonmusicians (aged 4, 6, 8 and 10 years and adult) and four groups of musicians (in groups of ages 6, 8, 10 years and adult) took part in a battery of rhythmic tasks. They found that across all tasks, accuracy and variability differed as a function of age. In forced tapping tasks, in which participants were asked to tap at their fastest and slowest rate, the spontaneous tapping rate declined between the ages of 6-10 years for non-musicians and musicians alike. They found that older participants had improved motor control but that accuracy did not necessarily improve with musical training. However, the musicians were better able to perceive a wider range of hierarchical beat levels. This supports evidence that shows that accuracy in tapping to slower tempi is related to the development of memory for longer temporal sequences (Drake and Bertrand, 2012; Grahn and Schuit, 2012; Krumhansl, 2000). In the same study, participants were asked to tap on a drum in time to an isochronous beat at their preferred tempo, from a selection of five tempi, and to a rhythmic sequence. The ability to synchronise in both tasks also improved with age and with musical training. Their study showed that the majority of children were unable to synchronise to the rhythm

before 10 years of age. However, all children entrained more accurately to the beat of Ravel's *Bolero* than to isochronous sequences. The authors suggest that the additional cues may support focal attunement to **beat hierarchies** in younger children.

In earlier studies, the authors reported that the ability to perceive meter develops with age and enhances rhythmic perception (Drake and Gérard, 1989; Gérard and Drake, 1990). In a study of 120 pupils aged between 5 and 8 years, the majority of 8 year-olds (83%) were able to perceive metrical accents and perception was most accurate at faster tempi. However, the ability to reproduce metrical accents was observed in just 50% of the older children. The authors concluded that motor control was not refined until 7 years, but also that performance was reduced when attention was divided between rhythmic form and dynamics. In their earlier study, Drake and Gérard (1989) found that rhythm imitation also improved with age, but they noted no difference in the ability to produce metrical rhythms between 7 year-olds and adults.

Several studies illustrate that the mode of instruction and production affects performance whilst motor and attentional skills are developing. For example, at 5 years, rhythms based on word syllables are more easily produced than tapped rhythms (Sloboda, 1985); rhythms are more easily produced if they are accompanied by sung words (Gérard and Auxiette,1992); and rhythms are reproduced more easily when they are tapped, rather than presented as part of a melody (Hargreaves, 1986). In their study of children aged between 6-8 years, Schleuter and Schleuter (1985) reported that younger children are better able to perform rhythms when chanting on a neutral syllable ('loo') than by clapping or stepping.

Singing development

Singing ability develops significantly during the ages of 7-11 years: just 30% of 7 year-olds are judged as 'inaccurate' singers, and this reduces to 4% by the age of 11 years (Welch, 2009). A large study by Welch et al. (2008) examined the progress in singing abilities of over 3000 pupils aged between 7-10 years, in 77 UK schools (Welch, Himonides, Saunders, Papageorgi and Sarazin, 2008). Participants were assessed against the models developed by Welch (1998) and Rutkowski (1997) (see above). The 10 year-olds demonstrated more advanced skills than the 7 year-olds, at statistically significant levels, confirming that singing ability continues to develop with age. However, this study also revealed the wealth of social, cultural and experiential factors that affect singing development. Upon completion of the project by Welch et al. (2008), which involved 8000 pupils and 150 Primary schools, data also revealed that accuracy in singing is related to teacher expectations, to the learning environment and to the importance ascribed by the school to singing (Welch, 2009). Their data complements that of other research that has demonstrated that

singing ability during later childhood and adolescence is affected by sex (Welch et al., 2008), and by the individual's perception of the task and how they identify as a singer (Welch, 2016).

2.1.3 Learning and development beyond middle school years

The evidence shows that most children follow a similar developmental trajectory in music and singing until about five to seven years. As discussed, many factors can account for variation between individuals as they develop and during schooling, and the extent to which this ability develops beyond maturation depends upon continued experience and opportunities (Demorest and Pfrodresher, 2015; Gooding and Standley, 2011). However, the ability to further develop specific musical skills may not be age-dependent. Adults retain enough plasticity to develop musical skills (Trainor, 2005) and singing skills (Demorest and Pfordresher, 2015; Mang, 2006), but success depends on the number of hours of practice (Gooding and Standley, 2011; Lamont, 2009; Trainor, 2005; White, Hutka, Williams and Moreno, 2013) and, in singing, on the maintenance of skills (Demorest and Pfordresher, 2015). As with children, the modality may continue to affect learning for visual learners, who learn more effectively when auditory information is supported with visual representations (Korenman and Peynircioglou, 2007; Mishra, 2007).

2.1.4 Section summary and relevance to the study

The literature indicates the primacy of experience and opportunity in the development of musical ability, coupled with teaching approaches that are developmentally appropriate. Therefore, it should be possible to develop the singing skills of any child or adult regardless of their cognitive impairment or age, and regardless of initial singing or musical limitations. Recent studies by Ockelford, Welch, Jewell-Gore, Cheng, Vogiatzoglou and Himonides (2011) have confirmed that progress can be made in people with severe and profound cognitive impairments if teaching activities are delivered at the developmental, cognitive and age-appropriate level. The development of these skills could have broader implications for a wide range of cognitive skills, and for speech and language in particular. This will be discussed in the next section.

2.2 Music development and its relevance for speech and voice

This section summarises research evidence that participation in music activities may affect aspects of development that are relevant to speech perception and production, and that music and singing activities may affect speech and voice production.

2.2.1 Musical ability and the perception and production of speech

Recent research comparing the brains of musicians to those of non-musicians suggests that extensive music practice leads to neurological, anatomical and functional differences in musicians, specifically in motor and auditory brain structures (for reviews, see Chandrasekaran et al., 2015;

Corrigall and Schellenberg, 2016; and Schellenberg and Weiss, 2013). Studies involving people with and without musical training show that musical strengths and weaknesses appear to have linguistic correlates (e.g. Anvari, Trainor, Woodside and Levy, 2002; Corriveau and Goswami, 2009; Cumming, Wilson, Leong, Colling and Goswami, 2015; Overy, 2000; Peter and Stoel-Gammon, 2008; Forgeard, Winner, Norton, Schlaug and Fitch, 2008). Much of the research involves correlational studies, in which the cause of differences cannot be determined (Chandrasekaran et al., 2015; Habib and Besson, 2009; Kraus and Chandrasekaran, 2010; Schellenberg, 2016; Wan and Schlaug, 2013). However, longitudinal studies and techniques that allow non-invasive studies of brain functioning are beginning to unravel the relationships between musical learning and other types of learning, with implications for intervention (for a review, see Besson, Chobert and Marie, 2011). In comparison to non-musicians, the brains of musicians typically display faster neural responses to the onset of speech sounds, and demonstrate enhanced processing and representation of sounds in the brainstem (Kraus and Chandrasekaran, 2010; Wan and Schellenberg, 2013). Functionally, musicians show superior abilities in a range of auditory skills, including: enhanced processing of pitch patterns (Besson, Schön, Moreno, Santos and Magné, 2007; Magné, Schön and Besson, 2006); enhanced processing of lexical stress (Kolinsky, Cuvelier, Goetry, Peretz and Morais, 2009); superior perception of the metrical structure of words (Marie, Magné and Besson, 2011); enhanced perception of speech in noise (Parbery-Clark, Skoe, Lam and Kraus, 2009); and faster processing of variation in rhythm pitch and harmony (Besson et al., 2011; Kraus and Chandrasekaran, 2010).

To date, the evidence that music training causes changes in speech-related skills is limited (Hallam, 2010; Chandrasekaran et al., 2015; Kraus and Chandrasekaran, 2010; Wan and Schellenberg, 2013). Furthermore, much of the research that indicates superior speech and language-related skills in musicians focusses on musicians who have received high levels of instrumental training and private instruction (Kraus and Chandrasekaran, 2010). However, a number of recent experimental studies demonstrate that changes are evident in the speech domain for groups of children who attend music groups, in comparison to matched groups who receive alternative sessions. For example, François, Chobert, Besson and Schön (2013) pseudo-randomly assigned twenty-four 8 year-old children to either music or painting sessions over a period of two years. Prior to taking part in sessions, the groups were matched on a number of cognitive abilities, and on age, sex and SES. After two years, the group who had received musical training, based on a combination of Kodaly and Orff approaches, performed better in identifying pseudowords from a speech stream of nonsense syllables. They conclude that music training develops skills in rhythm perception that are also required for speech perception.

The majority of correlational studies, and emerging longitudinal studies have introduced bespoke music programmes. Two recent studies demonstrate that pre-existing musical provision may be sufficient to lead to neurological and behavioural changes in speech processing (Kraus, Slater, Thompson, Hornickel, Strait, Nicol and White-Schwoch, 2014; Tierney, Krizman and Kraus, 2015). Over a three-year period, Kraus et al. (2014) studied the impact on 44 socially-disadvantaged children (aged between 6.7 and 9.3 years) of between one and two years' participation in an existing community music project. They measured annually the auditory brainstem responses to /ba/ and /ga/ and found that neural responses improved in line with the duration of instrumental training. Kraus et al. (2014) did not measure behavioural changes, but a longitudinal study by Tierney et al. (2015) demonstrated that participation in a music ensemble led to changes in neurological responses and statistically significant changes in phonological awareness. They collected neurological and behavioural data over a four-year period of 40 adolescents (mean age 14;7 years) who were enrolled in either a musical-ensemble group or in a Junior Reserve Officers Training Corps. Their study measured the brainstem response to a /da/ syllable and test scores on phonological processing. Although both groups showed improvements, the musical group developed superior neural encoding and also demonstrated superior phonological processing skills.

These experimental studies provide compelling evidence that music training programmes cause changes in speech processing that could not be accounted for by pre-existing levels of auditory perception or cognitive ability. The authors of all three studies (François et al., 2013; Kraus et al., 2014; Tierney et al. 2014) suggest that such effects might have consequences for speech outcomes. Studies into the musical skills of children with language deficits show a similar connection between musical deficits and speech and language disorders. Some correlational studies have investigated which specific musical elements are most closely correlated with linguistic skills, with conflicting results. Rhythmic disturbances have been found in some studies of children with Specific Language Impairment (SLI) (Corriveau and Goswami, 2009; Cumming et al., 2015), and in children with Developmental Apraxia of Speech (DAS) (Peter and Stoel-Gammon, 2008). The apparently strong links between developmental disorders of speech and reduced musical ability suggest a degree of shared neural and/or cognitive processing. In this case, it would seem that any person with a deficit in either the musical or linguistic domain would also be impaired in the other domain. This topic is widely disputed (Peretz and Coltheart, 2003; Peretz and Zatorre, 2003; Patel, 2008; Zatorre and Baum, 2012). Studies into amusic subjects, who demonstrate deficits in musical pitch processing, indicate that their speech is unaffected by their poor perception of 'fine-grained' pitch (Peretz and Coltheart, 2003; Peretz, Gagnon, Hébert and Macoir, 2004). However, evidence of overlap may depend upon the testing procedure as well as upon individual cases. For example, Lui, Patel, Fourcin and Stewart (2010) found that compared to

controls, the majority of their participants with amusia were impaired in their ability to distinguish between statements and questions, as signalled by small pitch excursions (c 2.4 semitones) at the end of sentences. Accordingly, Liu et al. (2011) argue that music and speech share resources for processing.

To date, most evidence suggests that music and speech are separate domains in the adult brain, with the result that deficits may be observed in one domain but not in the other. However, the difference between song and speech is often blurred (Mang, 2001; Tierney, Dick, Deutsch and Sereno, 2012; Welch, 1994), and may be learned. It is not known at which point or to what extent individuals begin to categorise music and language as being separate, nor what experiential or cognitive factors lead to this division. According to Rinta and Welch (2008), children of six years and below may not differentiate between 'singing' and 'speech' unless the culture in which they live makes a distinction. Singing and speech may remain essentially 'the same' process in some contexts (Deutsch, 2013; Tierney et al., 2013), cultures (Brandt, Gebrian and Slevc, 2012; Rinta and Welch, 2008) and for some individuals (Mang, 2001; Welch, 2016). The possibility of shared neurological processing, and the undisputed social and emotional benefits that can be gained from singing and music-making (Clift, 2012; Deutsch, 2013; Welch, Himonides, Saunders, Papageorgi and Sarazin, 2014) may make singing-related music activities a highly effective tool for supporting speech processing or production.

2.2.2 Therapeutic effects of singing on speech and voice production

In singing, the prosodic and vocal qualities come to the fore, exaggerated or manipulated for rhythmic and emotive effect. Because singing is carried mostly by sustained vowels (Rinta and Welch, 2008), it requires a greater percentage of sustained phonation than speech. It also utilises higher vocal intensity, greater control of pitch, and different use of breath rhythms (Wan, Ruber, Höhmann and Schlaug, 2010). Singing uses the same physiological processes as speech, and shares some neural mechanisms (Jäncke, 2012; Özdemir, Norton and Schlaug, 2006), particularly areas associated with motor perception and production (Callan, Tsytsarev, hanakawa, Callan, Katsuhara, Fukuyama and Turner, 2006). The degree of physiological overlap and the increased demands that are placed on the cognitive system when learning to sing are such that several authors argue that singing may be a useful tool for supporting speech perception or production (Christiner and Reiterer, 2013; Patel, 2008, 2014; Rinta and Welch, 2008). Rinta and Welch (2008) recommend singing as a potential tool for voice disorders, specifically for prepubertal children, arguing that singing may potentially benefit the physiological functioning and psychological factors that impact on effective vocalisation. In addition, they suggest that the rhythmic elements of singing may lead to improvements in the rhythmic production of speech.

Singing has long been recommended as support for speech and language therapy (Zoller, 1991). Historical studies suggest that singing and musical activities encourage better vocal functioning and articulation in children (Murphy and Simons, 1959) and may encourage expressive speech (Seybold, 1971), but the evidence is inconclusive due to lack of controls. Evidence from music therapy does indicate that musical activities might be a useful tool for supporting speech in individuals with learning disabilities (Jellison, 2016; Hooper, Wigram, Carson, and Lindsay, 2010). Communication is a key focus for music therapy (Stephenson, 2006) and individual studies indicate that music- and song-based activities may encourage and support vocalisation. However, historically, there is a lack of developmentally-appropriate assessment and a lack of rigorous experimental evidence (Hooper et al., 2010; Stephenson, 2006). Stephenson (2006) suggests that where changes have been noted, this may be a result of the musical context being a positive experience and inherently motivating. However, she concludes that musical activities may be suited to developing early communication skills within students with developmental disabilities if a programme is designed in conjunction with teachers and speech therapists.

Melodic Intonation Therapy (MIT: Sparks and Holland, 1979) has been used in a number of recent studies to develop spoken communication in people with acquired or developmental speech disorders. The approach uses simplified, musically-structured intonational speech contours, and tapped patterns derived from word rhythms in order to gradually develop expressive communication in patients (Sparks and Holland, 1979). Carroll (1996) used MIT with four young children with DS and reported increases in the amount of vocalisation compared to a control group receiving a comparable, non-melodic intervention. Schlaug, Marchina and Norton (2008) used MIT with two non-fluent patients with Broca's aphasia who had previously undergone a year of speech and language therapy. Each received a programme of intensive therapy: one received MIT and the other received a comparable non-song based programme. Both showed improvements in their speech but the patient receiving MIT showed increased improvement and greater activation of the brain’s right hemisphere. The authors suggest that the melodic activities engage sustained phonation and greater activity in the right hemisphere in order to process melodic and metrical detail. They also suggest that the tapping activity, in which word rhythms are tapped onto the patient’s left hand, triggers sensorimotor coordination in the brain that benefits oromotor and facial coordination. A more recent study by Stahl, Henseler, Turner Geyer and Kotz (2013) obtained similar results without the melodic component and suggest that the rhythmic tapping aspect alone may produce the same changes in patients with comparable difficulties.

Wan et al. (2010) reviewed the literature detailing the effects of singing on the quality of speech for people with acquired and developmental speech-motor difficulties. Their research suggests that singing leads to physiological improvements in vocal function and pulmonary function for both

populations. The authors suggest that the successful application of singing as voice and speech therapy may be due to a number of factors relating to improvements in brain functioning, motor output and perception. In particular, they highlight the potential for singing to engage speech-motor functions; engage different brain areas; encourage slower processing of sound; heighten awareness of speech sounds; facilitate vocal output through the use of beat to cue sounds; and to develop neural pathways linking sounds and motor patterns. The apparent ability of music and singing to tap into the same motor-networks as speech is highlighted by Thaut (2008), who cites several therapy programmes that employ singing and specific musical activities (e.g the use or whistles or kazoos) to improve speech production.

A study involving music with children with developmental speech delay indicates that singing may improve perception of prosody (Groß, Linden and Ostermann, 2010). The study considered the effects of music on the understanding of language and interest in communication in 18 young children who took part in singing and musical activities with two music therapists. Scores on phonological memory for words and understanding sentences increased during the treatment period, but mostly during periods of music sessions, with statistically significant improvements. The authors interpret the results as an effect of increased prosodic awareness. They also note that children with speech delays are often also delayed in their ability to 'grasp' the prosodic qualities of speech and argue that this deficit places additional demands on working memory. They conclude that an improved awareness of prosody may therefore lead to improvements in speech development. However, without a control group, the effects of their intervention may be attributable to other non-musical factors, such as the children's emotional response to an increase in attention to their needs or a behavioural response to being observed.

There is a small body of research to suggest that singing-related activities may help improve phonation in children and adults with the functional voice disorder, Muscle Tension Dysphonia (MTD). MTD is characterised by a 'harsh' or 'strained' quality of voice that results from abnormal tension within the laryngeal muscles or in the external muscles (Aronson and Bless, 2009; Fawcus, 2009; Laver, 1980; Roy, 2003; Sataloff, 2014). In a study of eight children with MTD, Lee and Son (2005) reported that over a period of 1-2.5 months, bi-weekly exercises in producing vowels with a soft onset, humming and instruction in laryngeal relaxation resulted in a reduction in perceived hoarseness and strain. Acoustic and stroboscopic analysis confirmed reduced voice perturbation measures (jitter, shimmer, NHR) and reduced supraglottic tension, post-therapy. In a similar study, Ogawa, Hosokawa, Yoshida, Yoshii, Shiromoto and Inohara (2014) observed the immediate effects of humming on the laryngeal tension of 23 adults with MTD and 15 controls with normal voice production. The resulting laryngoscope videos were analysed in order to measure the degree of supraglottic tension in each production. Voice quality was judged against a standard assessment

(GRBAS) by four speech and language therapists. The results indicated statistically significant differences between the groups, and humming resulted in reduced supragalottic compression and in lower judgements of 'roughness' for 18 of the participants with MTD. Ogawa et al. conclude that for adults with MTD, humming corrects anterio-posterior compression of the aryepilglottis and reduces lateral compression of the vocal folds.

2.2.3 Section summary and relevance to the study

Evidence from music therapy literature suggests that singing and musical activities may be successful in developing awareness of speech sounds and words in children with developmental difficulties with potential benefits for production. Research also indicates that the effects of habitual hyperfunctional voice disorder may be reduced by the types of vocal exercises that are typically found in singing pedagogy (e.g. humming, soft onset: Phillips, 1996). The literature indicates that music and singing has the potential to support aspects of speech perception, speech production and voice production in people with learning disabilities and developmental speech and voice disorders. The current study aims to examine the potential use of music and singing to support speech perception and production and voice production in adults with DS. Therefore, the next sections will explore the characteristics of DS in order to determine the main barriers they face in speech domains, in learning and in developing musical skills.

2.3 Down Syndrome

In the UK, approximately 1 in 1000 infants are born with DS (Morris and Springett, 2014; Wu and Morris, 2013). DS is a medical condition caused when cells contain an extra copy of the chromosome 21, leading to an over expression of certain genes associated with the chromosome (Capone, 2001). The syndrome results in learning difficulties, and a high incidence of birth defects and medical conditions (Alexander, Petri, Ding, Wandel, Khwaja and Foskett, 2015; Shapiro, 2003; Wu and Morris, 2013). The syndrome is characterised by a range of physiological features and learning differences that set apart Down Syndrome from other types of genetic or developmental difficulty (Silverman, 2007). However, is important to note that there is considerable within-group heterogeneity. The following section will provide an overview of the physiological differences associated with the syndrome that are most relevant to the development of speech, voice and motor-movement. Sections 2.4 and 2.5 will review in more detail the characteristics of speech, voice and cognition that are typical of the syndrome.

2.3.1 Physiological characteristics

Infants born with DS have distinctive facial and physiological characteristics. Anatomically, many DS children and adults are small in stature, with most adolescents being 2.5 standard deviations shorter than their peers (Davis, 2008). Individuals typically have small arms, legs and fingers

relative to TD peers (Sacks and Buckley, 2003) and broad hands (Davis, 2008). Physiological and muscular differences within the face are particularly relevant for speech as they can affect hearing, the control of muscles involved in speech production and the resonance of speech sounds. Relative to TD individuals, those with DS typically have a high, narrow palatal arch, a small lower jaw, and a small oral cavity (Bhagyalakshmi, Renukarya and Rajamgam, 2007) whilst the tongue and soft palate are of average size, making the tongue seem relatively large for the oral cavity (Xue, Kaine and Ng, 2010). Many people with DS have teeth occlusions and poor muscular control of the tongue, lips and lower jaw (Bhagyalakshmi et al., 2007; Cleland, Wood, Hardcastle, Wishart and Timmins, 2010; Uong, McDonough, Tayag-Kier, Zhao, Haselgrove, Mahboubi et al., 2001). Evidence suggests that DS children and adults typically have anatomical differences in the ears, nose and throat, which include: a narrower external ear canal (Sacks and Buckley, 2003); small nasal sinuses (Cawson and Odell, 2008); and a reduced vocal tract volume (Xue et al., 2010). There may be malformations within the respiratory tract (Cawson and Odell, 2008) and reduced pulmonary function (da Silva, de França Barros, de Azevedo, de Godoy, Arena, and Cipriano, 2010; Laibsirinon, Jarusurin, Kokoi, Manakiatichai, 2012; Salguerinho, Venâcio, Martín-Nogeuras and Ribero (2012). Facial-cranial physiological features may affect articulation of sounds, vocal cord control, and resonance (Bhagyalakshmi et al., 2007); and speech and language development may be further affected by hearing difficulties (Bennetts and Flynn, 2002; Delage and Tuller, 2007; Kent and Vorperian, 2013; Pettinato, 2009). The consequences of these characteristics for speech will be discussed further in Section 2.4.

2.3.2 Peripheral sensory impairments

Approximately 70-80% of children with DS have intermittent hearing loss primarily as a result of infections in the ear that lead to recurring bouts of otitis media (Alexander et al., 2015; Laws and Hall, 2014; Maris, Wojciechowski, de Heyning, and Boudewyns, 2014; Pettinato, 2009). In a cross-sectional study of 107 children with DS, Maris et al. (2014) reported that peaks of prevalence in hearing loss occurred in the first year and between 6-7 years, with median losses of 28.3 dB (unilateral loss) and 36.7 dB (bilateral loss). A reduction of just 10 dB affects the ability of children with DS to perceive speech in noise (Bennets and Flynn, 2002). The long-term consequence of this temporary hearing loss has been debated in relation to speech and language development in children with average and above-average cognitive abilities (Klein and Rapin,1993). However, for children with DS, even a small loss may have a significant effect for long-term development of speech and language (Bennetts and Flynn, 2002; Rondal and Buckley, 2003; Laws and Hall, 2014). Comparisons of children with DS to children with cochlear implants, who receive a reduced auditory signal, show that both populations exhibit comparable reductions in hearing acuity, similar speech difficulties, and difficulties in short-term memory for phonological sounds (Pettinato, 2009). Pettinato suggests that a reduced quality in early speech processing may be correlated with poor

memory for speech sounds in the DS population. Hearing loss will also alter the auditory feedback processes requires to produce voice and can affect laryngeal control, control of respiration, phonation and resonance (Coelho, Medved and Brasolotto, 2015). Given the high incidence of hearing loss in infants with DS between 1 and 8 years (Maris et al., 2014), a prolonged period of hearing loss could potentially affect voice function in children with DS.

Impairments in sight are more common in DS children than in TD children (Alexander et al., 2015). Alexander et al. (2015), studied over 6000 UK medical records from 2004-2013 and reported that 28.1% of DS infants are treated for eye disorders. Over 70% of people with DS have persistent difficulties in focussing on near objects, and visual acuity lags behind TD infants and is not always corrected by wearing glasses (Woodhouse, 2005). Differences in eye-gaze and in establishing joint attention with care-givers have been reported and may contribute to the development of different styles of interaction with caregivers, with long-term consequences for language development (e.g. Chapman and Hesketh, 2001). Difficulties in visuo-motor perception have also been reported (Virji-Babul, Kerns, Zhou, Kapur and Shiffrar, 2006). Virji-Babul et al. (2006) investigated visual-motor perception skills in DS children and concluded that they are impaired in their ability to perceive and interpret the complex movements involved in gait and in postural cues signifying emotions, with consequences for social development.

Many children with DS have poor proprioception and rely more on visual cues to maintain physical postures, such as when standing and sitting (Sacks and Buckley, 2003). Difficulties in proprioception are also implicated in learning speech movements. Hamilton (1993) used palatography and visual feedback as speech intervention and reported improved motor-control for some speech sounds, implying that proprioceptive deficits may be involved in co-ordinating movements for speech. Proprioceptive difficulties will make it difficult to lay down stable motor-memory for some sounds. These factors may account in part for inconsistencies in production, which are typical within the population (Cleland et al., 2010; Stoel-Gammon, 2001).

2.3.3 Motor development

At birth, infants with DS are described as having reduced muscle tone and ligaments that are more elastic than in other populations, allowing a greater range of movements and giving the appearance of ‘floppy’ limbs (Sacks and Buckley, 2003). Research within the population has confirmed that excessively relaxed muscles (‘lax’) lead to differences in joint movements and postural control, with consequences for control of gait, arm movements, ankle and head movements (Latash, Wood and Ulrich, 2008). People with DS are also delayed in the acquisition of early motor skills, such as crawling and walking (Latash, 2007; Vicari, 2006; Palisano, Walter, Russell, Rosenbaum, Gémus, Galuppi and Cunningham, 2001).

Differences in the quality of some movements remain evident throughout the lifespan of people with DS. Whole-body movements, limb movements and fine motor control all appear ‘clumsy’ and slow in comparison to peers (Latash, 2007). When walking, children with DS tend to use a wide stance (Palisano et al., 2001; Rigoldi, Galli, and Albertini, 2011a) and when standing, adolescents and adults with DS maintain a less stable posture in comparison to age-matched controls (Rigoldi, Galli, Mainardi, Crivellini, and Albertini, 2011b; Shumway-Cook and Woollacott, 1985). Motor tapping tasks have revealed differences in the trajectory of arm movements in adults with DS, in comparison to age-matched controls (Ringenbach, Allen, Chung and Jung, 2006; Vimercati, Galli, Rigoldi, Ancillao, and Albertini, 2013). Vimercati et al. (2013) required participants to tap between five dots that were arranged on a semicircular pattern on a flat surface. They reported that the arm movements of participants with DS were slower and less linear, and that they used trunk movements to move between dots, rather than elbow and arm movements. Their findings support those of Ringenbach et al. (2006) who found that adults with DS also used less linear arm movements in bimanual drumming tasks.

Motor functioning has been linked to cognitive ability in DS (Chen et al. 2015; Malak, Kotwicka, Krawczyk-Wasielewska, Mojs and Samborski, 2013; Sourtiji, 2010). Chen et al. (2015) found that fine motor control, as measured by a manual dexterity task was positively correlated with the time taken to complete a problem-solving task and with verbal memory, with statistically significant results. They suggest that remediation in motor skills may affect cognitive development. However, several studies reveal that the mode of instruction can affect performance in motor tasks (Ringenbach et al. 2006; Bunn et al., 2007). Bunn et al. (2007) reported that verbal-only instructions resulted in poor imitation of pantomime gestures (Bunn et al., 2007) but that the accuracy of motor movement was enhanced with the presence of a symbolic tool (Bunn et al., 2007). Similarly, visuo-spatial instructions resulted in improved accuracy in bimanual tapping tasks, in comparison to verbal or visual-only instructions (Ringenbach et al., 2006).

Research indicates that motor-movements can be developed in accuracy and efficiency, even if movements remain slower than in TD populations (Almeida, Corcos and Hasan, 2000; Capio and Rotor, 2010; Latash et al., 2008; Rigoldi et al., 2011a; Sacks and Buckley, 2003). More time is required for individuals to learn specific motor movements, especially as they increase in complexity (Palisano et al., 2001). However, for 16 young adults with DS, measurable improvement was recorded in the speed and accuracy of simple elbow movements after five days of practice, and of finger movements after three days of practice (Latash, Kang and Patterson, 2002). Latash (2007) argues that even short periods of focussed training is sufficient to improve motor coordination at the neural level.

2.3.4 Section summary and relevance to the study

People with DS have physiological and neurological characteristics that affect their perception of sensory information, and their development, acquisition and production of a range of motor skills. These affect the quality, co-ordination and speed of limb-motor movements. The integration of sensory information and the co-ordination of limb-motor movements and oral movements are required for proficiency in music and in singing. It is therefore to be expected that individuals with DS will find it more challenging than TD peers to master aspects of musical production. However, evidence suggests that for people with DS, practice in motor skills enhances performance for manual-motor tasks. Therefore, it is also anticipated that people with DS will be able to develop rhythmic-motor movements. This is essential, if they are to benefit from musical activities.

2.4 Speech and voice in DS

This section will summarise typical difficulties associated with the syndrome in terms of segmental and suprasegmental perception and production. The first section will give a brief overview of the characteristics of segmental speech production in the DS population, primarily in the context of articulation and phonological production, and of the role of prosody in speech intelligibility. The next section focuses on voice: it begins with an overview of what is required in order to produce voice in healthy populations. This is followed with an overview of the perceived characteristics of the DS voice. It will then examine the evidence that people with DS have difficulties in producing phonation. The concluding section will examine the evidence that people with DS have difficulties in perceiving and producing speech prosody.

2.4.1 Speech characteristics and intelligibility in DS

Onset of speech is delayed, but the acquisition of sounds in DS children follows a similar pattern to that of TD children (Kumin, Councill and Goodman, 1994; Stoel-Gammon, 2001; van Borsel, 1996). Cleland et al. (2010) report a high incidence of developmental errors in speech and some non-developmental errors, such as nasal emission. Differences in phonological development are likely to stem from the complex interactions of perceptual, physical and cognitive factors that accompany the syndrome. For example, fluctuating levels of hearing loss are likely to interfere with acquisition of speech sounds (see Section 2.3.2); and a reduced auditory memory (see Section 2.5.2) may affect accurate storage and retrieval of sounds. Anatomical difficulties affect sound production and inconsistencies in production suggest motor-programming difficulties (Stoel-Gammon, 2001; Rondal, 2009). People with DS typically have facial and oral structures that may limit oro-motor movements and affect the quality of speech (see Section 2.3.1). It is therefore not clear to what extent delayed and deviant oral motor control affects speech production in this population (Alcock, 2006; Kumin, 2006). Barnes, Roberts, Mirrett, Sideris and Misenheimer (2006) found that 75% of 34 DS boys tested had abnormalities in the relationship between lips and other

oral structures, teeth gaps, teeth occlusion, tongue carriage, and palatal vault height, and they had difficulties in controlling lip, tongue and jaw movements. Their subjects performed more poorly during speech tasks than oro-motor tasks, and improvements in oro-motor performance failed to transfer to speech. The authors argue that motor-programming difficulties lead to speech difficulties, rather than anatomical difficulties or reduced motor-control. These results support Kumin's findings (2006) that motor-planning difficulties do not always correspond to oral-motor difficulties. In a parental survey of over 1600 families with DS children, 60% had received diagnoses of oral motor control difficulties; 15% had been diagnosed with dyspraxia, and just 2% had not been diagnosed with motor control difficulties. However, she suggests that symptoms show that dyspraxia is under-diagnosed in the population.

For people with DS, although the production of segmental speech sounds affects the clarity of words, intelligibility may also be affected by atypical prosody (Bowen, 2015; Ramig, 1992; Greenberg and Arai, 2001). Prosody is the melodic and rhythmic pattern that arises from changes in pitch, syllable duration, and variations in tone and intensity that occur at a usually subconscious level to help a speaker convey linguistic and emotional meaning (Peppé, 2009). Nooteboom (1997) argues that prosodic cues 'maintain' intelligibility when other, segmental, cues are poor. The production of prosody requires an ability to control the duration, pitch and relative intensity of syllables within words and to plan rhythmic and intonational contours at the whole phrase level. Difficulties in prosodic production may therefore arise from the rhythmic disturbances to speech production, such as dysfluency or altered duration of speech sounds, and from difficulties in controlling the production of voice. The next section discusses the requirements for healthy voice production and the characteristics of voice production in DS. Prosodic perception and production in DS will be discussed in Section 2.4.3.

2.4.2 Voice production in healthy populations and in DS

Voice production in healthy voices

In a healthy vocal system, the application of breath to the vocal muscles causes them to open (abduct) and close (adduct) in a regular cycle and with minimal air wastage (Gick, Wilson and Derrick, 2013; Mathieson, 2001). The regularity of the movement generates glottal pulses, known as **phonation**, or voicing. One cycle of abduction and adduction is the 'period' of the vibration: the number of periods produced per second is the **fundamental frequency** (F0: measured in Hertz), which is perceived as 'pitch'. Physiological properties of the nervous system mean that in healthy vocal systems there are small variations in the duration between successive cycles ('**jitter**') (Baken and Orlikoff, 2000; Mathieson, 2001). Similarly, there are minor fluctuations in the intensity of the breath stream as it is applied to the vocal muscles and as it passes through them ('**shimmer**') (Baken and Orlikoff, 2000; Mathieson, 2001). Jitter and shimmer values vary between individuals

but higher values along the continuum are associated with pathology and with changes in perceived voice quality (Baken and Orlikoff, 2000; Fourcin and Abberton, 2008; Gick et al., 2013; Mathieson, 2001; Teixeira and Fernandes, 2015). When the opening and closing movements of the vocal folds are efficient and co-ordinated with a steady breathstream, the resulting signal from the glottal pulse is stronger than any additional 'noise' that may be generated by excessive airflow, giving a measure of **harmonics-to-noise ratio** (HNR, dB). Additional noise results in a comparably reduced harmonic component and lower values of HNR: this is sometimes associated with vocal fold pathology (Mathieson, 2001) and with breathy quality (Yap, 2012; Yap, Epps, Ambikairajah, Choi and Wales, 2015).

The signal from the glottal pulse is filtered and amplified by the resonators within the supralaryngeal tract: this includes the pharynx, oral cavity and nasal cavity. The resonators amplify harmonics produced by the voice to produce resonance frequencies known as **formants**, the first and second of which are crucial to the clear production of vowels (Baken and Orlikoff, 2000; Mathieson, 2001). Formant frequencies 1 and 2 are closely related to the tongue height (F1) and to the anterior-posterior position of the tongue (F2). Formant frequencies and voice perturbations (HNR, jitter, shimmer) are easy to measure numerically and to represent visually using programs such as *Praat* (Boersma and Weenink, 2011). This allows comparison between healthy norms and people with atypical phonation.

When the vocal system is working to optimum conditions, voice quality is 'modal' (Laver, 1980). However, there are many interacting factors that can disrupt this process at the laryngeal and supralaryngeal levels. These include changes to aerodynamic functioning (breath control, breath pressure); differences in muscle tone and functional muscle control within the respiratory or vocal system; differences in the shape of the vocal tract and articulators; and neurological disorders that interfere with muscle movement (Aronson and Bless, 2009; Gick et al., 2013). Any of these can affect the quality of the voice as it is produced and perceived.

Voice production in DS

Individuals with DS are frequently described as having voices that are characteristic of the syndrome (Moran, 1986). Descriptors vary between 'harsh', 'guttural', 'raucous' and 'monotonous' (Benda, 1965; Novak, 1972; Pentz and Gilbert, 1983). In comparison to TD populations, the voices of people with DS are perceived as more nasal, breathy, or hoarse (Moran, 1986; Moura, Cunha, Vilarinho, Cunha, Freitas, Palha et al., 2008; Rodger, 2009). In an early study by Beck (1988), two judges assessed the voice production of 28 young adults with DS (mean age 28;9) and an age-matched TD control group (n=16), using a vocal profile analysis checklist. Beck recorded statistically significant differences in 11 of the 18 measures, which included reduced

movement of articulators (mandible, tongue); perceived differences in nasal production; lax supralaryngeal tension settings; pharyngeal constriction; and higher degree of harsh or whispered phonation. The exact causes of the differences in DS are unclear, as voice quality in DS depends on the configuration of organic properties within the speech organs (Beck, 1988). In addition, voice production depends upon how the organs and muscles involved are used, which can be affected by hormonal, emotional and physiological factors (Laver, 1980). The characteristic voice quality of people with DS is therefore likely to arise from a combination of physiological, anatomical, neurological and functional difficulties and differences. The sections below describe the key resonance and phonatory characteristics and the factors that are believed to affect them. A summary of research into the acoustic correlates of voice production in DS will then be given.

Resonance

Pryce (1994) hypothesised that hypotonicity within the oral cavity or pharyngeal walls may result in the characteristic voice quality (Pryce, 1994). Reduced movement of the soft palate or inadequate closure of the velo-pharyngeal muscles is commonly reported in DS (Section 2.3.1) and will affect the balance of nasal vs. oral resonance (Beck, 2010). A 'lax' pharyngeal wall may dampen the sound and affect resonance. Findings by Rodger (2009) support this. In her study, two speech and language therapists (SaLTs) judged the voices of 22 children and young people with DS against the Vocal Profile Analysis Scheme. In comparison to 52 TD age-matched controls, each SaLT judged the voices of participants with DS to have higher nasality and more lax pharyngeal settings, in common with Beck's findings (1988). Additionally, the smaller oral cavity of people with DS may affect the resonant quality of the voice (Xue et al., 2010; Laver, 1980). Many DS people have a large tongue relative to the oral cavity, which sits back in the mouth, further affecting the size and shape of the vocal tract (Xue et al., 2010; Venail et al., 2004). This difference in the shape of the oral resonator, coupled with difficulties and limitations in moving the articulators and in phonation are likely to contribute to the distinctive voice types of people with DS (Xue et al., 2010).

Phonation

A degree of dysphonia is commonly reported in the syndrome and may be present from birth (Beck, 1988; Kent and Vorperian, 2013). Physiological differences contribute to this. The larynx of people with DS is small and is placed higher within the vocal tract than in non-DS populations (Venail, Gardiner and Mondain, 2004). Pharyngeal and laryngeal anomalies are present (Venail et al., 2004) which may also affect the movement and control of the larynx. Differences in the supralaryngeal tract, such as excessive nasal emission, can affect the balance between supra- and sub-glottal pressure with consequences for the act of phonation (Baken and Orlikoff, 2000). Lax muscles will affect control of the external and instrinsic laryngeal muscles, including the vocal muscles which in turn affects the control of Fundamental frequency (F0) (Beck, 2010; Laver, 1980).

Restricted tongue movement also impacts upon phonation, as a result of changes to sub-glottal pressure and fundamental frequency (Gick et al., 2013; Lin, Jiang, Noon and Hanson, 2000; Stevens, 1997; Honda, 1983; Honda, 1995). Furthermore, the vocal folds of people with DS may be affected by atypical mucosa (Beck, 1988), by hypotonia and by motor-programming difficulties (Howell, 2010; Heselwood, Bray and Crookston, 1995).

Pryce (1994), in studying the activation of the vocal cords in people with DS, concluded that people with DS require more energy to activate their vocal cords as a result of hypotonia. Excessive effort in producing voice may result in hyperfunctional voice disorders, in which the false vocal folds are activated (Aronson and Bless, 2009; Fawcus, 2009; see also Section 2.2.2). This can result in unvoiced ventricular production, or in the generation of two simultaneous pitches (**diplophonia** or biphonia: Titze, 1994) if the vocal folds also oscillate (Cavalli and Hirson, 1999; Dejonckere and LeBacq, 1983; Titze, 1994). The latter has been reported in one case of DS (Beckman, Wold and Montague, 1983; Rodger, 2009).

Acoustic characteristics

Studies into phonation have measured the acoustic properties of the voice. Many have tried to reconcile the acoustic data with the perceived vocal qualities and known neuro-physiological differences. A summary of key findings is given below.

Moura et al. (2008) measured seven voice parameters of 66 children with DS, aged 4-8 years, as they sustained five Portugese vowels, and compared this to data collected from TD children. The DS children performed statistically differently to the control group on all voice measures except for the fundamental frequency of the vowel /u/. The results showed that DS children sustained vowels at a lower fundamental frequency, with greater deviation, and that they had higher perturbations in amplitude and frequency and a greater degree of noise in the vocal signal. Decreased measures in spectral tilt suggested a different use of the glottal source, and are consistent with breathiness and use of force in phonation. The authors observed a qualitatively different spread of formant frequencies in DS vowels compared to controls. They found that F1 was low for /a/ and /o/ which suggests that subjects had limited tongue movement in the oral cavity that prevented a higher tongue movement. The second formant, F2, was lower in frequency for vowel /e/ by 8% and for vowel /i/ by 7% but F2 was higher for /u/ by 22%. The authors suggest that the results reflect restricted anterior tongue movement. Bunton and Leddy (2011) also reported reduced tongue movement in the production of vowels in two adults with DS. This affected vowel quality in their participants, reducing contrast between vowels (Bunton and Leddy, 2011).

Moura et al. (2008) argued that a lower-placed tongue and restricted forward movement may have resulted in lower fundamental frequency in the participants in their study. They also argued that the effect of reduced tongue movement upon laryngeal tension could in turn affect jitter, shimmer and HNR. The DS group in their study produced higher measures of perturbation than the controls, and the degree of perturbation changed with each vowel in a similar pattern observed to that in the control group. The authors state that the link between the tongue and larynx, via the hyoid bone is preserved in the DS group. They proposed that the lower F0 and perturbations might also result from poor functioning of the thyroarytenoid muscle, due to irregular neuromuscular firing rates, and that this might be related to generalised hypotonia, as frequently observed in DS children. However, it is interesting to note that not all vowels were equally impaired, relative to controls. In the DS group, the vowels /i/ and /u/ resulted in HNR and shimmer values that more closely matched those observed in the control group, whereas the /a/ and /e/ vowels were produced with higher relative degrees of jitter. This suggests that the tongue position and tension may differentially affect phonation for some vowels.

In contrast to the findings of Moura et al. (2008), Lee, Thorpe and Verhoeven (2009) found less difference in voice perturbation measures between young DS adults and controls. They measured aspects of phonation in the speech of nine British adults with DS, aged 17-24 years, who were asked to sustain an /ɑ/ vowel, and to read the Rainbow Passage. The participants with DS produced speech with a higher mean F0, normal levels of shimmer, but lower levels of jitter in comparison to age-matched controls (CA) and gender-matched controls. The difference for jitter was statistically significant. The DS group had smaller pitch ranges (of an octave, as compared to 2.5), and spoke at a higher pitch than the controls, with DS males speaking with less variation in pitch than females. In terms of phonation, the DS group were able to produced a sustained vowel for a comparable length of time to controls, although the DS females produced lower Maximum Phonation Time. Jitter and shimmer values were comparable in both groups, although DS males had lower levels of perturbation. The authors suggest that phonation in the individuals studied was within normal limits and that other, supralaryngeal, factors contribute to the perceived vocal quality. However, they also acknowledge that the subjects' previous training in speaking as part of a drama group may have affected results, or that the group may have had shorter vocal cords.

In line with the findings of Lee et al. (2009), Albertini, Bonassi, Dall'Armi, Giachetti, Giaquinto and Mignano (2010) reported some differences between the voices of Italian people with DS and TD controls. They compared the acoustic features of 48 children (mean age: 9;6 years) and 30 adults (mean age: 28;7 years) with DS to those of a TD control group consisting of 60 adults and 46 children. Their results showed that voice production was similar between DS children and controls, who were matched on age and sex, with the exception of spectral quality (coefficient of variation),

which was lower in DS children. However, they noted differences between the adult groups, and more so in males for some parameters. In comparison to older TD controls (mean age 48.1 years), the voices of the adult males with DS were characterised by statistically significant lower energy levels (in decibels), by higher mean F0, and by reduced spectral quality (coefficient of variation). The authors also noted some difference between males and females with DS. Males in the DS and control groups produced higher shimmer values after puberty, and higher values than females in both groups. However, in contrast to findings by Lee et al., (2009) the value of shimmer in adult male participants with DS was significantly lower than that of male controls. Albertini et al. (2010) relate these differences in performance before and after puberty to changes in the anatomical structure of the palate - which they state become more noticeable in males after puberty - and to differences in respiratory function, affecting phonation. However, the authors acknowledge that children with hearing loss were excluded from their study, which may have affected the data.

Whilst most research agrees that the voice of those with DS is perceived as atypical, research using acoustic measurements presents conflicting reports as to the link between perceived voice and the physical production of voice. The role of articulation, and the resonance properties of the vocal system are argued to influence the perceived sound (Kent and Vorperian, 2013), and developmental and individual differences in vocal production may further influence overall voice quality (e.g. Lee et al., 2009). However, studies have traditionally excluded those with additional difficulties such as hearing impairment (HI) which is prevalent in DS, especially in children (Laws and Hall, 2014; Maris et al., 2014: and see Section 2.6). Depending upon the severity of hearing loss and upon the age of onset and management, HI can significantly affect phonation (Coehlo et al., 2015; Bolfan-Stosic and Simunjak, 2007; Monsen, 1979) and voice quality (Coehlo et al., 2015; Fourcin and Abberton, 2008). Given that phonation in fluent speech depends upon the production, co-ordination and auditory monitoring of voice for both segmental production and suprasegmental production (Fourcin and Abberton, 2008), its role is fundamental to the examination of prosody in DS. In particular, it has implications for the timing of speech sounds (affecting rhythm), for the control of pitch (affecting intonation) and for the combination of rhythm, pitch and intensity required to produce lexical stress.

2.4.3 Prosody

The perception of prosody in speech is necessary for the understanding of speech, and difficulties in perceiving prosody will affect the intelligibility of produced speech. The perception of prosody may also contribute to the learning process by enabling the segmentation of speech into phrases and words (e.g. Mason-Apps, Stojanovik and Houston-Price, 2011). This section will examine the evidence that people with DS have difficulties in perceiving or producing speech rhythm and intonation.

Speech rhythm

Research suggests that people with DS may have difficulties in perceiving word rhythm. A study by Mason-Apps et al. (2011) found a delay in the ability of DS infants to perceive weak-strong syllables, relative to TD controls, indicating delayed sensitivity to the rhythm of English speech. Their study does not confirm when weak-strong and strong-weak rhythms are perceived, but there is some evidence for a deficit in the perception of word stress in older children. Pettinato and Verhoeven (2009) examined the abilities of 16 children and adolescents with DS (aged 11-20 years) to repeat and to discriminate nonsense words with simple and complex stress patterns. Their DS participants each had a receptive vocabulary age of over 4 years (4.06-7.07 as measured by BPVS II), the age at which most English children have mastered native complex stress patterns. A group of TD children matched on sex and BPVS scores (aged 4-7 years) acted as controls. The controls were at ceiling in both tasks. They found that the accuracy of repetition deteriorated in the DS participants as the target syllable length increased in words of 2, 3 and 4 syllables. However, in words of 4 syllables, weak-initial words and words with complex stress structures were produced less accurately than weak-initial words. Although the differences as a result of stress pattern were not statistically significant, they concluded that the development of stress patterns in the DS subjects was delayed, relative to their receptive language abilities. A second experiment was conducted to rule out the effects of auditory memory on performance. The study required participants to identify, using an animated computer task, which of two imitations of a nonsense word was accurate. The results showed a statistically significant deterioration in the DS group for word length and for position (weak-initial, and word stress in compound structures). Pettinato and Verhoeven concluded that these stress structures were poorly encoded in their participants with DS. They also argued that the acquisition of these patterns was delayed, relative to TD controls, with consequences for the perception of words in continuous speech, and for the rhythmic production and intelligibility of their speech.

Dysfluencies are common in DS and it is estimated that between 10% and 45% of people with DS have problems in producing fluent speech (Kumin, 2006; Kent and Vorperian, 2013). Cluttering is another motor-speech disorder associated with stuttering, in which the speaking rate may be rapid and uneven (Alm, 2004). It is believed to be prevalent in the speech of people with DS, often in addition to stuttering (Howell, 2010; Kent and Vorperian, 2013; van Borsel and Vandermeulen, 2008). Even in fluent DS speech, studies indicate that rhythm is atypical (Kent and Vorperian, 2013) and that this occurs primarily at the syllabic level, which may play a crucial role in maintaining intelligible speech (Greenberg and Arai, 2001). For example, the deletion of syllables is common in DS speech which may simplify rhythmic production (Bray, 2008). In people who stutter,

the reduction of linguistic content can support fluency, which Bosshardt (2006) attributes to a reduction in the demands imposed by speech planning.

A case study by Heselwood et al. (1985) demonstrates that syllable deletion may be linked to linguistic planning, and to rhythmic planning in particular. Their subject, a 34 year-old DS male, simplified the rhythm of many words, including words whose initial syllables were strong. Although fluent, he demonstrated some features of cluttering, which affected the rhythm of his speech. His production of words before a pause were articulated either more accurately, or more intelligibly. They speculate that the rhythmic and articulatory disturbances they found in their speaker resulted from a disturbance in planning intonation groups. In particular, in this subject, the inclusion of a semantically empty 'tag' (*isn't it*: [ɪnti]) that he used at the end of phrases was produced most clearly but appeared to 'squash' the rhythm of the preceding phrase. This resulted in shortened syllable length and segmental errors. Rather than being the consequence of articulatory problems or motor-control problems, Heselwood et al. argue that the error arises from an error in phonetic planning, in which the final section of the phrase is 'downloaded' before the preceding section. Furthermore, pre-pausal phrases were produced with less breathy phonation, which they argue may result from reduced initiator power and poor respiratory control.

Flipsen (1999) examined some of the features raised by Heselwood et al. (1995) by analysing pre-existing transcriptions of speech from six adults with DS. He found some evidence that his subjects produced syllables more intelligibly, but not more accurately, just before a natural speech pause, partially supporting the earlier findings of Heselwood et al. (1995). Flipsen recognises that methodological differences may account for the different findings. In particular, the pre-existing narrow transcriptions may have affected the calculations of segmental accuracy, but not intelligibility, which was based upon whether the word had been glossed in the transcript.

In summary, the studies indicate the presence of perceptual rhythmic deficits in DS. These may affect rhythmic production of some words. However, the research suggests errors of motor planning that may affect the production of syllables through the simplification and reduction of speech sounds, and by reducing syllable length. This may be further affected by planning demands, and by physiological limitations, such as reduced breath control. The latter, in turn, can also affect phonation, which will have consequences for the control of intonation in speech.

Intonation

Studies into speech intonation in DS are limited but research indicates that that the majority of individuals use a smaller vocal range than TD populations and produce inflectionless speech (e.g. Moran, 1986; Montague and Hollein,1973; Moura et al., 2008), which may affect their ability to use

pitch changes to convey linguistic or emotive meaning. Lee et al. (2009) found that in non-emotive speech, their nine young adults with DS (see Section 2.4.2) used significantly less variation in pitch at sentence accent and changes in pitch at boundaries than controls. They also found that male and females with DS had a significantly narrower pitch range than age- and sex-matched controls, of less than one octave. Compared to controls, the DS group had a higher average speaking pitch, and reduced variation from the mean pitch, implying less natural melodic variation; and the study showed a difference in the amount of declination - the gradual reduction of speaking pitch at the end of phrases - in the DS subjects. However, as the authors acknowledge, the use of the 'Rainbow Passage' is challenging for this group in terms of speech production and comprehension, both of which may have affected the results.

Stojanovik (2011) assessed both comprehension and production of prosody in nine children with DS. She found that compared to TD controls matched for age, they performed poorly in all tasks but scored more highly in comprehension than production. Compared to controls, the DS children were significantly poorer at discriminating contrasting intonation patterns and they failed to use a fall or rise in pitch to signify meaning. The DS children were able to perceive the pitch changes, however, and in contrast to controls, they were better able to perceive lexical stress than to produce it. Likewise, Zampini, Fasolo, Spinelli, Zanchi, Suttora, and Salerni (2015) found that Italian children with DS are less able to signal interrogative sentences with a rising tone. They reported differences in some prosodic abilities between nine young DS children and TD children of comparative age and vocabulary size. In comparison to controls, the DS group produced fewer spontaneous multisyllabic utterances, used a significantly lower F0 and their highest F0 was also lower than controls.

Intensity

Finally, changes in vocal intensity, perceived as volume changes, combine with changes in syllable duration and pitch to create word or syllable accent, or 'stress'. Control of vocal intensity for speech intonation for both pitch and volume depends on the co-ordination of vocal cords and breath flow (Baken and Orlikoff, 2000; Gick et al., 2012; Ladefoged, 2005; Strik and Boves, 1992). Whilst there is little direct evidence regarding control or use of speech intensity in DS, lung capacity is often small, and hypotonia may affect the respiratory muscles (da Silva et al., 2010; Laibsinion et al., 2012; Salgueirinho, Venâcio, Martín-Nogeuras and Ribero, 2015). The effect of this on phonation in people with DS has not been studied, but it is believed to affect both the quality of phonation and the intensity of the sound (Howell, 2010; Heselwood et al., 1995). Furthermore, the control of intensity for speech is affected in people with HI (Coelho et al., 2015). This puts people with DS at risk of producing atypical intonation, particularly during early childhood when hearing loss is most prevalent (Maris et al., 2014).

2.4.4 Section summary and relevance to the study

The speech and voice production of people with DS is affected by structural, muscular and neurological differences. As a result of intermittent or congenital hearing loss and neurological differences, they may also have difficulties in perceiving speech sounds and aspects of prosody. Previous sections have shown that singing and speech activities may support aspects of speech perception and production in other populations, and voice production in people with functional voice disorders. Given that people with DS are able to develop motor skills (Section 2.3), people with DS might therefore benefit from singing in terms of voice production. Similarly, the slower, exaggerated presentation of words in song might support their perception of speech sounds, words or prosody.

2.5 Intelligence and cognition in DS

On average, DS adults reach a MA-equivalent of 6-8 years but there is considerable variability within the population and uneven development of particular cognitive sub-skills (Capone, 2004; Couzens and Cuskelly, 2014; Silverman, 2007). Cognitive development in people with DS is not fully understood and may depend upon a range of environmental and personal factors (Couzens and Cuskelly, 2014), including maternal education, the severity of medical conditions, personal temperament, and early school experience (Couzens, Haynes and Cuskelly, 2012; Määttä, Määttä, Tervo-Määttä, Taanila, Kaski and Iivanainen, 2011). The majority of people with DS are also at higher risk of co-morbid conditions, such as autism, ADHD and epilepsy, that can further impact upon learning ability and quality of life (Alexander et al., 2015; Schieve, Boulet, Boyle, Rasmussen and Schendel, 2008). Despite the considerable variation in cognitive ability of individuals, the syndrome is characterised by common differences in intelligence, cognitive development, and in cognitive processing (Chapman and Hesketh, 2000; Capone, 2004; Dierssen, 2012; Silverman, 2007). The following section will begin with an overview of intelligence and cognitive development in DS. This section will be followed with an overview of sensory and cognitive processing abilities in people with DS. It will then examine the evidence that they process sound and verbal information differently to TD populations. Their abilities to perform executive functions, such as the ability to attend to tasks and to process information, will then be discussed, with a specific focus on the difficulties that DS people face in verbal memory.

2.5.1 Intelligence

Studies that compare abilities to TD children show a general cognitive delay (for a review, see Patterson, Rapsey and Glue, 2013). The range of IQ within the DS population is approximately 60 points between average maximum and minimum, similar to that of the TD population (Buckley and Bird, 2002). The majority of people with DS achieve IQ scores of 25-55 equivalent to a label of moderate-severe learning difficulty (Tsao & Kindelberger, 2009), but a small percentage can

achieve scores of up to 85 (Tsao and Kindelberger, 2009). In terms of MA, Iacono, Torr and Wong, (2010) reported scores of 2 years -10 years in 55 adults aged 19-58 (mean age = 38 years), using a measure of receptive vocabulary (Peabody Picture Vocabulary Test-III). The mean MA score was 5;2 years, with a Standard Deviation of 2;2 years. These results were not related to age, once the researchers excluded scores for ten adults with Alzheimer's Disease.

In children with DS, most studies report that IQ declines relative to development (Capone, 2004; Channell, Thurman, Kover and Abbeduto, 2014; Patterson et al., 2013). Sigman and Ruskin (1999) found that in comparison to controls with autism and developmental delay, the majority of their 71 participants with DS showed a decline in IQ of 20 points between 2-3 years and 10-11 years. Thus, their intelligence as measured by general intelligence tests do not increase in line with their chronological age (CA), unlike TD populations (Chanell et al., 2014). Although some studies report a plateau in cognitive development by the age of 10 years (Capone, 2004), other studies show an overall increase in MA during the ages of 9 and 15 years (Iacono et al., 2010) and that some aspects of cognitive skills may continue to improve beyond 20 years of age (Wuang and Su, 2011).

Recent research shows that different aspects of intelligence may develop differently in DS subjects, in comparison to TD children or those with other learning disabilities (Patterson et al., 2013). In a cross-sectional study, Couzens, Cuskelly and Haynes (2011) examined the test results of 208 children, adolescents and young adults on the Stanford-Binet (SB)-IV subtests. The participants who were aged 4 years were at least 8 months behind TD peers in measures of vocabulary, as measured by the Stanford-Binet Intelligence Test IV (SB:IV) (Couzens et al., 2011). However, they found different patterns of ability in older participants: the performance of DS subjects on the visual reasoning and pattern analysis subtests of the SB:IV continued to develop beyond the age of 24 years, as measures of verbal reasoning and vocabulary decline (Couzens et al., 2011). In a longitudinal study, Channell et al. (2014) reported a change in visual reasoning and visuo-spatial perception over time in adolescents with DS. They measured the nonverbal intelligence of 20 young males (aged 10-15 at the onset of the study) over a three year period, using the Leiter-R scale. Their mean MA-equivalent was 4.70 (range 3.13-6.08) at their first assessment, with seven participants scoring at floor level. Although the IQ scores did not change with age, the raw composite growth scores increased annually, at statistically significant levels. However, statistically significant changes were found in just two of the four subtests. The data indicated that aspects of verbal intelligence continued to develop during adolescence. The data from both studies suggest a specific weakness in 'fluid' reasoning, that is associated with executive function, such as working memory and the ability to shift attention. These issues will be discussed further in the next section.

2.5.2 Cognition and language

Cognitive Processing

Wishart (2001) reported on a series of experiments that traced the development of 38 DS infants and children between birth and 3 years. She recorded differences in how they responded to tasks that required them to locate objects in tasks of increasing difficulty. The studies found that infants used a variety of 'avoidance' tactics during more challenging tasks, and showed inconsistencies in their performances. Wishart proposed that instability in response, inadequate motivation and inefficient learning processes contribute to difficulties in consolidating new skills. Such differences in behaviour may relate to differences in how people with DS process information (Will, Fidler, and Daunhauer, 2014). Infants learn through active exploration and engagement with their environment (Kolb and Whishaw, 2006), and difficulties in perceiving or integrating information will interfere with their ability to make sense of their environment. In addition to peripheral sensory impairments (see Section 2.3), people with DS may have brains that process sensory information differently, or more slowly, than TD populations (Glenn et al., 1981; Silverman, 2007; Will et al., 2014; Wuang and Su, 2011). The sections below examine the evidence that people with DS respond differently than TD individuals to cognitive tasks, from infancy.

Global vs. local processing

Porter and Coltheart (2006) identified differences in how DS individuals attend to visual information, relative to people with William's Syndrome, autism and to age-matched TD controls. Fifteen people with DS (aged 6;9-31;3 years) were assessed on their ability to copy shapes that were within the capabilities of four-year old TD children, and to name either the global shape of a pattern or to find and name a small shape within a global pattern. In comparison to all controls, the group with DS showed a bias for the global form when identifying shapes.

Auditory processing

Differences in processing sound have been noted in DS infants, with differences noted in auditory brain-stem response (Roizen, Wolters, Nicol and Blondis, 1993; Pekkonen, Osipova, Sauna-Aho and Arvio, 2007) and speech processing (Keller-Bell and Fox, 2007; Marcell, Cohen, Weathers, Wiseman, Croen and Sewell, 1990; Marcell, 1995). Behavioural studies indicate that people with DS process speech sounds differently than TD populations: in dichotic listening tests, in which sounds are simultaneously presented to both ears, they are better able to discriminate sounds printed to their left ear, suggesting that the right hemisphere may be more strongly involved in processing speech, than the left (Chua et al., 1996; Elliott and Weeks, 1993). Elliott and Weeks proposed that communication between hemispheres is reduced or slow (Elliott and Weeks, 1993).

In behavioural studies, Marcell et al. (1990) and Marcell (1995) reported that 26 adolescents and young adults with DS had greater difficulties than intelligence-matched and age-matched controls in discriminating speech sounds. Although they reported a link to HI, their results also indicated a processing deficit. In comparison to controls, they had specific difficulty in repeating digitised words when a masking noise was presented within 40 ms. The results indicated that the DS participants were slower in processing the words than controls when imitation was required (Marcell et al., 1990). Marcell (1995) further reported that when participants were asked to identify words by pointing to pictures, the DS group produced fewer correct responses than controls when the digitised words were masked within 40 ms of their production. He reported a correlation within the DS group between performance on both word-discrimination tasks and the degree of hearing loss. The DS group also had greater difficulty than controls in identifying minimal pairs that differed in initial stop-consonants, but this was not correlated to HI. Marcell concluded that low auditory acuity might interact with slower processing of speech sounds. He argued that this difficulty extends to the processing of consonant sounds, but not to vowel sounds. However, he acknowledged that the link between hearing and performance on these tasks may be influenced by other individual cognitive or health-related factors, and may not be unique to DS.

Although the effect of hearing on auditory processing has not been followed up in the literature, atypical processing of speech has since been confirmed by electrophysiological methods that measured changes in the brain's response (Groen, Alku and Bishop, 2008). Groen et al. found that in comparison to age-matched TD controls, 19 DS children without hearing loss showed reduced response time and differences in the patterns of lateralisation of sounds in response to semi-synthesised vowel sounds, simple tones and complex tones. Chen et al. (2015) also reported atypical processing of speech, but not of rhythm or melody. They examined the neurological response of 14 DS adolescents and adults (mean age 21.57 years) as they responded to different modes of instruction in a drumming task. Instructions to drum were given every 500 ms under three different conditions: verbal, rhythmic or melodic. Tasks were trialled and repeated until 12 seconds of data was produced on the EEG computer. The DS participants showed a stronger response to verbal stimuli in the RH. However, as for TD groups, they processed melody in the RH and rhythm in the LH. The authors conclude that processing of speech is atypical, but that processing of music is typical.

Higher-order cognitive factors

Language

People with DS have a difficulty in expressive language, relative to their non-verbal cognitive ability (Cleland et al., 2010; Chapman and Hesketh, 2001). The development and mastery of phonology and syntax are delayed (Chapman et al., 1991), and utterances are simpler in structure (Laws and

Bishop, 2003). Receptive language is a relative strength (Chapman and Hesketh, 2001), but performance on sub-tests shows an uneven profile (Abbeduto, Peveto, Kesin et al., 2001; Naess et al., 2011). Abbeduto et al. (2001) reported deficits in receptive grammar, and in tasks involving word classes and word relations in 25 adolescents and young adults with DS. In their review of 15 studies, Naess et al. (2011) found that DS children performed 1SD below controls matched for nonverbal MA in receptive grammar tasks, but that receptive vocabulary skills were not significantly weaker. Their meta-analysis also confirmed a productive deficit in expressive vocabulary tasks, with DS children performing more than 0.5SD below controls. Naess et al. (2011) suggest that whilst receptive vocabulary skills appear to be less impaired than expressive skills, this may be an effect of the relatively reduced processing demands required of receptive language tasks. Other sources also suggest that non-linguistic factors may affect expressive language in the population: Pryce (1994) argues that the effort required to initiate phonation may lead to shorter utterances; and Martin et al. (2009) suggest that people with DS might simplify sentence structure in order to simplify speech production (see Section 2.4.1).

Executive Functions

Executive functions are the set of cognitive processes that are involved in processing information and in regulating and adapting behaviour. These include basic responses, such as attention and memory, and more complex processes such as the ability to inhibit responses to distractors, the ability to retain information in memory for processing and planning, and the ability to shift attention between stimuli (Will et al. 2014). For TD children, performance on tasks that measure these functions shows that skills emerge in early infancy and develop during early childhood by about 5 years (for a review, see Will et al., 2014). Difficulties in memory are well-documented in the DS population and a specific deficit in verbal memory is evident (Baddeley and Jarrold, 2007; Carney, Brown and Henry, 2013; Costanzo, Varuzza, Menghin, Addona, Gianesin and Vicari, 2013; Lanfranchi, Jerman, Dal Pont, Alberti, Vianello, 2010) and will be discussed further in the section below. In comparison to MA-matched controls, people with DS demonstrate difficulties in sustaining attention (Lanfranchi et al., 2010) and in shifting attention in the verbal domain (measured by producing words in alternating semantic categories) (Carney et al., 2013; Costanzo et al., 2013). They also demonstrate difficulties in sustaining auditory attention by silently counting auditory tones (Costanzo et al., 2013), and in verbal inhibition tasks (the Stroop test: naming words that are printed in a different colour) (Costanzo et al., 2013). Carney et al. (2013) found that in older children and adolescents, most EF skills are in line with overall cognitive development. However, their subjects were delayed in comparison to TD children in their abilities to process verbal and visual information in memory, and in shifting responses.

Memory

Components of long term memory (LTM), short-term memory (STM) and working memory (WM) are impaired in persons with DS (Baddeley and Jarrold, 2007; Chapman and Hesketh, 2001; Jarrold, Baddeley and Phillips, 1999; Laws, 1998; Marcell and Weeks, 1998; Vicari and Carlesimo, 2006). In LTM, implicit memory, requiring no conscious recall, is relatively unimpaired; but explicit memory, which requires deliberate recall of information, is weak (Vicari, Bellucci and Carlesimo, 2000; Jarrold, Thorn and Stephens, 2009). There is a distinct STM profile in people with DS in terms of both how the senses are processed, and in the relative strengths and weaknesses of the conceptual memory systems themselves. For example, auditory memory appears to be a specific weakness; components of visual memory, by contrast, are relatively unimpaired (Baddeley and Jarrold, 2007; Jarrold and Baddeley, 1997; Lanfranchi, Carretti, Spanò and Cornoldi, 2009; Vicari, Belucci and Carlesimo, 2005, 2006). Working memory – the system, based on the model by Baddeley and Hitch (1974), that holds and processes information in different sensory modalities – is specifically weak in DS individuals. There is believed to be a particular weakness in the phonological loop, a component of WM that stores phonological information temporarily, allowing subvocal rehearsal. This, according to Purser and Jarrold (2005), is a deficit of storage, rather than of rehearsal. Several studies have shown that DS individuals have a reduced ability to repeat verbally-presented sequences of numbers (the 'digit span'), in comparison to MA-matched controls (Kay Raining-Bird and Chapman,1994; Kay Raining-Bird, Cleave and McConnell, 2000; Laws, 1998; Seung and Chapman, 2000). In a study by Seung and Chapman (2000), the mean digit span of 35 DS children and adults (aged between 9.8-24.3 years) was measured at 3.68, whereas the TD children in their study scored 5.15 digits (aged between 2.3 and 6.8 years). Differences between the groups were statistically significant. Research shows that digit span does not change significantly over time (Patterson et al., 2013).

The ability to process information in the central executive — the theoretical component of WM that regulates attention — appears to be diminished in children with DS, especially for verbal information (Carney et al. 2012; Lanfranchi, Baddeley, Gathercole, and Vianello, 2012). Furthermore, performance is weaker when attention must be divided between two tasks, whether both tasks are in the same or different mode. Lanfranchi et al. (2012) conducted a study with 45 DS individuals matched to 45 younger TD children for verbal MA. Participants were assessed on a battery of tasks that required them to remember one piece of verbal or visual information and to perform a second task on cue, requiring attention to be split between two tasks of one or two modalities. For example, in a task requiring both verbal and visual modalities, they were asked to remember the first word in an aurally-presented list of two or three words and to tap a table when shown a red card. The DS participants showed impairments in verbal memory in comparison to younger TD controls but no deficit in their performance on visual tasks. All DS participants were

impaired on all dual tasks, but with poorer performance on verbal tasks. The findings support other evidence of an impaired phonological loop in persons with DS, which affects verbal memory (e.g. Baddeley and Jarrold, 2007), and of a deficit in the central executive when two tasks are processed simultaneously (Lanfanchi et al., 2009, 2010). The authors suggest that the latter may be a feature of learning disabilities in general.

A specific deficit in verbal WM may be implicated in the persisting language and speech difficulties observed in DS. In their meta-analysis of fifteen studies, Naess, Lyster, Hulme and Melby-Lervag (2011) found that verbal memory was impaired in 145 DS subjects compared to 201 controls matched for nonverbal MA, with DS subjects performing one SD lower. They suggest that this may affect development of grammar. It has also been suggested that a deficit in verbal STM may explain why receptive language is often superior to expressive language (Gathercole and Baddeley, 1993). However, Jarrold, Thorn and Stephens (2009) contest the role of impaired phonological memory in reduced language on the grounds that both lexical development and the ability to learn novel words quickly are spared; they report poorer performance on tasks requiring the recall of novel words, an act of memory that might require 'finer-grained' memory. This weakness in the auditory memory will slow the process of learning the mental representation of the word and its appropriate motor movements, with implications for the learning of speech sounds (Stackhouse et al., 1997). Using the visual channel to assist auditory learning may improve this learning process: a study by Gibbon, McNeil, Wood and Watson (2003) showed that speech remediation was more successful when accompanied by a form of visual feedback, for a 10 year-old female with DS.

2.5.3 Section summary

Research shows that people with DS exhibit different behavioural responses to cognitively challenging tasks. These difficulties may result from differences in how their brain processes information, and from specific areas of difficulty or ability perceiving, processing and storing sensory information. In particular, a weakness is noted in processing verbal and auditory information. However, people with DS are able to continue to develop new skills into their thirties, and may be better able to learn when information is visual. If teaching is tailored to their strengths, they may be able to learn new skills in the musical domain. The next section will examine the evidence that people with DS, and with other learning disabilities, are able to develop musical skills, and whether there is any precedence for such skills supporting adaptive behaviour or verbal communication skills.

2.6 Musicality in people with learning disabilities and DS

The term ‘musicality’ is used in this section: as in Kodaly teaching, this signifies sensitivity to music as well as musical ability and musical learning (Houlahan and Tacka, 2008). Whilst there is very little evidence of the musical abilities of people with DS or with general learning disabilities, there is evidence of musical sensitivity and an ability to learn. As there is a dearth of research in both populations, this section will review the evidence for both, beginning with an overview of musical ability in people with developmental difficulties. This information will be used to inform the design of assessment and of teaching.

2.6.1 Musicality in people with developmental difficulties

Children and adults with a wide range of developmental and cognitive disabilities benefit from musical activities in terms of communicative abilities (Carroll, 1996; Coyne, Dwyer, Kennedy and Petter, 2000; Groß et al., 2010; Hoelzley, 1993; Lim, 2010; Perry, 2003; Simpson and Keen, 2011). Despite this, research into musical abilities and learning in people with developmental disorders is limited (Hooper, Wigram, Carson and Lindsay, 2008; Stephenson, 2006). There are just three key studies that have examined musical ability or development in people with different levels of cognitive ability. These will be examined in some detail.

MacDonald, O’Donnell and Davies (1999) considered the educational value and therapeutic value of musical activities for 20 people with mild to moderate learning disabilities, aged 17-58 years. The intervention group received training in playing the gamelan; one control group received training in cooking and art; and a second control group did not receive additional training. Each group was tested before and after the intervention periods on rhythm production and pitch perception, on the Communication Assessment Profile for Adults with Mental Handicap (CASP), and on their perception of their ability to play the gamelan. All programmes ran once a week for ten weeks. The music group made statistically significant improvements in measures of imitating simple rhythms (clapping), instrumental rhythms, and section 3 of CASP, which measures pragmatic communication, and involved picture identification and picture naming tasks.

A study by DiGiammarino (1990) gives some insight into the abilities of learning-disabled adults and the relationship between specific skills and level of cognitive ability. She studied 45 musical skills in 120 adults aged between 21-80 years, with mild-severe intellectual disability. A checklist of ‘functional’ listening and performance skills were drawn up in collaboration with music therapists. Care staff completed the checklists against descriptors such as ‘performs single rhythms in time to music’. The data was analysed according to cognitive descriptors of learning disability (LD): Mild LD, Moderate LD (MLD), or SLD. All groups found complex motor and vocal skills difficult and the percentage of people able to perform more complex musical skills decreased with overall cognitive

ability. For example, the ability to play rhythmic patterns was observed in just 22.5% of those with Mild LD, 35.8% of those with MLD, and 15.4% of those with SLD. However, the high percentage of participants credited with being able to move in time to the music (over 80%) suggests that the method of assessment by staff was not reliable, and there was a lack of data regarding previous musical experience.

Recent research by Ockelford et al. (2011) and the *Sounds of Intent* project (http://soundsofintent.org) confirms that children with complex and profound learning disabilities and SLD are able to develop musically. Teachers worked alongside researchers to chart the development over a period of 6 months of 20 individuals (aged 11;11 to 17;7 years) who took part in weekly 45-minute sessions of song-based musical activity. The teacher recorded observations on six occasions of the group's abilities, across six descriptive levels in three dimensions of learning: 'listening and responding', 'causing, creating and controlling' and 'interaction'. The group made progress in all areas of development. Although the authors attributed much of the group's rapid development to increased confidence and familiarity for the staff and pupils alike, the data showed an ability for learners to develop.

2.6.2 Musicality in people with DS

The research demonstrates that children with learning disabilities have a capacity to develop musical skills through music, regardless of the level of impairment. It has long been noted that people with DS have musical interests, are often motivated and more musical than other groups of people with intellectual and developmental disabilities (Lense and Dykens, 2011). Despite the well-documented interest in music, very little research has been conducted into musical abilities with DS. Early studies were conducted with 'institutionalised' persons with DS: one by Blacketer-Simmonds (1953) observed 42 DS subjects and 42 with other learning disabilities. Six in each group showed an interest in music and could copy rhythms and move in time, but none were able to hum or sing 'satisfactorily'. A tapping task (Cantor and Girardeau, 1959) showed that DS children and adolescents were impaired in their ability to tap to a beat, compared to TD toddlers, who were not matched for MA. However, detail is lacking concerning actual musical ability and the level of cognitive ability of those described.

DS infants show similar interest to singing as infants, as measured by eye gaze (de l'Etoile, 2015; (Glenn, Cunningham and Joyce, 1981). However, they display atypically prolonged eye-gaze when their mothers sing to them, which may indicate that they may require more time to process the words or musical information in songs than TD infants (Glenn et al., 1981). Singing has been recommended as a means to supporting communication (Barker, 1999) and working memory

(Bennet, 2009) in DS children but there are minimal published accounts of their abilities in this regard. In 1991, Edenfield and Hughes reported that no studies were available that referred to singing ability in DS, prior to their own study. They studied the singing abilities of 13 students (aged 13-23 years) with DS who attended a school that provided a choral curriculum. A control group of nine DS subjects, matched on average IQ and age, were used in order to test for differences. The participants in the singing group had taken part in choral lessons between two and four times a week, and had attended such sessions for an average of 4.2 years (range of 1-6 years). Participants in both groups were assessed against five categories of singing that were designed by the authors: articulation, melodic rhythm, melodic contour, steady beat, and pitch. There was no statistical difference between the composite scores or in sub-tests for each group. However, the group who undertook choral lessons scored more highly on all tasks than the control group. Their data do suggest a relatively higher ability in the students who undertake teaching. This is in line with an earlier study that indicated that the use of singing software with a DS male resulted in an increased singing range (Spitzer, 1989, as cited in Peters, 2000).

Research into musical ability in people with DS is similarly limited. Stratford and Ching (1983) investigated the abilities of ten DS children (CA 13.2 years; MA 3.8 years) to entrain to simple binary rhythms (*ta ta*; *tee-tee*; *ta tee-tee*), and compared their abilities to a TD group matched for MA (CA 4 years; MA 3;10 years). They concluded that the DS children were not impaired in rhythm repetition tasks compared to MA matched controls. A further study by Stratford and Ching (1989) found that in comparison to non-DS MA-matched controls, individuals with DS were significantly impaired in making rhythmic 'hammering' movements to music. The DS group were also less accurate in marching and clapping, but not statistically so. However, the DS group were scored more highly on 'butterfly' movements than non-DS controls. The authors concluded that the DS subjects have a sense of musical timing, but that they are hampered by motoric limitations in some movements.

Based on observations of community music and music therapy sessions that took place across a range of regions from Wales and the West of England, Picard (2009) documented the rhythmic characteristics of people with DS of different ages (toddlers to old age pensioners). She comments on the abilities of a teenager and a young adult with DS to produce and sustain quite complex rhythms when tapping on drums, in accompaniment to a music therapist. She reports that a younger child (aged 7 years) was able to maintain a simple beat, once the therapist adjusted tempo to match their playing. The boy, identified as C, was one of three boys of the same age who were also able to anticipate the end of the music, and to signal this with a change in behaviour by becoming still, or by slowing their own playing to end with a flourish. Picard commented that C, demonstrated difficulty in controlling his motor movements when drumming, but that he showed

awareness of tempo changes. Rhythmic patterns were observed in a structured 'call and response' activity, in which participants were required to clap back rhythmic patterns. In observing girls with DS, Picard wrote that the rhythmic production was 'relatively accurate' but that responses were inaccurate in the onset of their timing relative to the beat. She noted that in this session, and in others, some participants with DS responded sooner than required — sometimes, before the rhythm had been fully demonstrated — whilst others responded later. Individuals were also inconsistent in the timing of responses relative to the beat. Despite poor alignment of rhythms, Picard deemed that the rhythmic reproduction of most people with DS that she observed were accurate. However, as with the studies by Stratford and Ching (1983, 1989), these judgements were based on perception only.

Recent studies into bimanual drumming tasks confirm that DS subjects are impaired in their motor-timing abilities, relative to controls. A study by Ringenbach et al. (2006) revealed differences in the motor movements of those with DS, in comparison to CA-matched controls, but not to MA-matched controls. The 10 adults with DS (mean CA 30.2 years; mean MA 7.2 years) produced curvilinear motor movements when using bimanual movements to simultaneously strike two drums when asked to play to verbal instructions (the word 'drum', repeated at 1000 ms intervals) or to the sound of a drum, or to a visuo-spatial instruction. They found that the movements of the DS participants were more linear and in line with those of the TD-control group when the visuo-spatial stimulus demonstrated the movements. It is notable that in all conditions, the DS group were less accurate than TD controls in their ability to play to the beat, or in timing movements consistently. Although this is consistent with the findings of Stratford and Ching (1989), the slow tempo (60 bpm) in the study by Ringenbach et al. (2006) may have adversely affected their performance (see Section 2.1.2: Drake et al., 2002).

Reports from music therapy have indicated that visual teaching methods are successful in improving piano skills in individuals with DS (Velasquez, 1991, cited in Peters, 2000). Data from the drumming task by Ringenbach et al. (2006) confirmed that people with DS improve aspects of their performance in response to visuo-spatial instruction, relative to auditory and verbal instruction. In a similar study, Chen et al. (2015) confirmed that DS subjects may process verbal instruction differently to TD controls. They assessed the abilities of 14 DS children and adults (aged 12-39) to drum to verbal, rhythmic, and melodic stimuli at intervals of 500 ms. Neuroimaging was used to measure which area of the brain was dominant in each condition. The data indicate that the DS participants processed musical instruction bilaterally, with rhythm being processed primarily in the left hemisphere and melody in the right hemisphere, which is typical of TD populations, However, their data indicated atypical processing of verbal information, in line with earlier findings (Bunn et al., 2007; Chua et al., 1996; Elliott et al.,1994). The data indicate that

performance and development in music tasks may depend upon the mode of instruction, and that visual methods may enhance performance in this group, as it does for young children (see Section 2.1).

2.7 Summary and hypotheses

Benefits have been observed in TD populations in response to music and singing activities, especially in terms of improved auditory processing (Kraus and Chandrasekaran, 2010; Wan and Schellenberg, 2013). There is emerging evidence that this may lead to enhanced processing of speech sounds (Kolinsky et al., 2009; Kraus et al., 2014; Tierney et al., 2015), and to improved syllable segmentation (Marie et al., 2011). In addition, evidence from music therapy indicates that singing may support speech production and voice production in people with developmental speech (Groß et al., 2010; Wan et al., 2010) or voice disorders (Lee and Son, 2005; Ogawa et al., 2014). Relative to TD controls, the speech of people with DS may be affected by perceptual weaknesses, such as reduced perception of speech sounds (Keller-Bell and Fox, 2007; Marcell et al., 1990; Marcell, 1995), poor representation of speech sounds (Cleland et al., 2010); differences in perception of word stress (Pettinato and Verhoeven, 2009; Mason-Apps et al., 2011); differences in how the brain processes auditory information (Bunn et al., 2007; Chua et al., 1996; Elliot et al., 1994); and by production difficulties such as motor muscle weaknesses in the vocal and respiratory system, and motor programming differences and difficulties (Heselwood et al.,1985; Pryce, 1994; Stoel-Gammon, 2001; Rondal, 2009). Studies indicate that DS subjects are motivated to listen to and perform music (Lense and Dykens, 2011), and that they have a capacity to improve performance in motor-rhythm tasks (RIngenbach et al., 2006; Velasquez,1991, cited in Peters, 2000) and in terms of singing accuracy (Edenfield and Hughes, 1991). If people with DS can indeed learn and develop in the musical domain, as others with intellectual disabilities can, then they may also be able to benefit physiologically in terms of improved auditory perception for speech, and in terms of vocal and oro-motor production in speech, subject to individual physiological and cognitive constraints. For the DS population, this method may be more appealing than traditional intervention, given their enjoyment of music and singing activities.

Singing and speech overlap in terms of physiological processes and in some neural activities, in TD populations (Christina and Reiterer, 2013; Patel, 2014). Christina and Reiterer (2013) describe singing as a ‘hybrid’ of music and speech, and suggest that the additional demands that singing places on perception and vocal motor control make it most suited to supporting speech. Similarly, Patel (2014) argues that any transfer of learning from song to speech represents within-domain plasticity. Patel also maintains that cross-domain transfer is possible: that is, that music training that does not rely on singing may also result in changes in the brain’s plasticity in neural networks that are common to music and speech; and that these changes can transfer to speech, if a specific

set of five conditions are met. As defined in his original hypothesis (Patel, 2011), these conditions are that there is an overlap (O) in processing areas within the brain; that training places more precise demands on shared networks than speech alone (P); that the training evokes an emotional connection (E); that there is repetition (R); and that the learner focusses their attention upon the musical activities (A) (Patel, 2011). The hypothesis states that if these conditions are met, the shared neural networks involved are driven to a higher state of precision than is required for speech, with benefits for speech processing. Thus, hypothetically, precision in drumming may be enough to alter the brain's perception of temporal information and enhance the brain's ability to process syllables or consonants. His hypothesis has since been extended to consider the effects of the distance of transfer in the auditory domain (Miendlarewska and Trost, 2013), together with the level of processing required (Moreno and Bidelman, 2013) and its application to sensori-motor output (Fujii and Wan, 2014).

Hypothetically, then, musical or singing activities may support the reduced auditory processing abilities in people with DS; and singing in particular may support speech perception and aspects of speech and voice production. However, any change in vocal or speech perception or production depends in the first instance upon the ability of people with DS to learn to develop musical awareness or skills. Whilst research shows some musical ability and capacity to develop specific skills in singing (Edenfield and Hughes, 1991) and in rhythm (Lense and Dykens, 2011), their capacity to develop musical abilities have not been adequately researched. In TD children and adults, there is a wealth of environmental and biological factors impacting on development and learning of musical ability (Bispham, 2006; Corrigall and Schellenberg, 2016; Hargreaves, 1986; Reifinger, 2006; Lamont, 2009). People with DS are subject to the same influences, but have additional confounding factors that may affect their learning trajectory. It is also not known the extent to which musical ability in DS may be constrained by cognitive factors, or to what extent musical ability might constitute another aspect of general intelligence (Schellenberg, 2016).

This study aims to investigate the nature of speech and musical ability in four adults with DS and with SLD, with a view to establish the feasibility of using music to support voice and speech production. The DS phenotype is useful for informing and guiding the teaching of musical skills for the population. This pattern of relative strengths and weaknesses needs to be considered when designing singing and musical teaching for individuals with DS. These issues are discussed in Chapter 3, section 3.4.2. Interpretation of results must also consider the possibility of deviation from this profile as a result of unique genetic, medical, personal and experiential factors. For this reason, the current study will draw on individual case studies, and on group data. It will address the following hypotheses:

1. For adults with DS, musical development will be in line with their overall nonverbal cognitive development and vocal abilities;
2. Given the shared physiology between singing and speech (Rinta and Welch, 2008; Patel, 2014) and correlation between certain speech and musical abilities in terms of perception and production (e.g. Cumming et al., 2015; Kraus et al., 2015; Peter and Stoel-Gammon, 2006), difficulties that affect the speech domain will correspond to difficulties in the musical domain;
3. Adults with DS and SLD are capable of learning songs and of developing their melodic-rhythmic skills in singing. In line with other populations with developmental disabilities, individuals with DS can develop in both awareness and production of musical features when these skills are explicitly taught (MacDonald et al., 1999; Ockelford et al., 2011) and when teaching is tailored to strengths (Ringenbach et al., 2006); *and*
4. If the group is able to transfer learning between domains, and if the appropriate functional, emotional and cognitive conditions are met (Patel, 2011; Patel, 2014), then the development of musical skills may be employed to develop aspects of speech perception and production in adults and children with DS, of all abilities.

Data provided in the case studies (Chapters 4-7) will be used to support or refute hypotheses 1-3 as they relate to each individual in the study. Chapter 8 will respond to these hypotheses in the context of the group, and will address hypothesis 4.

This study therefore seeks to answer the following four questions:

1. What are the verbal and nonverbal abilities of the individuals in relation to each other and to the DS population as a whole?

The case studies (Chapters 4-7) will address the questions in relation to individuals. Chapter 8 will make comparisons between individuals in the study, and to the DS population.

2. What factors affect the individual's and group's musical and speech skills in terms of rhythm, melodic intonation and phonation?

The case studies (Chapters 4-7) will address this in relation to the four individuals. Chapter 8 will examine factors that are common to the group.

3. To what extent can the individuals and the group learn to improve their musical ability in terms of perception and production of musical features through song-based teaching?

The case studies (Chapters 4-7) will address the questions for individuals. Chapter 8 will make comparisons between individuals in the study, and to the DS populations.

4. For people with DS, what is the potential for rhythmic and song-based music tuition to develop aspects of auditory processing, voice and speech production?

The case studies will address this question for individuals. Chapter 8 will consider the implications of these findings in relation to the group and to the DS population.

Chapter 3: Methods

This study used a case-study approach to assess the impact of singing and music lessons on the speech and voice of four young adults with DS. The study was designed as a pilot study initially, the aim of which was to assess the suitability of an assessment battery and teaching programme for people with moderate-severe learning disabilities. Initially, it was intended to follow-up the pilot study with a larger-scale study in order to examine potential associations between musical and speech abilities, and the potential for individuals to learn and develop their singing skills. However, the pilot study revealed that the participants faced several difficulties in learning to reproduce speech and song that warranted further investigation. The pilot study also generated sufficient data to address the intended research questions. It is therefore the methodology of the pilot study that is given here.

Section 3.1 explains the study design and the procedures used for gaining ethical approval, and for recruiting and selecting participants. Section 3.2 explains the selection, design and delivery of the battery of assessments. Section 3.3 explains the procedures used when administering the tests and for recordings. The design and delivery of the song-based curriculum and teaching programme are explained in Section 3.4. Section 3.5 explains the methods and procedures of analysis used for cognitive and musical tasks. Section 3.6 explains the methods used for the analysis of speech sounds, voice, phonation and prosody. It also presents the results of the reliability tests. Section 3.7 provides the rationale for the selection of data for individuals in the case studies, together with a description of any additional methods that were used for individuals.

3.1 Study design and participants

3.1.1 Study design

The study was exploratory in nature. A multiple case study design was used, in combination with a small-scale longitudinal element, over a period of ten weeks. Over a four-week period, four participants with DS participated in a battery of cognitive, speech and musical assessments in order to profile their levels of cognitive and speech ability (Section 3.2.1), and to establish baselines for their musical and singing ability (Section 3.2.2). Both sets also provided data for assessment of voice in a range of tasks. The participants then engaged in four full sessions of group song-based activities and in one partial session, during which their progress in learning and performing songs was tracked (Section 3.2.3). In the final week, each participant was reassessed individually against measures of pitch- and interval-production (Section 3.2.3) in order to determine any changes in musical or vocal performance.

3.1.2 Recruitment and selection of participants

Following Ethics approval from the Human Communication Sciences' Ethics Committee, University of Sheffield (see Appendix 1), informed consent was sought and obtained from a residential care home for people aged 18-65 years with learning disabilities.

The Principal of the care home identified possible participants using the inclusion criteria, which were:

1. a diagnosis of DS;
2. aged between 11-25 years;
3. moderate-severe learning disability;
4. a recognised degree of speech impairment; and
5. an interest in singing.

Exclusion criteria were a diagnosis of profound and multiple learning disabilities, and a requirement for 1:1 support, as this would affect the teaching and learning process.

Staff within the organisation distributed letters to parents/carers and to potential participants, together with information sheets (see Appendix 1). The same staff sought consent from the participants after a period of two weeks. Four young people were identified and approached, and all four gave informed consent. The participants consisted of 2 females and 2 males with DS, all aged 23 years, with SLD. All participants were familiar with the researcher, who had worked with them in formal and informal music and singing classes over the course of their three years at an FE College. These sessions had ended one year prior to the study, when the participants left college.

3.2 The battery of assessments

3.2.1 Speech and cognitive tests

A battery of ten standardised tests (numbered 1-10, below) was used to enable comparison with other populations, and to estimate each participant's cognitive abilities and abilities in speech perception and speech production. Most of these speech, language and cognitive tests have been used with older people with learning disabilities, including young adults with DS: for example the British Vocabulary Picture Scale (BPVS) (Roch and Jarrold, 2008); Diadokinetic rate (DDK) (Cleland et al., 2010). This allows for direct comparison with a limited number of studies.

Six non-standardised speech tests (numbered 11-16, below) were included in order to generate data on connected speech, including spontaneous speech prosody and voice quality. Of these, tests 13 (Minimal Pairs test) and 14 (Simple Prosody test) were devised

specifically for the study. These were created in order to assess auditory processing of speech at a level below or beyond that tested in the standardised tests.

The test descriptions below explain the purpose of each assessment and the methods used for its administration.

Standardised tests

1. **British Picture Vocabulary Scale (BPVS**) (Dunn et al., 1992). The test measures receptive vocabulary, and correlates to verbal intelligence. It has been standardised for use for children and adults to 17;11 years, and has been successfully used with people with DS to estimate Mental Age, across a range of ability levels (Carr, 2005; Glenn and Cunningham, 2005). Participants were required to listen to a word and identify the corresponding picture, from a set of four.
2. **Goodenough-Harris Draw-a-Person test (DAP)**. The test has been used to estimate non-verbal cognitive development in people with learning disabilities. The outcome correlates to visual-motor development and intelligence (Dykens, 1996; Sourtiji, Hosseini, Soleimani and Hosseini, 2010). Participants were asked to draw from memory the picture of a man or woman, without assistance. No time limit was given.
3. **TAPS-R.** The task is used as an indicator of verbal short-term memory (standardised for participants of 4-18 years). Auditory memory is often reduced in people with DS, and this information may shed light on limitations in auditory tasks and to inform teaching practice. Participants listened to and repeated a sequence of numbers that were spoken at an even rate of approximately one digit per second, by the examiner.
4. **Test of Rhythm and Comprehension of Language (TRaCoL)** (Treharne, 1999). The assessment has been used to assess disorders in auditory processing, and reveals information relevant to both speech and musical processing. Two subtests were chosen to assess temporal perception and production: 1. listening to and copying clapped rhythms; 2. judging whether a tapped rhythm is the same or different to a rhythm played on piano (standardised 4;10-11;11 years). Task 1 was presented live: the researcher clapped the rhythm and asked each participants to copy. One repetition was given if the participant did not respond. Task 2 was modified, as below:

    Modification: Task 2 was modified in order to reduce the reliance on auditory working memory and to maintain consistency in administering the test. The Rhythm Matching Test 1 was presented as a Word 2007 Powerpoint slideshow (Appendix 2, p. 334) with pre-recorded sound samples, amplified through an *Amethyst* iPig speaker adjusted to a volume level that was comfortable for the participant. Images of two different puppet characters

were presented alongside images of a piano and a drum in order to show the participants that there are two sounds of different timbres. The researcher explained that one puppet would be 'playing' the piano and that the second would be copying the pattern on the drum: the participant's task was to assess whether the second puppet correctly imitated the pattern. 'Same/different' cue cards were present on the slides. The rhythm samples were generated from existing piano and drum samples using Garageband 08 (Version 4.1.2: Apple Inc. 2002-2007), on an iMac.

The TRaCoL test also includes a prosody perception test but this was considered too difficult for the participants recruited: a simpler, non-standardised test was designed (see Test 14, below).

5. **Diagnostic Assessment of Articulation and Phonology (DEAP)** (Dodd, Zhu, Crosby, Holme and Ozanne, 2002). The Phonology Assessment only was used, which includes two subtests:1. Phonological Picture Naming; and 2. Picture Description. The test consists of fifty coloured pictures, which are designed to elicit all vowels and diphthongs, and all consonants in syllable initial and syllable final position. The Picture Description subtest allows individuals to use fourteen words in connected speech, for comparison. The test includes assessment and analysis of an individual's inventory, error patterns and intelligibility ratings. The images were shown to the participants, and they were asked to name each picture: prompts were given as necessary, and imitated words were identified on the record sheet.
6. **Children's Test of Nonword Repetition (CNREP)** (Gathercole and Baddeley, 1996). CNREP measures phonological memory, as opposed to digit span, and does not draw on lexical or semantic knowledge. The test was used to compare phonological memory with digit span, and also to generate speech and voice data. The words were presented live for imitation, with one repetition allowed.
7. **Auditory Discrimination Task 1**: Same/Different, S-Cluster Sequences, Words and Non-words – Shortest Form (from Bridgeman & Snowling, 1988, in Stackhouse, Vance, Pascoe, & Wells, 2007). A short list of real words and non-words was presented, and participants were asked to say if the pairs sounded the same or different. Two cards were used so that participants could choose a 'same' or 'different' image. Participants were familiarised with the task by using the cards to indicate whether their own names were the same or different to other names, and to make judgements between pairs of words: *dog dog*, *dog cat*, *pig dog*. The 'same/different' name test was also used intermittently to check attention during the task. Norms are available.
8. **Ss/zz test** (Shipley and McAfee, 2009). The task requires participants to sustain /s/ and /z/ sounds on one breath, for as long as they can. It is used as a measure of vocal cord functioning and can indicate vocal fold pathology; it also gives an

indication of breath capacity and management. The task was demonstrated and subjects were asked to copy. Norms are available.

9. **DDK test**. The task measures an individual's ability to make rapid oral motor movements (Shipley and McAfee, 2009). Following a demonstration, participants were asked to repeat the sounds /p/, /t/, /k/ in succession, as quickly as possible for a period of 10 seconds. Norms are available.
10. **Phrase Imitation**. The Connected Speech Task 2: Connected Speech Processes (CSP) Repetition (Newton,1999, as cited in Stackhouse et al., 2007) was used to assess the ability of individuals to repeat the prosodic and segmental content of phrases. The first ten phrases of the phrase imitation test were spoken aloud for participants to repeat. Sentences were repeated if the subject requested it, or if they had not responded after a substantial pause (c. 10 seconds).

Non-standardised speech/voice tests

11. **Describe a picture**. In order to assess spontaneous production of speech and voice, and syntactic ability, subjects were asked to talk about a picture of a farmyard, a picture of a picnic, and a picture of a fairground (Shipley and McAfee, 2009).
12. **Imitate words (Syllables test)**. The task assesses the ability to repeat words of increasing target syllable length (Syllables test: Shipley and McAfee, 2009). A set of nine words, each consisting of words of one, two and three syllables that commenced with the same syllable (e.g. *love*, *loving*, *lovingly*), were presented to each participant, one word at a time, in order of the number of syllables. Words were repeated if the subject requested it, or if they had not responded after a substantial pause (c. 10 seconds).
13. **Minimal pairs test**. A set of slides were created and presented using a Word 2007 Powerpoint. Each slide showed four images of words that differed by one phonetic contrast. A set of five slides presented words that differed in initial consonant position, and a further set of five were delivered that showed images representing contrasts in diphthong and vowel length (see Appendix 2, p. 328).
14. **Simple prosody test**. The test was designed to learn whether participants are able to choose between two words based on a) speech rhythm, and b) speech rhythm and melody (see Appendix 2, p. 330). The television characters Sooty, Sweep and Soo were used to create a forced choice task based on a 'shopping' scenario: participants had to help Soo decide which of two items each character wanted to eat, buy, or watch on television. The character Sooty answered with rhythmic patterns, and the character Sweep gave a melodic and rhythmic response. Sounds were generated for the rhythm test by a synthesised drum sound, chosen from the Garageband 08 'pop drum kit'. The voice of Sweep as created by the researcher recording the words and phrases using a /ŋ/ sound. All

samples were created using Garageband software, and soundfiles were imported into a Microsoft Word 2007 powerpoint.

15. **Reading**. Participants were asked to read a short passage ('Swimming': Shipley and McAfee, 2009). This allowed comparison of production in task that did not require auditory memory and for comparison to spontaneous speech production.
16. **Speaking the words to a song**. Participants were asked to speak the words to a song they knew well, such as *Happy Birthday* or *Twinkle Twinkle Little Star*. The task was included as an indication of speech and voice production when drawing on long-term memory and to allow comparison to the sung version.

3.2.2 Musical tests

The Primary Measures of Music Audition (PMMA: Gordon, 1979) test is designed to assess the musical perception skills of primary-grade students, and has been standardised for children aged 5-8 years. Whilst there is a version available for older students, it was felt that the Primary version would be more developmentally appropriate for the pilot test. Due to the limited availability of standardised music tests, eleven non-standardised tests (numbered 18-28, below) were designed to measure perception and production skills of early musical skills and singing ability. The design of the tests was informed by previous teaching experience, using the Kodaly method (see Section, 3.4.1), and by developmental data pertaining to musical development in children (see Chapter 2, Section 2.1). Imitation tests included aural presentation of stimuli, and aural stimuli coupled with a visual representation in order to compare performance when participants did not need to rely solely on auditory memory.

Given the unknown cognitive abilities of the DS participants it was possible that the PMMA test would present difficulties in task understanding. It was also anticipated that the demands on auditory memory inherent in PMMA (see test 17, below) and TRaCoL may pose a barrier to the subjects in completing the tasks, and therefore obscure their perceptual abilities or difficulties. As a result, perceptual tests of pitch direction and of rhythm and melody matching were created to complement the TRaCoL and PMMA tests, using a test format that incorporated visual stimuli and used simpler auditory stimuli (see tasks 26, 27 and 28, below, for details, and Appendix 2, pp. 332-336). The tests were designed to reduce the load on auditory working memory, as participants with poor working memory would be likely to perform poorly on the standardised tests, regardless of perceptual ability.

Standardised test

17. **Primary Measures of Music Audiation (PMMA)** (Gordon, 1979) is a computer-based assessment of musical perception based on the ability to discriminate rhythmic and melodic differences between two samples. The test consists of a training period, introduced by a dog character, and two subtests consisting of

forty tonal and forty rhythmic sequences, each consisting of 2-4 pure tones. Each sample is introduced verbally by 'first' and 'second', and the child selects an image representing 'same' or 'different' (this same image was used across all same-different tests). The test was not administered, as the group failed to pass the simpler 'same/different' test (27, below).

Non-standardised tests

18. **Gross Motor Score**. A checklist was used (Appendix 2, p. 324) in order to assess the abilities of individuals to perform a selection of limb-motor movements in time to the beat of a song. The same song was used for all participants (*Beat It*, by Michael Jackson), which has a median tempo (139 bpm, bpmdatabase) and was known to all participants. Judgements were made based on observation of their live performance. The checklist had been developed by the researcher for previous teaching sessions with groups of comparable abilities.
19. **Beat entrainment.** This task assessed the ability to synchronise to a beat at different tempi. The task is based on the model used by Peter and Stoel-Gammon (2008) who explored its use as an indicator of developmental apraxia of speech. This test used the same tempi of 104, 132, and 160 beats per minute (bpm). A simple metronome sequence was created on Garageband using a 'clave' sound from within the Garageband loop section. The loop was edited to place the 'strong' beat in first position, and the second beat of each sequence was copied and placed on beats 2, 3 and 4. The new loop was copied to create a metronome with a strongly metrical beat. Participants were asked to listen for four bars (16 beats), tap in time to the beat on the table for five bars (30 beats) and to maintain the beat when the sound stopped. Appendix 2 (p. 337) shows a screenshot of the test at 104 bpm. Research shows that TD adults tap 20-50 ms ahead of the beat (Repp, 2005)
20. **Beat and rhythm extraction.** Two tasks were used to assess how well each participant could show the beat in unaccompanied vocal music, and whether they could segment words and play the word rhythms.
    a) The song *Can you tap the beat?* (Brindle, 2005) was demonstrated to each participant. Each phrase was then demonstrated and the participant was requested to copy it: a physical gesture was used to symbolise *my turn*/*your turn,* and to cue the response. Participants were then asked to clap the rhythm of the words in each phrase, by listening to a phrase at a time.
    b) Participants were asked to say and clap the words *cat*, *caterpillar* and *monkey*. If they were unable to do this independently, a demonstration was given. They were then given a visual representation of the word, and of its rhythm using standard music notation with enlarged 'heads' (Appendix 2, p. 327). Participants were told that two dots represents two claps, and one dot represents one clap. The participant was then asked to play the word rhythm

as the assessor said the words and pointed to each symbol on the score. If they were successful in this, the images were combined to create a rhythmic pattern (*monkey*, *caterpillar*, *monkey, cat).* They were asked to clap or drum the pattern as the researcher pointed to the picture with a steady beat.

21. **Rhythm entrainment**. Two clave samples were selected from the Garageband loops. Rhythm 1, a syncopated sequence of 8 notes, was looped for 20 bars. Rhythm 2 was a simple rhythmic pattern consisting of half and quarter notes, with each note synchronised to the beat for 15 bars, followed by a variant of the rhythm for a further 10 bars. This was designed to test production and to ascertain whether, and when, they would hear and respond to the change in rhythm. Appendix 2 (p. 338) contains a screenshot of the test of Rhythm 2.
22. **Matching pitch**. Participants were instructed to listen to and copy pitches within the range, G3, C4, D4, G4 and C5. Each pitch was played on a keyboard and then modelled by the researcher at a pitch deemed within the participant's vocal range. Following aural presentation, the task was repeated with a visual representation to reinforce the pattern they heard (Appendix 2, p. 326): this was in order to determine whether visual representation of the sound affected performance and to prepare them for a perceptual test in the next assessment session.
23. **Imitate sung patterns**

a) **Glides**. Participants were asked to imitate an ascending and descending vocal glide on /la/, which aimed to establish their highest and lowest vocal pitch.

b) **Matching Intervals.** Three patterns of two-note descending intervals were modelled (*so-me, me-do, so-do*) and two patterns of three-notes (*me-re-do, so-la-so*). The majority of the intervals were based on descending patterns and on the universal intervals that appear first within children's development (5^{ths}, 4^{ths}, major 3^{rds}, minor 3^{rds}) as the literature suggests that these are learned and perceived more readily (Brindle, 2005; Deutsch, 2012). Patterns were played on chime bars and were then modelled for participants at a pitch that was deemed most suitable to their vocal range. Once intervals had been presented with visual representations, a visual representation of an ascending and descending scale was presented (Appendix 2, p. 326). Participants were asked to sing from the image; if they were unable, the ascending section was modelled.

24. **Sing a song from memory**. Participants were asked to sing *Happy Birthday* or *Twinkle Twinkle Little Star* to assess ability to recall and reproduce words and musical information. If it was apparent that they were struggling, verbal prompts were given as necessary for melody and words. Such prompts are indicated in the results of each case study (Chapters 4-7).
25. **Learn a new song**. A simplified version of the song *Monkey Man* (Hibbert,1969) was introduced in week 1 to allow comparison between each individual's performance with and without visual support, which was available in the form of Makaton

signs. The song was demonstrated and participants were instructed to listen, then to listen to and repeat one phrase at a time. Phrases were then presented with Makaton signs. The song was chosen for its lyrical and melodic simplicity and popularity with previous teaching groups, and for its ease of adaptation in terms of both Makaton and musical accompaniment. Although the song had recently been re-recorded commercially, none of the participants were familiar with it.

26. **Pitch perception test**. This task assessed whether the participants were able to associate pitch direction with a visual representation. The task required them to listen to a recorded vocal glide and point to the image of an arrow (pointing upward, downward or straight) that best represented it. The powerpoint slides for the test are included in Appendix 2 (p. 332).
27. **Same/different music perception tests**. This test was designed to assess skills fundamental to, but at a lower level, than the PMMA test. It used the same 'same/different' images and format, and a pre-test task used the same/different images with groups of shapes (Slides 1-4, Appendix 2, p. 333). In the rhythm subtest a) (Appendix 2, p. 334), two sounds of up to three notes were played on a piano and a drum, after an introductory 'triangle' tone. The participants were told that the drum was 'copying' the piano and to say whether it played the same pattern. The participant was asked to choose the symbol that showed whether the sounds were the same or different in rhythm. The images were used to show that the sounds were different in timbre, in order to encourage focus on the temporal aspects. The same/different symbols were used for the pitch subtest b) (Appendix 2, p. 335), but participants had to judge whether two pitch patterns were the same. The test consisted of synthesised notes of single sustained pitches, intervals of two notes, and pitch 'bends' of a semitone. The sounds were presented rapidly in succession, but a second 'triangle' sound intervened at a lower volume. In each test, each stimulus spanned approximately four seconds, including the introductory sound.
28. **'Find the sound' interval and rhythm perception test**. The task aimed to establish whether participants could match patterns that differed in either melody or in rhythm to a target sound, with reference to a visual image (Appendix 2, p. 336). A pre-recorded sung pattern of one or two notes was presented, and was represented by a 'target' image. The participant was instructed to find which of three characters sang the target. The three characters were spaced around the target, to the left, to the right, and below. The characters changed position between samples to avoid association with one character, or one position, being 'correct'. Participants listened to each character sing, and then to the target. The participants were allowed to play the target and test samples up to three times each before answering.

3.2.3 Within-session and repeated assessments

An additional pitch-matching task was included in the second teaching session, as the environmental conditions were most suitable for recording voice during this session. Pitch and interval matching skills (task 30) were re-assessed in week 10 in order to measure any changes in ability. Progress in song-learning was tracked (task 31) during group teaching sessions (weeks 5-8) and in the final individual assessment (week 10) in order to measure learning.

29. **Single pitches on vowels**: In the second week of taught sessions (week 6), three single pitches were presented across at least one octave (C3, D4, G4 and C4) on three different vowels (/a/, /i/, /u/). Participants were asked to listen and repeat, and to sustain the note for as long as they could. The target note was played on a chime bar and each vowel was modelled, one at a time, by the researcher at a pitch that was deemed most suitable to the participant's vocal range.
30. **Summative assessment**: In week 10, participants were reassessed against Task 22 (matching pitch) and Task 23b (matching intervals), in auditory form only.
31. P**rogress in song-learning**. This task measured each participant's short-term memory for musical information and words, and charted their learning process for new songs. The singing game *Magic Book* (Brindle, 2005: Appendix 3, p. 358) was taught in session 1 (week 5) and was repeated every session thereafter. The song was simple in terms of melody, rhythm and words, and included an appealing magic trick. It was explained to the participants that if the individual sang the song to their best ability, the 'magic book', when opened, included colour images; if it was felt that they could perform it better (e.g. with more words, or more accurate rhythm), then the images revealed were black and white. The researcher retained control of the book, but this method was used to enable students to assess their own and others' performance on specific aspects of the song: for example, by asking 'what could *(name)* do better?'; and to give immediate visual feedback to learners on their performance.

3.3 Recording and assessment procedures

The battery of 28 tests was designed to take no more than three hours, with time for breaks as and when needed. Each participant was scheduled for two blocks of assessment of approximately 90 minutes, bi-weekly over a period of four weeks. Assessments took place during the early afternoon in a small room, separated from a communal living area and dining room by a door and a window, with blinds. Assessments that were presented via an iMac, were amplified through an *Amethyst* iPig speaker, which was placed above the computer and adjusted to a volume level that was comfortable for the participant.

Audio recordings for speech and singing assessments were made using a shock-mounted cardioid condenser *RODE NT1A* microphone, set to 'unidirectional' mode. This was

connected to the iMac via an *M-Audio* Mobile-Pre USB soundcard, and recorded using Garageband 08 Version 4.1.2 (Apple inc. 2002-2007). A pop shield was placed 8 centimetres from the front surface of the microphone to encourage participants to maintain a constant distance when speaking and singing into the microphone. The microphone was set at a distance of 30 centimetres from the participant and mounted on a microphone stand slightly to the left of centre so as to avoid obstructing the participants' view of the computer and to allow them freedom of movement for singing and musical tasks. The recording levels were adjusted as necessary using the 'input' dial on the soundcard, to maintain a constant level.

In addition, an *H4* Zoom (Samson Technology, 2005) handheld digital recorder with two in-built unidirectional stereo condenser microphones was used to record spontaneous speech samples, and performance of songs during teaching sessions in order to monitor learning. The recorded format was a 24 bit WAV file, with a 44.1 KHz sampling frequency. Data was stored digitally on the Zoom's memory card, and transferred via USB to the iMac.

Data from recordings were imported into Praat version 5.2.17 (Boersma and Weenink, 2011) and edited into segments, which were saved as .wav files for further analysis and editing in both Praat (Boersma and Weenink, 2011*)* and Sonneta (Mint Leaf Software, 2015), according to the data required (see Sections 3.5, 3.6 and 3.7).

3.4 Lesson design

There is some evidence that students with Down Syndrome can learn to sing (Edenfield and Hughes, 1991), but there is no specific literature regarding how to teach this population singing or musical skills, and limited research concerning what developments they can make in this field. The lessons were therefore informed by singing and musical pedagogy in TD populations. This information was adjusted for people with learning disabilities (LD) based upon the researcher's previous experience with this population in teaching singing and in developing musical skills using a Kodaly (Houlahan and Tacka, 2008) approach.

3.4.1 Song-based musical development

Musical abilities develop through singing in the early stages (see Chapter 2: Section 2.3): at approximately 2;6 years, children are able to imitate melodic shape and simple rhythms. Phillips (1996) proposes that early vocal development should reflect this natural development. His recommendations are similar to those of the Kodaly approach (e.g Brindle, 2005; Houlahan and Tacka, 2008). Both approaches stress the development of ability to move and synchronise to the pulse, to clap and vocalise simple rhythms, and to learn to sing simple melodic patterns, and to develop inner hearing. The Kodaly philosophy (Brindle, 2005; Houlahan and Tacka, 2008) and Phillips's (1996) developmental curriculum formed the basis of teaching in this study. Both

approaches are at the appropriate developmental level for people with DS, and the combination of kinaesthetic, visual and aural learning styles make them accessible for people with learning disabilities. It was ensured that teaching songs and teaching activities were age-appropriate by including the use of age-appropriate resources (drums, chime bars), and songs with limited but thematically-appropriate lyrical content. Although songs with a 'nursery' or child-based theme were avoided in teaching, the songs *Happy Birthday* and *Twinkle Twinkle Little Star* were included in the battery of tests in order to measure individual's abilities to sing from long-term learning. These songs are often used in other research as measures of singing ability (e.g. Demorest and Pfordresher, 2015; Pfordresher and Brown, 2009; Welch, 2009; Welch et al., 2015).

3.4.2 Lesson design

Lesson content

The teaching curriculum aimed to develop the physical techniques associated with traditional singing pedagogy, and to encourage inner hearing and attention to melody and rhythm through the combination of singing, physical gesture and visual images (i.e. by using inner hearing, clapping words, showing pitch using simple notation and pitch visuals). Each session incorporated vocal technique exercises based on Phillips' Grade 1 curriculum (1996), but the exercises were adapted to meet the abilities of the participants: for example, speech and articulation drills were included by singing fragments of popular songs at increasing tempos (see Appendix 3). Songs and singing games were used to teach or reinforce melodic and rhythmic concepts (Houlahan and Tacka, 2008; Brindle, 2005). The objectives for the curriculum were as follows:

- **Develop techniques in preparing to sing.** Practise exercises in postural control, breathing, phonation, resonance, diction, expression;
- **Sing a 'line'**. Associate simple up/down/straight visual representation of pitch movements with the sound – e.g. upwards pitch glides with ascending straight lines, sustained tones with a straight line;
- **Sing a 'shape'**. Associate visual and gestural movements with melodic shapes and simple intervals – e.g. associate a wavy line with a sound that moves up and down repeatedly between two pitches;
- **Use motor movements to show the beat at different tempos.** Use motor movements whilst simultaneously saying, singing or listening to words - e.g by tapping knees, clapping hands, or playing an instrument in time;
- **Learn a new song and perform from memory.** Demonstrate retention of musical and lyrical content when singing a newly-learned song; *and*
- **Develop individual singing skills.** Developmental targets were set based on individual performances, as assessed at baseline. At the start of each activity that was pertinent to their learning need the 'target' was explained in simple terms (for example, to hum for three

counts). Immediate feedback was given on which aspects they were successful in; if further development was required, the target was repeated and revised for the next session (see notes, Appendix 3).

Skill development and assessment

Lessons were designed following initial assessment in order to match activities and content to the abilities of participants. Skills in singing and rhythmic movement were assessed against a checklist (Appendix 2, pp. 324-325). Individual targets for improvement were based on this initial assessment, and group activities were selected that developed these. All participants were expected to work towards the following aims by participating in the relevant activities, regardless of their ability recorded at baseline:

- Develop accuracy in singing *so-me* and *so-me-la* intervals;
- Develop ability to sing pitch glides across vocal range;
- Develop sustained phonation by sustained humming or by using kazoos;
- Practise controlled exhalation through the use of pinwheels;
- Practise inner audiation (ability to listen and think before repeating, in order to develop rhythmic and/or melodic memory);
- Show or maintain a steady beat of familiar songs at different tempos whilst vocalising words; *and*
- Learn the new songs *Magic Book* (Brindle, 2005) and *Monkey Man* (Hibbert, 1969): perform one or both songs from memory by attending to visual/auditory prompts as necessary.

Individual targets are discussed in the case studies (Chapters 4, 5, 6 and 7). Assessment of progress was formative and summative. Notes regarding progress were taken during teaching sessions (see Appendix 3) and these were used to evaluate progress and to adjust targets for the following session if necessary. These were supported by field recordings of teaching sessions. Summative assessment was conducted by assessing the individual's abilities to sing a newly-learned song (*Magic Book*: task 25); to tap or clap the beat and the rhythm of the words to *Magic Book*, whilst singing it; and to reproduce aurally single pitches and melodic intervals (task 22).

Lesson design

Six sessions were designed (Appendix 3, p. 340), each lasting 90 minutes, but unforeseen circumstances resulted in the cancellation of the fifth session after completion of group warmup exercises, and of the sixth session (Appendix 3, p.341). Although the content changed between sessions, each session typically, followed the following format:

- **Physical warm-up**. Imitate gross motor movements and creating movements to the beat (e.g. Movers and Shakers, Phillips, 1996);

- **Controlled exhalation, using pinwheels**. Participants were asked to control their breath to keep a small pinwheel spinning (Phillips, 1996);
- **Phonation**. Kazoos were used to encourage sustained phonation, before moving onto 'humming movement' exercises (Phillips, 1996);
- **Range development and visual development of pitch**. Participants were instructed to imitate upwards and downwards pitch movements on /u/ and /ɑ/ vowels. Before imitation, they were asked to look at a visual image (an arrow showing direction) of the pitch direction, then to listen to the demonstration and watch the modelled hand movements that showed the direction of pitch change (a hand movement showing upwards, downwards or sideways movement);
- **Pitch practice and inner audiation**. Pitched instruments were distributed (chime bars, and boomwhackers) to encourage individuals to explore different pitches, prior to listening and imitating single pitches. Pentatonic notes were used from the C major scale, and after imitating all notes, the song *Do Re Mi* (Rodgers & Hammerstein, 1959) was used to begin song-based work;
- **Articulation and tempo**. Fragments of familiar songs were performed at increasing tempos to encourage the ability to maintain a beat when singing, and to focus on articulation of consonants. Actual songs used depended on the preferences expressed by the participants;
- **Melodic awareness**. Vowel sounds were 'drawn' in the air, and were represented visually (e.g Appendix 2, p. 326). Similarly, lines from familiar pop songs of the participants' choice, and sections of newly-learned songs were 'drawn' in the air with gestures and students were asked to imitate or model these;
- **Pulse and rhythm practice**. A simple singing game, involving passing objects in time to the beat and/or playing a drum to the beat, was used to encourage development of beat-keeping, and as a means to teach a new song (*Rocky Mountain*, Brindle, 2005);
- **Song-learning and practice**. New songs *Monkey Man* (Hibbert, 1969) and *Magic Boo*k (Brindle, 2005) were taught by rote over a period of weeks, with a different focus each week. Initially, songs were represented 'whole' then taught a phrase at a time (Houlahan and Tacka, 2008). Makaton and representative actions were used to support word recall.
- BREAK: 15 minutes;
- **Solo performances**. Individuals chose a song to perform from the available song-collection and gave and received feedback. Participants were asked to comment on the quality of each other's performances; *and*
- **Final solo song.** *Magic book* (Brindle, 2005) was used as a plenary, and to track on-going development of musical skills. The song consists of four repeating phrases (Appendix 3, p. 358) and repeating melodic patterns, and accompanies a magic trick (see task 25, section 3.2.2).

3.5 Analysis of cognitive and musical performances

This section will outline the procedures used for analysis of cognitive ability, motor movements, rhythm and pitch, accuracy of speech production, voice production, and prosodic features. Task numbers are given that refer to the descriptions in section 3.2.

3.5.1 Verbal and nonverbal intelligence

BPVS (1) and Draw-a-person (2) tests were marked by the researcher in accordance with the guidance given in Dunn et al., (1992) and Goodenough (1963), respectively.

3.5.2 Motor movements

Gross motor movements (17) were assessed in real-time, by observation against a checklist during live performance (Appendix 2, p. 324). A rating scale was devised to indicate the observed accuracy of the movements relative to the beat: this allowed the generation of an impressionistic score of each participant's ability in different activities.

3.5.3 Rhythm and beat entrainment

Beat and rhythm abilities were evaluated using Praat (Boersma and Weenink, 2011: version 5.2.17). Recordings of beat entrainment tasks (19), beat and rhythm extraction (20) and rhythm entrainment (21) tasks were imported into Praat and the position of target beats were marked on a Text Grid (Figure 3.1). The audio recording, waveform and intensity peaks were used to determine

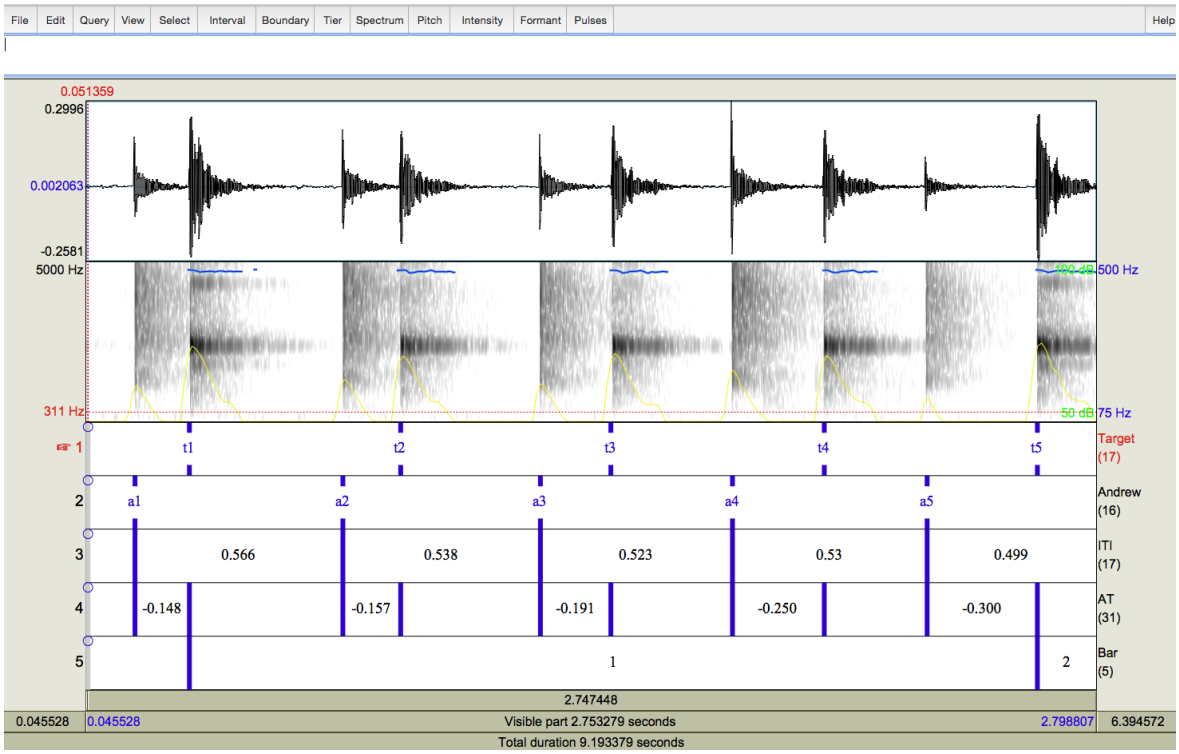


Figure 3.1: Screenshot of the method used in Praat for measuring timing accuracy in a beat entrainment task. Tier 1 shows the target beat marked and numbered and Tier 2 shows the beat the participants produced. The timing between successive beats played by the participant is given in Tier 3, and the difference in synchronisation (t1-a1) is shown in Tier 4. Tier 5 marks the bar numbers.

the onset of target beats. The participant's taps were identified in the same manner and were marked on the text grid. This enabled measurement of the timing between subsequent taps (*inter-tap interval*, ITI) and the delay between the participant's timing and the target (*anticipation time*, AT) (Figure 3.1; Tiers 3 and 4). The AT score was recorded as negative if it was produced in advance of the target tap and as positive if it followed the target tap. This allowed comparison to norms, in which most TD adults tap ahead of the target beat by 20-50ms (Repp, 2005).

3.5.4 Rhythm and melody in songs

Recordings of songs and rhythmic tests were imported from Garageband files into Praat for visual analysis. Field recordings from the *H4* Zoom were also used, as necessary: these were suitable for analysis of rhythmic timing, melodic shape and word production.

Rhythm

The rhythm of songs was measured by comparing the relative duration of successive syllables to each other, and to the target relative duration of syllables in the song. For example, in *Happy Birthday,* the first syllable of *happy* is worth 3/4 notes (Figure 3.2). Based on this it was possible to determine the approximate ratio of successive syllable duration for the target production and for song productions. For example, in Figure 3.3, the duration of the syllable *birth* (577 ms) was given a value of one: the duration of each syllable in *happy* was calculated as a percentage of this, and expressed as a ratio.

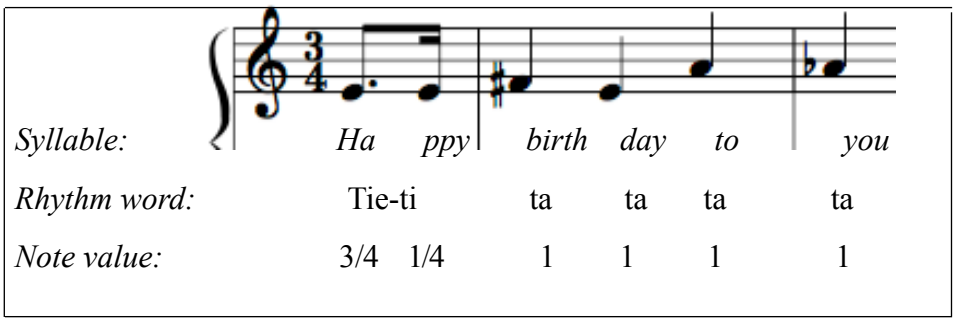


Figure 3.2: Notation for the opening phrase to 'Happy Birthday'. The rhythm names, based on Kodaly notation, and the rhythmic note values (generated in Garageband) are given. In this version, at three beats to a bar (3/4 time) each quaver (ta) is worth one count. The first word 'happy' is worth one count, but the first syllable is 1.5 times as long as the second syllable: the target production is therefore 3:1 (tie-ti) in this word.

Pitch and melodic contour

Melodic contours of intervals and songs were calculated by selecting the vowel in Praat, and measuring its mean pitch in semitones or Hz. Pitch settings were set to 'autocorrelation', pitch range was 50- 500 Hz for both generation of pitch in Hz and semitones (semitones re. 1 Hz). Pitch breaks were excluded from measurements (Figure 3.3). If the recording quality or vocal quality

resulted in an obvious miscalculation in Praat, the recorded vowel was played into a chromatic tuner in order to confirm the nearest note and the frequency. This procedure enabled the participant's melodic production to be measured and expressed as changes in semitones, or by plotting the mean F0 data into the Pages' chart generator, in order to produce a line graph. Examples of both methods are given below.

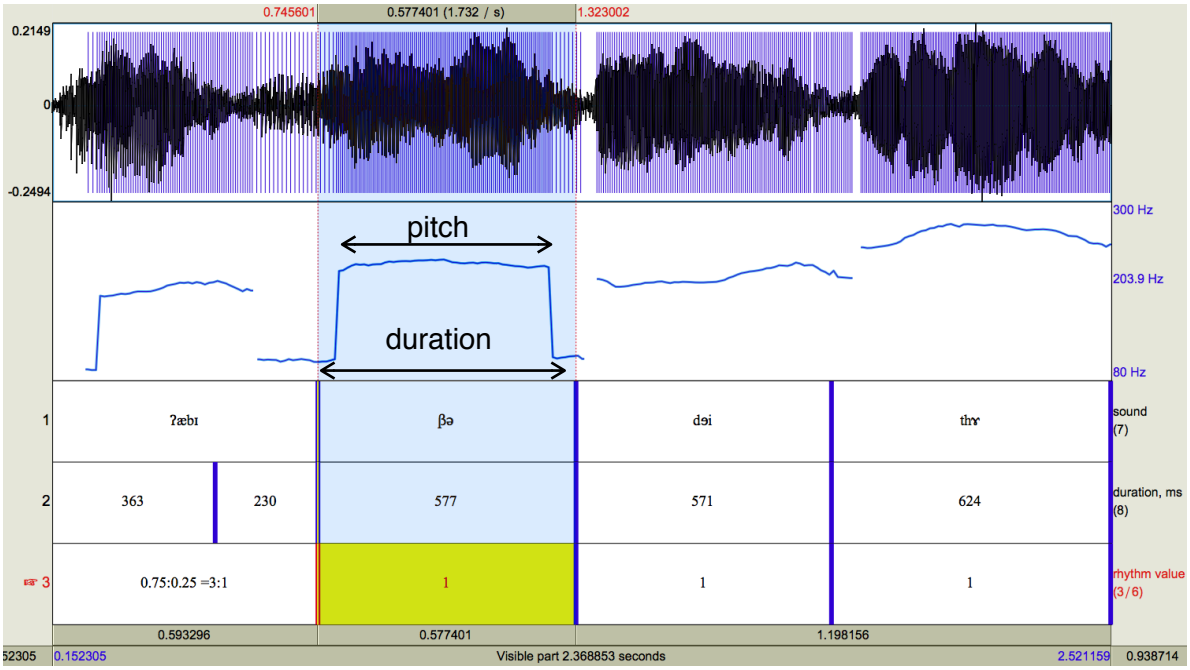


Figure 3.3: Method of calculation of duration and pitch in 'Happy Birthday'. The duration of the whole syllable (Tier 1) was measured (Tier 2, milliseconds), and was ascribed a rhythm value (Tier 3), according to the value of one note in the song. The mean pitch of the syllable was measured by highlighting the vowel, excluding sections of pitch break, and using Praat's 'get pitch' command. The mean pitch was then entered into a graph generator, in Pages.

Melodic contours in hertz

In order to compare the accuracy of songs to the intended melodic production, the participant's mean F0 was compared to the target F0. This method was useful when the recording quality contained too much noise for the changes to be illustrated in Praat. The participant's starting note was used as the target starting note, unless there was evidence that this was an error (see example Figure 3.4, below). The intended F0 of the next note could be calculated, with reference to the music notation. For example, in *Happy Birthday* (Figure 3.2), the second note should be the same as the first (no interval) and the third note should rise by an interval of a second; for *Twinkle Twinkle Little Star,* the first two notes are the same pitch and the second two notes are a fifth above these. In order to convert the intended semitone change into F0, the pitch of the starting note was entered into an online pitch converter (de Pijper, n.d.), and the target change (in semitones) was entered: the calculator then generated the target pitch, in Hz, of the second note. The data was entered alongside the participant's productions in order to generate a pitch graph.

Example

Figure 3.4 illustrates the process used for one participant (Andrew, Chapter 4): this also plots the accuracy of his timing, relative to the target production. First, the duration of the intended opening note was measured and used as a reference point. The intended starting pitch (F3: 175 Hz) was based on Andrew's second note, which was the note demonstrated on a keyboard. His opening note was in key, but was produced in error. The fourth column in the table (Figure 3.4) shows the target frequency of the sung pitch (top row), the frequency of Andrew's production (second row) and the point in time that he should be expected to change pitch. The anticipated timing was based on the duration of Andrew's first syllable (460 ms): based on this, he ought to change pitch at 1.38 seconds; and so on.

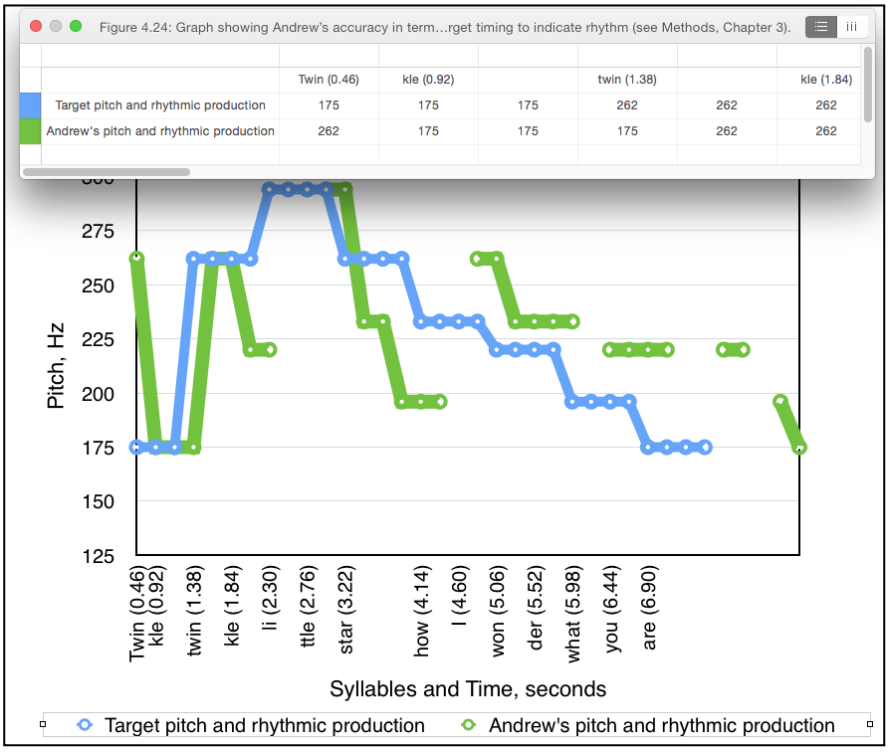


Figure 3.4: Method used in Pages to generate a graph of pitch and rhythm of a song produced by a participant ('Andrew', in green) against the target pitch and rhythm of a song (in blue).

Song or pitch imitation

Semitones were used in order to track pitch changes, regardless of absolute pitch. This allowed comparison between subjects and productions (e.g. of *Magic Book*). To measure pitch changes in semitones, the mean value of one note or vowel was subtracted from that of the previous note, giving either a positive (+) or negative (-) result that showed the direction of pitch change (Figure 3.5). In order to compare the mean pitch or pitch change that was produced by the participant to the target, a 'target' contour was generated. The pitch of the first note of the demonstration was measured in Hz or ST: the pitch of the next note was then calculated and entered into the graph;

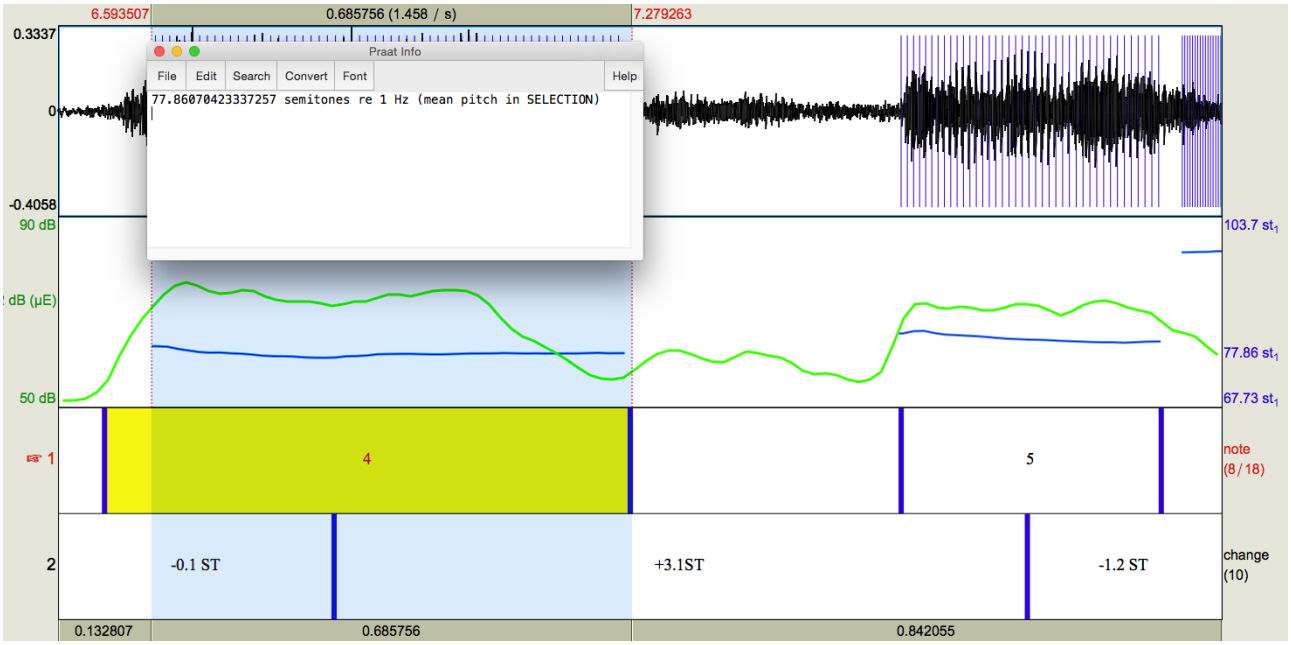


Figure 3.5: Screenshot showing the calculation of pitch changes in semitones, using Praat. The mean value of one note (Tier 1) was measured in semitones (re. 1 Hz), in Praat. The value of the next note was then subtracted and the difference was recorded (Tier 2) as positive (an increase) or negative (a decrease).

and so on. The mean pitch of the participant's syllables were then entered in the next row. The process differed for imitation tasks and production tasks, as explained above.

3.6 Procedures for analysis of speech data

3.6.1 Calculation of accuracy of speech sounds

Recordings were imported into Praat as .wav files and edited where appropriate to remove words that were contaminated with extraneous noise. A consistent broad-band setting was used for spectrograms. The settings were as follows: spectral frequency 0-5000 Hz; window length 5 ms; dynamic range 70 dB; automatic time-step strategy; intensity 50-100 dB; pitch 50-500 Hz. Text grids were used in order to demarcate segments and to allow annotation of speech sounds (Figure 3.6). Data from core speech tasks (5: DEAP; 6: CNREP; 10: picture description; 11: imitated words and phrases; 15: familiar songs) were transcribed phonetically, using IPA and EXT IPA symbols (Duckworth, Allen, Hardcastle and Ball, 1990) and VOQ symbols (Ball, Esling and Dickson, 1995). The transcriptions were informed by the audio recording, supported with information from the waveform shape, spectrogram, pitch contour and from the first and second formant contours. This enabled some sounds that were atypical or ambiguous to be more easily identified. However, some segments and words in continuous speech remained unintelligible or ambiguous. These were excluded from calculations of speech accuracy (below) but are included in transcriptions in the text, where relevant.

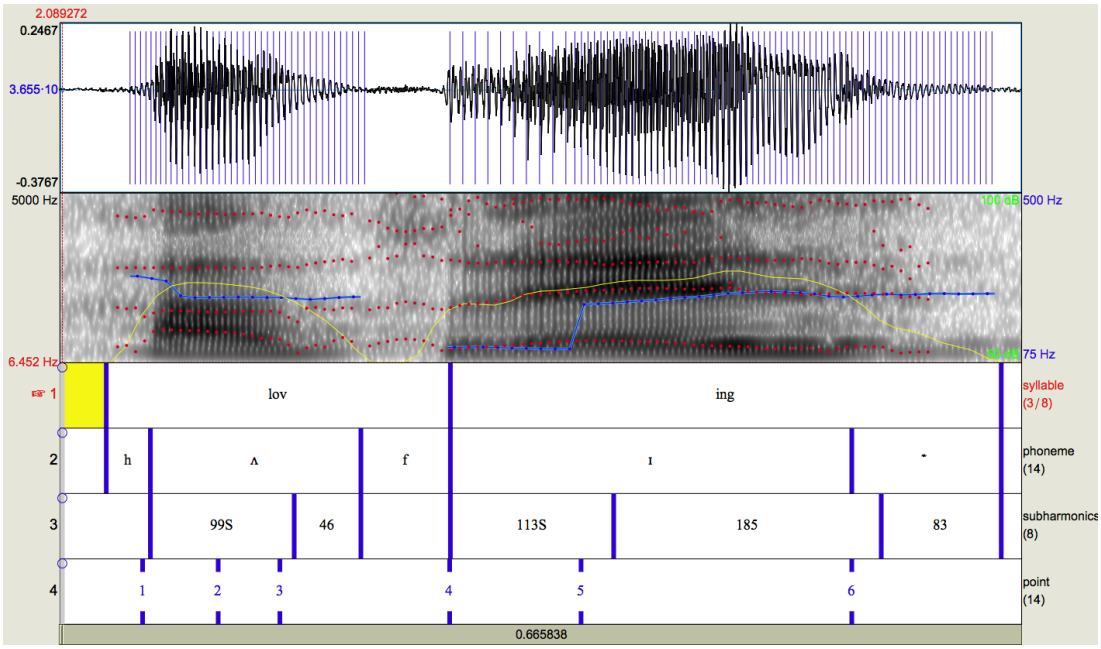


Figure 3.6: Screenshot showing the analysis of speech sounds in Praat, using broad-band spectrogram. The image shows the waveform (above) and spectrogram below, which is overlaid with the pitch contour (blue), formant contours (red) and intensity contour (yellow). Visual information from the waveform and the pitch contour (in blue) and formant frequency contours (in red) were used to isolate phonemes. The tiers below are annotated with the syllable, phonemes, duration of subharmonics (ms) and points of interest.

For each word that had a known referent (naming and imitation tasks), the number of sounds produced and the percentage of correct sounds were calculated. The sounds produced for each word were documented in an Excel spreadsheet, and the target sounds were written alongside these (Figure 3.7). Scores were tallied for each test. Within each test, words were grouped according to target syllable length of words and scores were also calculated for each group. The determination of sound errors followed the guidance in the DEAP test (Dunn et al., 1992). All single sounds, including diphthongs, were marked as either correct or incorrect. In Figure 3.7, speech

			Segmental information			**Dysfluency**
Percentage phoneme error=%	item no.	word	*speech sounds: phonetic description*			*sounds produced in dysfluency*
50.00%	3	train	tj	ɛĩ		t tʃ
66.66%	4	swing	tw	ĩ	n̥	h ts'w
25.00%	5	bread	bɹ	ɛ̰	t	
33.33%	6	duck		ʌ̰	kz	f k
0.00%	8	five	f	aĩ	v	ʔf f
33.33%	9	teeth	t	i:	uh	
0.00%	10	watch	w	ɔ̰	tʃ	

Figure 3.7: Method used for the calculation of speech errors in words of one syllable (DEAP). The image shows a screenshot from an Excel file. Boxes that are highlighted in pink were used to identify and tally errors in consonants and vowels. Boxes that are shaded in yellow identify errors of production (nasalised vowels, creaky vowels), but these were not included as errors. The image also shows the method used to calculate the degree of dysfluency that was present in tests.

errors are highlighted in pink (Figure 3.7), and impressionistic errors of vowel production (such as nasal production) are highlighted in yellow. As a result of coarticulation in speech, in which neighbouring sounds influence production, nasalised vowel production can reflect phonetic processes. These were not marked as errors due to this potential conflict. Furthermore, nasal production is common in DS, and the inclusion of such errors would raise the results to ceiling level.

The number of correctly produced sounds was compared to the target number of sounds to generate a Percentage Phoneme Error (PPE) for each word and mean Percentage Consonant Error (PCE) and Percentage Vowel Error (PVE) scores for each test and sub-test. The published guidelines were followed for DEAP (Dodd et al., 2002) in order to calculate the percentage of developmental and unusual errors and error patterns (where the error occurred in five or more instances). Where it was likely that a recurring error could be attributed to a habitual or learned production (e.g. substitution of the approximant /w/ for the alveolar approximant /ɹ/), these were removed to generate a revised calculation of consonant error (PCE-R) (Shriberg, Austin, Lewis, McSweeny, Wilson and David, 1997).

Phrases were treated in the same way: the target number of sounds was generated from the total number of sounds required to produce the entire phrase. If a word had been omitted, all missing target sounds were recorded as errors. The percentage of omitted words (PEO) and sounds was calculated separately for each phrase, and for the test as a whole.

The number of words affected by initial dysfluency were tallied and recorded as a percentage of the words produced in that test.

Use of SPSS for segmental data

The measures of PPE, PCE-R, PVE and PEO were copied into SPSS v.21. The target word, test name and number were entered. The data were coded in SPSS to indicate the following values, in order to generate mean data by test and by different parameters:

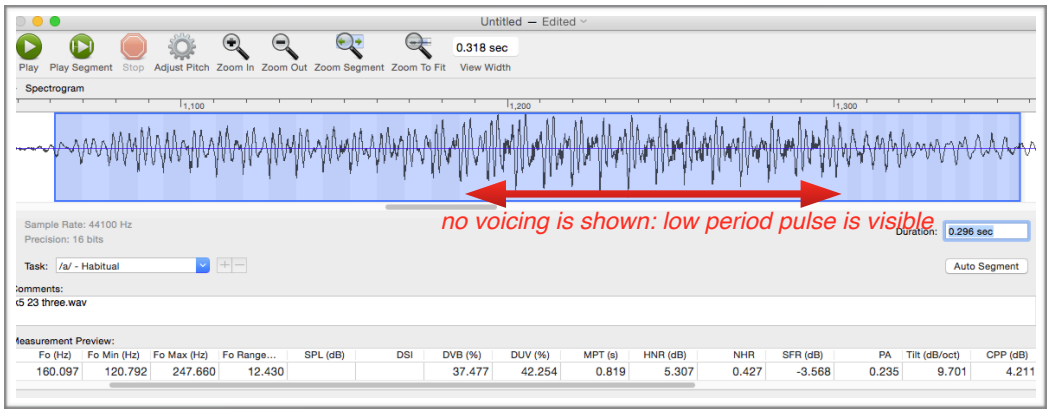
- the target number of syllables in the word;
- the category of target number syllables: one, two, three or more;
- the number of target words in the phrase (if relevant);
- the number of syllables produced; *and*
- the number of words produced.

3.6.2 Phonation and voice quality

Use of Praat and Sonneta

Initially, Praat was used for acoustic analysis of speech samples. However, comparison between measures for the same sample in Praat and in Sonneta showed a discrepancy in measuring the presence of voicing (Figure 3.8). Data from Sonneta were used for voice measures of all samples as it more accurately reflected the perceived changes and allowed generation of voice measures in speech, as well as for voiced vowels only. Praat was used for measures of intensity, however, as this feature was not available in Sonneta. Praat was also used for detailed examination of segmental and prosodic features and waveforms, and for the generation of Figures.

a)



b)

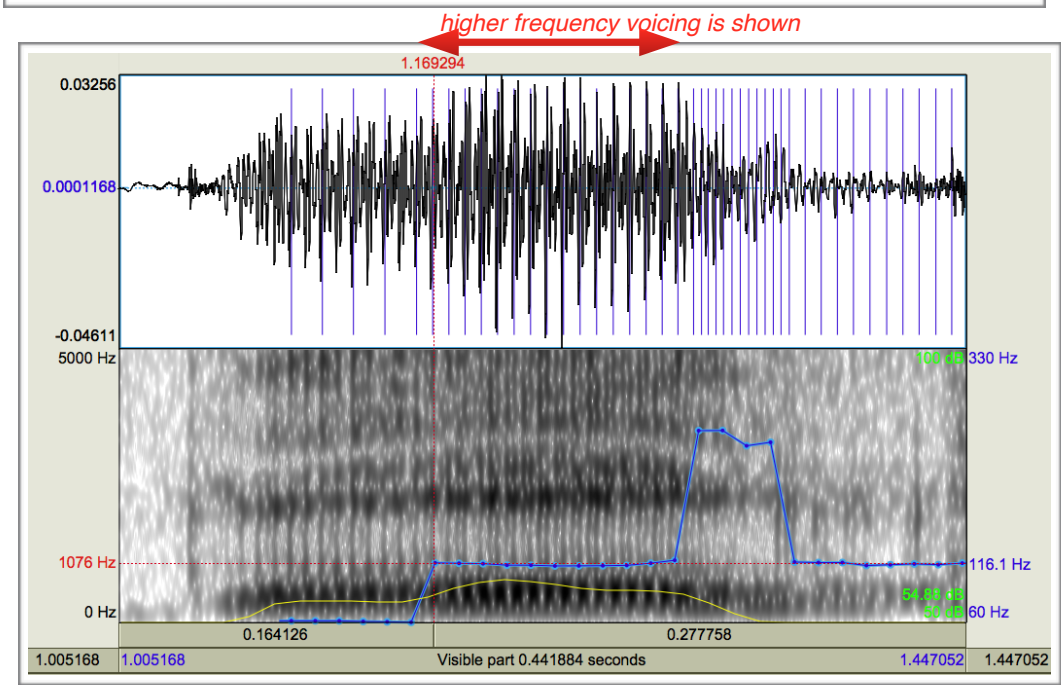


Figure 3.8: Screenshots of the word 'three' in Praat and Sonneta. The image shows a screenshot of how Sonneta (a) above) and Praat (b) below) present the voice data for the same recorded sample of Kerry's (Chapter 5) word 'three' at the same pitch settings. The programs differ in their measurement of fundamental frequency values and in the degree of unvoiced sections. Praat and Sonneta both report mean values of about 120 Hz, but Praat shows a rise in pitch after 0.164126 seconds, whereas a drop in pitch is perceived at this point. The red arrows highlight this section on the Praat and Sonneta image: Sonneta recognises this as an unvoiced section. The Sonneta measurements more accurately reflect the perceived changes to voice.

Selection and preparation of recordings

Voice perturbation measures and measures of fundamental frequency were calculated in Sonneta for recordings that were uncontaminated by background noise. Recordings for individual subjects were made at different times and days and so the available selection of recordings that were suitable for automatic voice analysis was different for individuals. However, for all participants, recordings from the sustained vowels (task 22b) and syllables tasks (task 11) were used to generate voice measures. Recordings of words within each test were imported as .wav files into Sonneta. Within these recordings, individual words were isolated, and examined aurally and visually. Contaminated portions were removed from analysis by excluding portions from the selection of the segment (e.g. the onset of sustained vowels), or by excluding the syllable or word in whole words.

Choice of voice perturbation measures

Measures of F0, standard deviation of F0, HNR, DUV, DVB, jitter (%), shimmer (%) were generated by Sonneta. These measures are commonly reported in voice research (e.g. Teixeira and Fernandes, 2015) and in voice research in DS literature (Albertini et al., 2010; Moura et al., 2008; Lee et al., 2009). These measures rely on periodic phonation, in which there is minimal variation between how the waveform repeats, but unhealthy voices may demonstrate atypical degrees of change from cycle to cycle (Baken and Orlikoff, 2000; Childers and Lee, 1991; Teixeira and Fernandes, 2015). Unhealthy voices may also be produced with additional frequencies (interharmonics) or with other irregularities that create 'noise' in the spectrum that make it impossible to measure features that depend upon a periodic signal (Baken and Orlikoff, 2000; Maryn, Roy, de Bodt, van Cauwenberge and Corthals, 2009; Fraile and Godino-Llorente, 2014). Given the high incidence of atypical phonation in DS (Kent and Vorperian, 2013), CPP values were also used in this study. CPP measures how prominent the harmonics are in the vocal source (Maryn et al., 2009) and the measure is appropriate for aperiodic voices which do not show 'well-defined' harmonic spectra (Maryn et al., 2009). CPP is shown to be a reliable indicator of both breathy and rough vocal qualities in vowels, of irregularities in periodicity of phonation (Dejonckere 1998) and of dysphonia in continuous speech (Heman-Ackah, Heuer, Michael, Ostrowski, Hoorman, Baroody et al., 2003; Maryn et al., 2009; Garrett, 2013). It is considered more reliable than jitter and shimmer for voices that are dysphonic (Fraile and Godino-Llorente, 2014; Heman-Ackah et al., 2003; Maryn et al., 2009).

The definitions of the measures that were used are as follows, and are derived from definitions given by Baken and Orlikoff (2000), Aronson and Bless (2009), MintLeaf Software (2015) and Teixeira and Fernandes (2015):

- **Fundamental frequency** (F0): a measure of the 'repetition rate of a recurring waveshape' (Baken and Orlikoff, 2000). The measure correlates closely to the perceived attribute of 'pitch';
- **Jitter (%)**: a measure of irregularities in the frequency of the vibration of the vocal folds. Relative jitter is the average difference between successive periods, divided by the duration of the average period (see Figure 3.9);
- **Shimmer (%)**: a measure of the average short-term variations in the peak-to-peak amplitude of the consecutive voicing cycles, expressed as a percentage of the average amplitude (see Figure 3.9);
- **Harmonic to Noise Ratio (HNR)**: a measurement of the proportion of periodic signal generated by the vocal cords in comparison to any aperiodic signal, such as that generated by turbulence;
- **Degree of Unvoiced Frames (DUV, %)**: the percentage of the frame that contains no periodic voicing;
- **Degree of Voice breaks (DVB, %)**: the proportion of a signal where there is a pause between glottal pulses that is atypically long for the fundamental frequency;
- **Cepstral Peak Prominence (CPP, dB)**: based on a logarithmic scale, CPP measures the prominence of the harmonic signal against background noise.

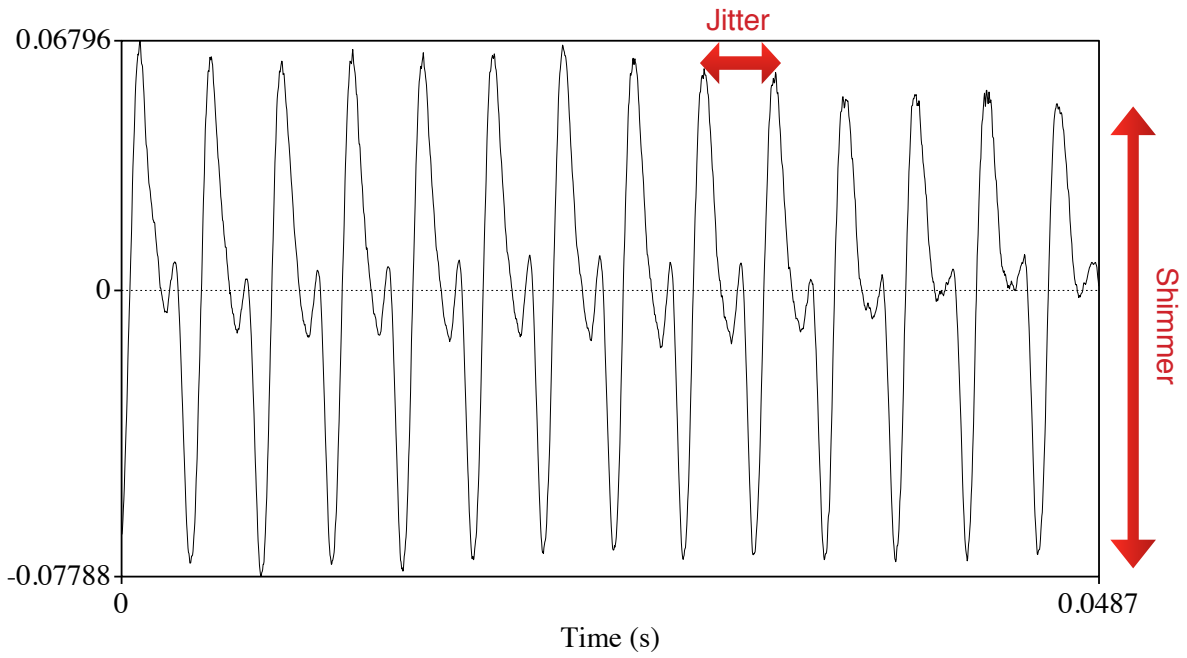


Figure 3.9: Jitter and Shimmer. The arrows indicate the dimensions of the waveform that are measured to calculate differences between successive cycles in period (jitter) and amplitude (shimmer).

Settings: Settings were predetermined by Sonneta. The program uses an automatic waveform-matching program that is considered reliable for recordings that contain some background noise (Goy, Fernandes, Pichora-Fuller and van Lieshout, 2013). The program allows pitch to be

automatically adjusted to capture different ranges, but the function was not used in order to ensure consistency between samples. Voice measurements for the vowels in each test were exported from Sonneta, using its 'export' function. This generates a text file of all available data for each vowel, along with identifying notes. The standard settings were used for all measurements and for the report template to ensure consistency of measurements and data between subjects and tests. The reports were opened as Excel files and the following data from the reports were exported from Excel into SPSS v 21.

Selection of voiced segments in sustained vowels

Most voice research uses medial portions of sustained vowels only (Fourcin and Abberton, 2008), as a result of the rapid changes in fundamental frequency and amplitude that occur at occur at onset and offset in healthy voices (Titze, 1994). However, the time taken for voices to stabilise can approach nearly 100 ms in most disordered voices and in a minority of healthy voices (Schaeffler, Beck and Jannetts, 2015). Furthermore, the precise selection of vowel segments can affect results (Shu, Jiang and Willey, 2016). Therefore, data were generated separately for the whole sustained vowel, for the initial 100 ms (onset) and for the final 100 ms (offset), and for the medial portion between the onset and offset (T-200 ms). This was calculated in Sonneta by first selecting the full vowel with reference to a broad-band spectrogram. The onset was determined by the point at which the pitch contour began (see Figure 3.10a): offset was determined by the end of the pitch contour or the end of a stable first formant contour, if phonation was absent. The same visual cues

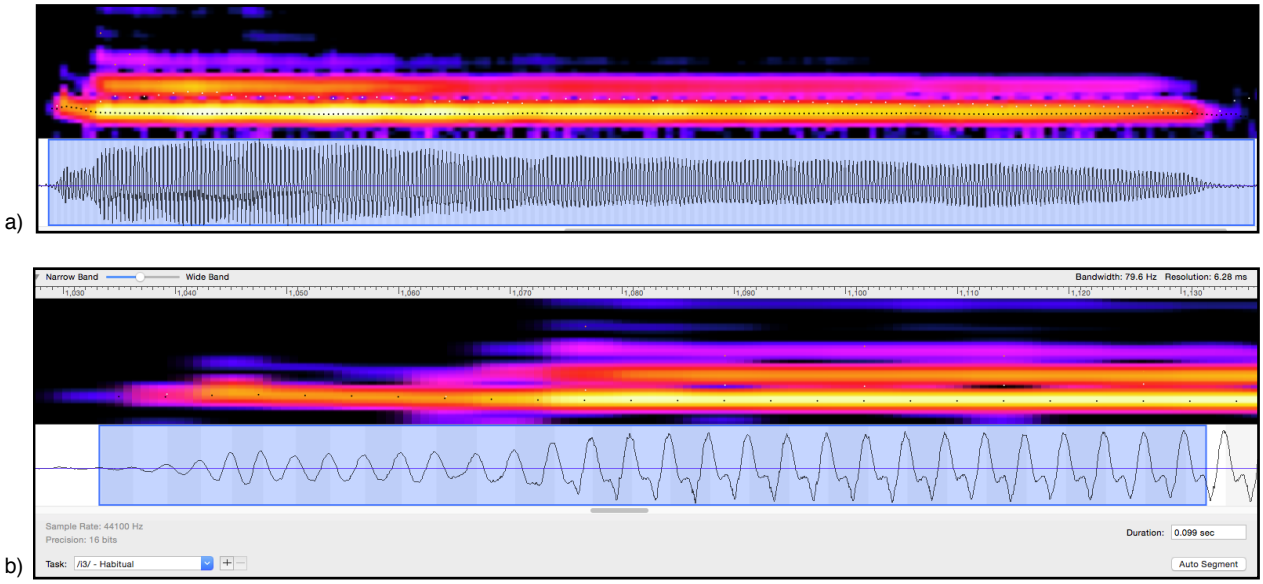


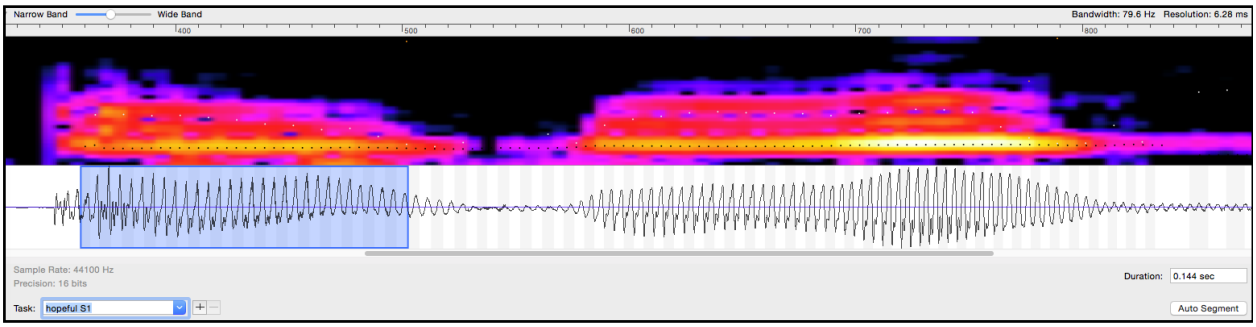
Figure 3.10: Screenshots showing the method used for voice measurement of sustained vowels in Sonneta. The images show the waveform and broad-band spectrogram of an /i/ vowel (0-1500 Hz; 35 dB), with voicing indicated on the waveform by the darker shaded vertical lines. Image a) shows a coloured spectrum in order to illustrate the presence of the first formant contour (a dotted white line) and the pitch contour (a dotted black line) that were used to inform the selection of onset and offset of the vowel, in combination with the pulse lines on the waveform. The blue box on the waveform shows the section of the vowel highlighted for measurement. Image b) shows a detailed view of the onset of the same vowel.

were used to select and generate data for the first 100 ms (Figure 3.10b); then the final 100 ms, and the duration between these.

Selection of voiced segments in words

For the generation of voice data in words, the vowel was selected visually by highlighting the section between the first and last waveform after any initial consonant, as indicated by the pale grey voicing lines on the waveform (Figure 3.11). The vowel was named in Sonneta as the word (e.g. *love*). The same selection process was used to measure each vowel in the multi-syllable words. These were coded with data that showed the word and the position of the syllable (e.g. Figure 3.11 a: *hopeful*, S1).

a)



b)

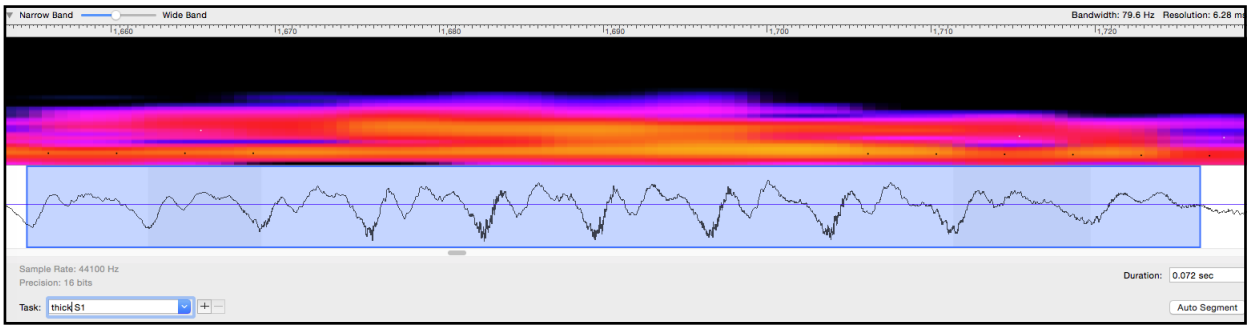


Figure 3.11: Screenshots showing the method used for voice measurement of voiced and devoiced vowels in words, in Sonneta. The images show the spectrogram and waveform of the word 'hopeful' (a) and 'thick' (b). The blue box shows the section of the vowel that was highlighted for measurement. In the voiced word (a), the onset and offset of the vowel were selected with reference to the audio recording, and the movement of the first formant contour (a white dotted line on the spectrogram), and the pitch contour (a black dotted line on the spectrogram). In the devoiced vowels, which were produced without periodic phonation (b), the shape of the waveform was used as a guide to inform the onset and offset of the vowel, in combination with the recording and formant and pitch contours, if available.

Calculation of interharmonics

In Praat, narrow-band spectrograms were used in order to determine the presence of bifurcations, such as **interharmonics**, which are noted in voices that are perceptually creaky, harsh or diplophonic (Cavalli and Hirson, 1999; Keating, Garellek and Kreiman, 2015). This study used the term interharmonics to describe the presence harmonics between main harmonics; it used the term **subharmonic** to describe a subset of interharmonics in which there is a well-defined single

harmonic. Often, a single subharmonic coincided with the acoustic phenomena, diplophonia, in which two tones are audible. This method allows for an objective presentation of the perceived voice quality, especially where voicing is atypical (Fuks, 1999). It enabled detailed analysis of the harmonics when it was not possible to generate acoustic measures as a result of aperiodicity in vowels. In combination with the recording, it was used to determine the presence of diplophonia and ventricular phonation, which may be present in the voices of people with Down Syndrome (Beckman et al., 1983; Rodger, 2009). For sustained vowels, measurements were also made of the duration of interharmonics in the initial and final 100 ms of phonation, in addition to the duration in medial portions (Figure 3.12). For vowels in words, the duration of phonation that was produced with visible interharmonics was measured and this was calculated as a percentage of the total duration of voice components in the word and test (Figure 3.13).

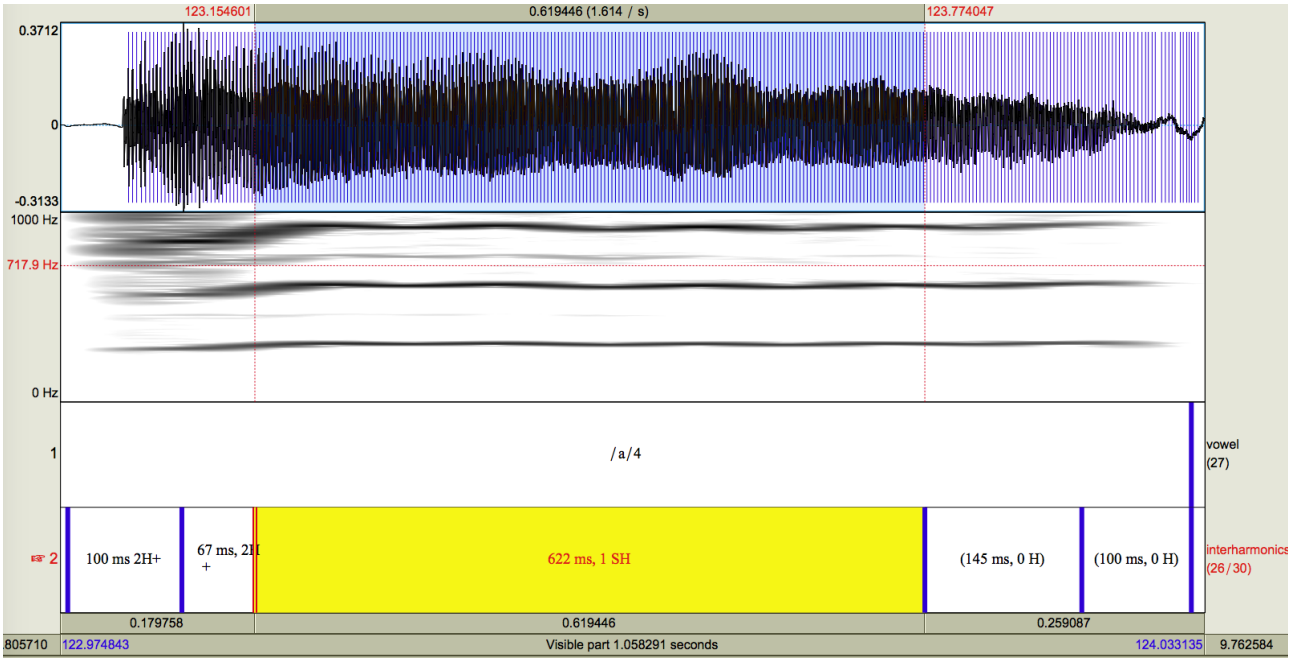


Figure 3.12: Screenshot of Praat showing the calculation of interharmonics present in sustained vowels. The spectrogram was set to a window length of 150 ms, frequency range is 0-1500 Hz, dynamic range is 35 dB. Tier 2 shows the duration of onset and offset, and sections that have clearly delineated interharmonics (2H+) or subharmonics (1SH) are annotated with the duration and magnitude of interharmonics (0H, 1 SH, or 2H+). In the above example, multiple interharmonics are visible during onset and during the subsequent 67 ms. In addition, a single subharmonic is visible between the 2nd and 3rd harmonic, above the red boundary marker at 717.9 Hz, and this continues for a further 622 ms.

Settings: A consistent setting for spectrograms was used to calculate the percentage of interharmonics for all participants, using narrow-band spectrograms. Settings were based on Cavalli and Hirson (1999): window length was set at 150 ms, the frequency range was 0-1000 Hz, and the dynamic range was adjusted to 35 dB to screen out the effects of background noise.

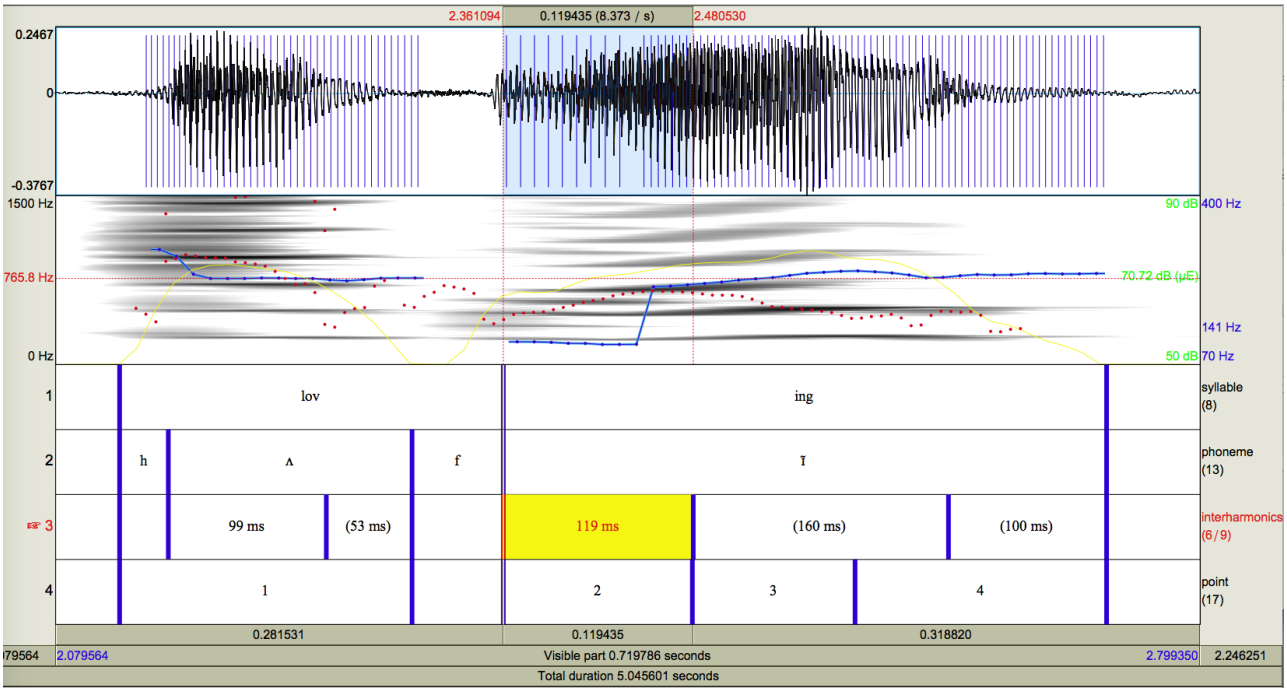


Figure 3.13: Screenshot of Praat showing the calculation of interharmonics present in words. A narrow-band spectrogram was used to reveal interharmonics in words (window length of 150 ms, frequency range is 0-1500 Hz, dynamic range is 35 dB). The duration of interharmonics in vowels are marked in Tier 3 (e.g. highlighted section) in milliseconds; the duration of vowels (ms) that did not have clearly defined interharmonics is given in brackets. The percentage of interharmonics was calculated as a percentage of the total vowel duration in each syllable.

3.6.3 Calculation and representation of prosody

Prosogram

Prosogram (Mertens, 2004) was used for two case studies (Rachel, Chapter 6; and Robert, Chapter 7). This Praat script allowed automatic generation of key prosodic data from recordings of suitable quality. A high degree of voice perturbations in the voice samples of Andrew (Chapter 4) and Kerry (Chapter 5) interfered with the accuracy of Prosogram and alternative means of calculation were used, as described below. The measures examined and their relevance to prosody are given in Table 3.1. Although these measures are typically used to allow comparison of a large corpus of speech to norms (Mertens, 2004), the measures were not used for this purpose in this study. Instead, they were used to assess the ability of individuals to reproduce prosodic features in comparison to known stimuli and for comparison between tasks and repetitions. In imitation tasks, comparison was made between the demonstration or target phrase and the individual's production. For the Syllables task and phrase imitation tasks (Newton, 1999), recordings of the stimulus and the imitation were separated in Praat and saved as new files. For the Syllables tasks, the files were split into separate files according to the target syllable length. The script was then run for the demonstration of words of one, two, and three syllables; and for the imitation of these. As the stimuli had been presented live, this procedure was followed for each participant in order to allow for differences between presentations.

Table 3.1: Description of prosody measures used in Prosogram. The table provides the definitions of the measures (adapted from Mertens, 2004) used to examine prosody using the Prosogram script. An explanation is given to demonstrate how these allow comparison to the demonstration in imitation tasks.

Measure	Definition	Application
nPVI-V	A measure of the variability between successive vowel durations in speech (Grabe and Low, 2002).	The measure allows comparison between the relative variation of vowels as spoken by the participants and the researcher. This can be taken as an indicator of the ability to reproduce syllabic rhythm.
TrajIntra	Pitch trajectory (sum of absolute intervals) within syllabic nuclei, divided by duration (in ST/s).	This enables comparison with the stimulus of mean pitch movement within the vowels.
TrajInter	Pitch trajectory (sum of absolute intervals) between syllabic nuclei (except pauses or speaker turns), divided by duration (in ST/s).	This enables comparison with the stimulus of mean pitch movement between the vowels and indicates the extent to which successive vowels within words are produced with intonational pitch changes.
PitchRange	Estimated pitch range (in ST) (2%-98% percentiles of data in nuclei without discontinuities).	Pitch range in semitones allows direct comparison of the individuals ability to imitate melodic changes, regardless of the fundamental frequency produced.
Gliss	Proportion (%) of syllables with large pitch movement (abs distance) >= 4ST).	These measure allow comparison of the trajectory and magnitude of pitch changes within nuclei.
Rises	Proportion (%) of syllables with pitch rise (>= 4ST).	
Falls	Proportion (%) of syllables with pitch fall (<= -4ST).	

The program is capable of measuring the relative duration of successive syllables, but the segmental and prosodic features of the group made determination of syllable boundaries unreliable. Instead, only vowels were used, as these were reliably identified by visual inspection of the waveform and spectrogram. The recommended procedure for Prosogram was followed: in Praat, text grids were used to demarcate and annotate the vowels in syllables in Tier 1 of the text grid; for the nuclei in vowels, automatic segmentation was run; and a program and prosodic profile was generated. The Prosogram was checked visually to ensure that phonatory difficulties had not affected the process. If they had, these samples were excluded from the data by removing the annotation for the vowel in the text grid, and the test was then repeated.

Settings: The default settings were used to run the script: these include the full time range and autodetection of F0, which identifies the median and range (median-12ST; median +18 ST); and a time frame parameter of 5 ms. The segmentation parameter was set to ‘nuclei in vowels in tier “phon”’.

Manual calculations of prosodic features

For most samples of speech produced by Andrew (Chapter 4) and Kerry (Chapter 5), phonation was intermittent. This resulted in Prosogram misreading a single nucleus as one or more, which generated false readings of nPVI and pitch changes. In this instance, key rhythmic, melodic and articulatory features were calculated manually, using text grids to measure key features in Praat and using Sonneta's 'continuous speech' measurements. The manual method was used for all participants for the sung and spoken productions of *Happy Birthday* and to track the development of the song, *Magic Book*. The recordings of these were from teaching sessions and the quality was not suitable for automatic processing. In song production, the target stimulus was calculated from the published musical notation (e.g. see Figure 3.3). For *Magic Book,* the duration of syllables was measured and pitch changes between subsequent syllables were also measured. This enabled calculation of rhythm, melody and speaking rate (syllables per second) (see Section 3.5.4).

Rhythm

In Praat, the relative durations of syllables in imitation tasks were annotated and measured. For example, in the Syllables task, in the first instance, the stimulus and the imitation were separated into different files and separated again into words of 1, 2 and 3 syllables. In words of two and three syllables, the boundaries between syllables of the researcher's production were marked and measured. Given the difficulties in determining reliable syllable boundaries (Bell-Berti and Raphael, 1983) the boundary was marked at the point at which there was a clear change in pitch, intensity and movement in the second Formant (F2). The orthographic transcription of each syllable was marked on the image of the demonstration and this was used as a template from which to mark and compare the imitation: that is, the syllables in the imitation were divided at the same phonetic point as in the demonstration (e.g. *lo-ving*). As the imitation tasks were presented live, the process was repeated for each participant in order to allow comparison to the stimuli they heard.

Intonation

In Sonneta, mean changes in F0 across phrases were measured in order to calculate the participant's voice range in speech in Hz. These were then converted to semitones via Hz-to-Semitones converter (Traunmüller, 2005). However, Praat was used in order to measure the semitone changes between words in songs or phrases. The mean value of the highest pitch contour on the image was generated using Praat's 'get pitch' function. The lowest value was then generated and subtracted from the highest measure. These measures were recorded in both Hz and in Semitones (see also, Section 3.5.4).

Articulatory speed

The percentage of articulation time was calculated for each production of the phrase *Magic Book*, in order to capture any changes in articulatory speed. The duration of consonant sounds (milliseconds) and the duration of vowel sounds (milliseconds) were measured and tallied for each word that each participant produced in the first phrase of the song (Figure 3.14). The ratio of C:V was then calculated, allowing comparison of changes in speed of consonant production between successive productions, and comparison to the demonstration.

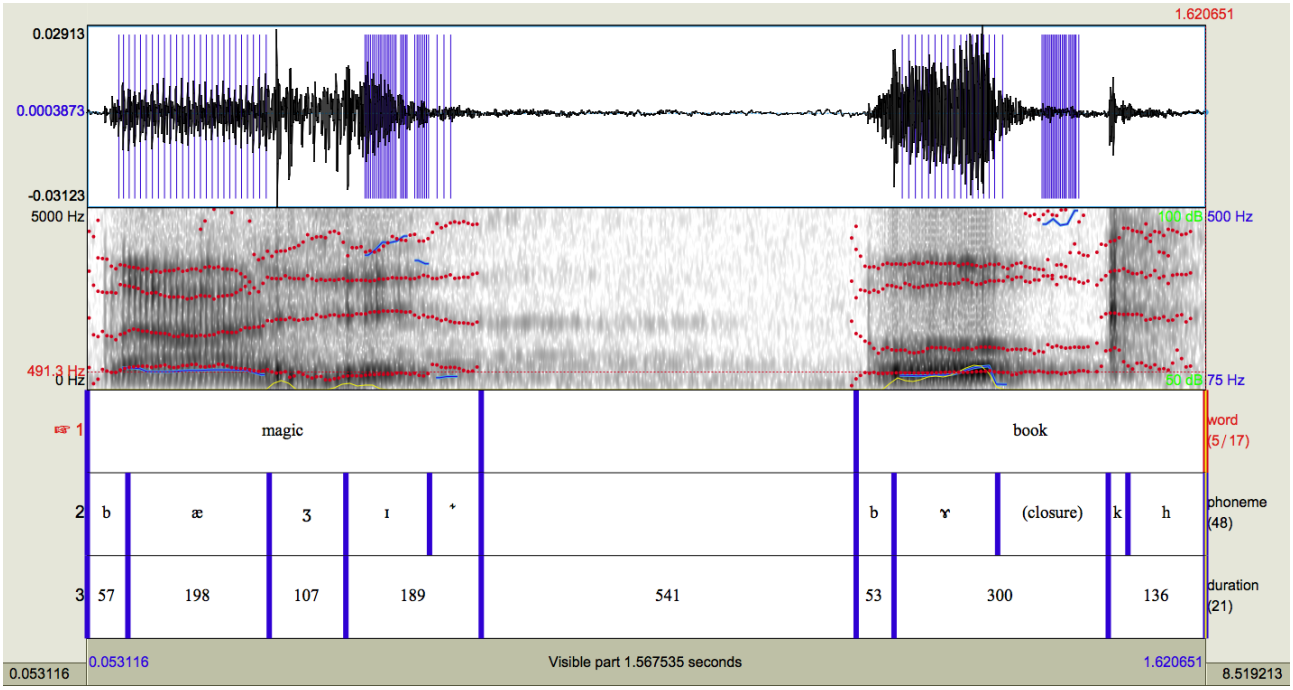


Figure 3.14: Screenshot of Praat showing the measurement of articulatory features for 'Magic Book'. The duration in milliseconds of consonants and vowels is shown in Tier 3.

Settings: For the calculation of pitch changes, a range of 50-500 Hz was used for F0, in order to adjust for the vocalisations that approached vocal fry. Autocorrelation was used as the aim was to measure intonation. For calculations in semitones, the unit was changed to 'semitones re 1 Hz'. All other parameters remained the same.

Perceived stress

Perceived stress was judged by ear. The dominant syllable was noted and is indicated in the text by the use of bold type for primary stress. Secondary stress was judged by ear, and is marked in the text by underlined text. The intensity contours and amplitude of the waveform were also used to support and inform perceived judgements.

3.6.4 Use of SPSS for voice data

Measures of fundamental frequency and voice perturbation were exported from Sonneta and copied into SPSS v.21. Data from all tasks were included within the same spreadsheet but the target word from which the vowel was taken was entered along with the identifying test name and number. The data were coded in SPSS to indicate the following values:

1. the position of the syllable in the word (first, medial or final). Words of one syllable were marked as 'first', final vowels in two-syllable words were marked as 'final';
2. the target number of syllables in the word;
3. the category of target number syllables: one, two, three or more;
4. the number of target words in the phrase (if relevant);
5. the number of syllables produced; *and*
6. the number of words produced.

Within SPSS, descriptive analysis tests were conducted for a combination of parameters: mean data was generated for each test, and according to the target number of syllables in the word, the actual number of syllables produced, and the position of the syllable in the word.

Voice tasks were recorded on different days and in different environmental conditions. Therefore, it was not appropriate to test for statistical significance for any data between tasks. Similarly, single tasks did not generate enough data for statistical analysis.

3.6.5 Intra-rater reliability

After a period of at least 12 months, a selection of non-live tasks were re-assessed. Tasks were chosen for re-assessment whose reliability was considered most critical to the arguments presented in the case studies (Table 3.2). The following measures were selected, which allow comparison to other studies:

- Measures of PPE in DEAP: the first five words were retested (20% per participant);
- Beat entrainment at three tempi: all measures at 104 bpm, whose results were the most variable within participants; and at least two bars (8 beats: 25%) at 132 bpm and 160 bpm for each participant. These were taken from the onset of playing that was at the tempo appropriate for the task: that is, if a visual cue was given, the measurement was taken from the participant's first tap;
- Voice perturbation measures in the first six of each participants sustained vowels (58%).; *and*
- The duration and percentage of interharmonics in sustained vowels. Five sustained vowels were retested for Andrew (55%), Kerry (45%), Rachel (50 %); and four (36%) were retested for Robert, who produced fewer vowels that were suitable for measures at onset.

The data from the original and final assessments were entered into SPSS, and a Pearson's correlation test was applied. The results indicated high intra-rater reliability for all tasks, with statistically significant ($p<0.05$*) or highly-significant ($p<0.001$**) effects (Table 3.2). However, there was individual variation in the validity of measures for individuals. Although the group data is reliable for the percentage of interharmonics that were produced in initial, medial and final portions of sustained vowels ($p<0.05$), the data were less reliable for Andrew, and validity did not reach significance for the percentage of interharmonics in final position of vowels (Table 3.2). On examination of the data, this result was influenced by the initial inaccurate measurement of the duration of interharmonics at offset in one vowel only, of the order of 12 ms. There was not sufficient data for re-testing of this measure in other tasks, which had simply indicated the proportion of interharmonics in each test, rather than measured each vowel in each test. Furthermore, the data for interharmonics is supported with visual evidence in the case studies.

Table 3.2: Results of the intra-rater reliability test scores. The table shows the mean reliability scores for core tests, and the results for each participant, as generated by Pearson's correlation test in SPSS. The number of items in the group data and the percentage of each set of data that this represented are given. Results that are statistically significant are marked with * ($p<0.05$: 2-tailed) and ** ($p<0.001$: 2-tailed). The text in red show results that did not achieve statistical significance.

	Test	Measure	n=	% of data (per test)	All	Andrew	Kerry	Rachel	Robert
Speech	**DEAP**	**PPE**	**40**	**20**	0.827**	1.000**	0.709*	0.733*	0.890**
Rhythmic accuracy	**Beat Entrainment at 104 bpm**	**ITI**	**81**	**100**	0.921**	0.958**	0.999**	0.968**	0.850 **
		AT	**69**		0.998**	0.999**	0.998**	0.992**	0.997**
	Beat Entrainment at 132 bpm	**ITI**	**43**	**25**	0.995**	0.932**	0.938**	0.939**	0.932**
		AT	**43**		0.998**	0.997**	1.000**	0.951**	0.999**
	Beat entrainment at 160 bpm	**ITI**	**33**	**25**	0.995**	0.845**	0.988**	0.830*	0.845**
		AT	**31**		0.997**	0.997**	0.999**	0.994**	0.869*
Voice perturbation measures	**Sustained vowels**	**HNR**	**24**	**58**	0.997**	1.000**	0.999**	0.999**	0.998**
		CPP			0.991**	0.988**	0.988**	0.995**	0.997**
		F0			0.999**	1.000**	0.996**	1.00**	1.000**
		DUV			0.944**	0.996**	0.776*	0.945**	0.936**
		Jitter			0.996**	0.999**	0.998**	0.967**	0.959**
		Shimmer			0.992**	1.000**	0.994**	0.989**	0.988**
Interharmonics	**Sustained vowels: percentage of interharmonics at onset, medially, and at offset**	**Initial**	**19**	**46**	0.993**	0.903*	1.000**	0.942*	0.998**
		Medial	**19**		0.990**	0.979**	0.981**	1.00**	0.986*
		Final	**19**		0.979**	0.379 (sig. .530	0.924*	n/a- constant	n/a - constant

3.7 Preparation of the case studies

3.7.1 Selection of data for inclusion in the case studies

For each participant, the core assessments were analysed in order to generate mean values on standardised tasks. These would allow comparison between individuals and between the group and the DS population. Chapter 8 draws on the core data in order to compare the abilities of the individuals within the group to the wider DS group. Although the data are summarised in Section 1 of each case study (Chapters 4-7), their use in each case study is limited to informing the discussion and to inform the selection of tasks to examine in greater detail. Given the scope and amount of data available, it was necessary to select examples and omit others in order to present the case studies. However, there is a degree of homogeneity between the participants, in line with the DS phenotype. Therefore, each chapter has a core set of tasks in common (beat entrainment, rhythm imitation, sustained vowels, words of increasing syllable length, pitch development and song learning, and production of phrases and songs from long-term memory). However, each individual also demonstrated differences in ability, and the data suggested that factors such as behavioural response, previous learning, memory and HI may account for some of these. As a result, each case study also contains detail that may be absent in that of another. For example: Chapter 4 presents data regarding Andrew's voice production according to the target number of syllables and their position in the word; Chapter 5 presents data on Kerry's voice production when reading; Chapter 6 examines Rachel's ability to segmental speech syllables when given visual support; and Chapter 7 gives additional detail of Robert's ability to learn and develop, when given visual support.

3.7.2 Adaptations and additional methodology for case studies

Additional methods of analysis or adaptations of methods were required for individual case studies. These are given below.

Andrew (Chapter 4)

Not all Andrew's vowels in speech tasks were produced with periodic phonation, which resulted in 'missing' values for HNR, jitter and shimmer, which depend upon periodicity (see Section 3.6.2). The percentage of 'unvoiced' vowels in tasks were therefore calculated in order to capture any changes in task or task demand. CPP values were generated and are given for all vowels.

The proportion of unvoiced segments and vowels in syllables affected the accuracy of both Praat and Sonneta to produce automatic measures for pitch. Andrew's pitch changes were therefore calculated manually. Where available, voice data from Sonneta were used to measure pitch changes in Hz: these were then converted to semitones via Hz-to-Semitones converter

(Traunmüller, 2005). Where no pitch was recognised, the vowel was played into the internal microphone of a chromatic tuner on iPad to measure frequency: this was adjusted if necessary to the correct perceptual octave, by emulating the perceived pitch of Andrew's production. This allows an approximate representation of Andrew's changes in perceived relative pitch and comparison to the stimulus.

The majority of Andrew's speech productions were produced with a high number of interharmonics. As a result, several samples were at ceiling level, when the standard settings (after Cavalli and Hirson, 1999: see above) were used. Therefore, a second set of calculations were applied to capture the intensity of interharmonics by reducing the dynamic range to 15 dB. The duration of the interharmonics that remained visible in voiced components was calculated and expressed as a percentage of the total duration of the voiced segments, as detailed above (see Figure 3.8).

Rachel (Chapter 6)

Rachel produced a number of utterances during tasks that suggested frustration or unwillingness (e.g. '*I can't do it'*). These were tallied in tasks and used as an indicator of her stress-response to the task. During assessments, and in line with ethical procedures (see Appendix 1) judgements were made on the basis of her comments whether to continue with tasks. For this reason, Rachel was not asked to imitate all phrases in the Newton (1999) task.

3.7.3 Selection of examples of voice production from individual tests

Samples were chosen from each test to illustrate both typical and unusual features of voice production and to confirm the qualities of phonation when a measure (such as CPP or HNR) may indicate either breathy phonation or harsh phonation. Titze (1994) recommends that spectrograms are the best form of analysis for Type 2 voices, that contain qualitative changes (bifurcations) and multiple harmonics that approach the fundamental frequency. For presentation as Figures, examples were chosen that reflect the variability of voice production across tests and within tests, and that illustrate similarities between examples.

3.7.4 Presentation of data in Figures

Speech segments were annotated in Praat using IPA: however, annotations for atypical voice production do not appear faithful on the text grids and diagrams. For this reason, figures are annotated with broad phonetic transcriptions. Where appropriate, a narrow transcription is given in the text to illustrate details of voice production, using extensions to the International Phonetic Alphabet (extIPA) (Duckworth et al., 1990) and Voice Quality Symbols VoQs (Ball, Esling, and Dickson, 1995).

For images, the pitch, intensity and formant contours are presented for illustrative purposes, only, in order to indicate prosodic features. Pitch settings were set to auto-correlate for all images. The above sections detail the settings used for generation of data, but settings were chosen for the images that best illustrated the features discussed in the text. The specific settings used are given in the heading of each Figure. Likewise, for visual representations, the window-length, frequency range and intensity settings of the spectrograms were adjusted to demonstrate the clearest resolution. These settings are also given in the Figure headings.

Where the environmental noise adversely affected the recording, the 'robust' features of intensity and pitch contour are presented on diagrams but the spectrogram is not. In order to examine examples of voice production that were recorded against a noisy background, recordings were filtered in Praat. This used the standard settings within the 'remove-noise' filter for spectral subtraction (window length 25 ms; frequency range 80-10,000 Hz; smoothing, 40 dB).

Copyright permission for use of lyrics.

Permission was sought from the copyright owners to reproduce the lyrics of the songs *Monkey Man* and *Magic Book*. Lyrics from the song *Monkey Man* (Hibbert, 1979) are reprinted with permission of Hal Leonard Corporation. Judith Brindle granted permission to reproduce the lyrics to the song *Magic Book*, which she had adapted from a song of unknown authorship.

Chapter 4: Andrew

This chapter will present the results of Andrew's core tests, and will include details of his voice production in imitation, recall and spontaneous speech tasks. Section 4.1 will present a summary of his background, of his learning behaviour and progress during the teaching sessions (Chapter 3, Section 3.4). It will also summarise the data from his performance in core speech, cognitive and musical tasks, and highlight his key characteristics. Section 4.2 will present results from motor movement tasks. Section 4.3 will focus on Andrew's ability to reproduce prosody and to produce voice in imitation tasks. Section 4.4 will provide examples of his prosodic and vocal production speech and singing tasks that drew on long-term recall. The results will be discussed in Section 4.5 and the implications for the use of music to support Andrew's speech will be discussed in Section 4.6.

4.1 Summary of history, learning and performance in core assessments

According to his college records, Andrew is a young man of 23;9 years with SLD, moderate speech intelligibility and a history of behavioural difficulties, related to non-compliance and poor attention. According to his speech and language therapist (SaLT), his voice is gruff and harsh in quality, and his hearing is 'thought to be adequate' for speech. His SaLT record states that he responds to high frequency sounds but has some difficulties discriminating low frequency sounds, which affects his ability to perceive whispered speech. No formal test results were available. According to his SaLT report, he understands sentences of up to three content words and his expressive language is limited to functional words. His speech is described by his SaLT as 'quite clear' in terms of articulation and fluency and he has no reported phonological difficulties. His college records state that Andrew is a keen singer and performer with a preference for musicals and pop music and that he has over four years' experience of singing and of choreographed dance to music.

4.1.1 Cognitive and perceptual abilities

Test results (Table 4.1) suggest a verbal MA equivalent of 3;4 years and a nonverbal MA of 5 years. However, Andrew was impatient to complete the DAP test and this may not reflect his actual ability. When imitating sequences of numbers in TAPS test, he was accurate in both trials of three-digit sequences and one at four digits.

Perceptual tests (Table 4.1) indicate that Andrew has no difficulties in speech-sound discrimination (test 13) or in discriminating words on the basis of rhythm (test 14), but he

Table 4.1: Profile: a summary of Andrew's performance in core tasks. The table gives the results of core assessments in nonverbal and verbal ability, speech perception and production and musical perception and production. Test numbers refer to the test descriptions in Chapter 2, Section 3.3.

Test number		Test name	Results
Cognition	**1**	BPVS RAW score	28
	1	BPVS Age equivalent	3;4 (2;11 – 3;9)
	2	Draw-a-person	5;0
	3	TAPS - Digit Span	3-4
Perception (speech)	**7a**	Same-different discrimination nonwords (Stackhouse et al., 2007, App C.2)	4/10 = 40%
	7b	Same-different discrimination real words (Stackhouse et al., 2007, App C.2)	5/10 = 50%
	13	Minimal pairs perception: discrimination of consonants	5/5 = 100%
	13	Minimal pairs perception: discrimination of vowels	5/5 = 100%
	14	Simple speech prosody test rhythm melody	4/4=100% 2/4=50%
Speech Production	**5**	DEAP: Percentage of Phonemes Correct (PPC)	78
	8	s/z	n/a
	9	DDK	n/a
Perception (music)	**17**	PMMA	n/a
	26	Pitch perception test: pitch matching	1/3 (pitch) = 33%
	27	Rhythm perception test (same/different)	4/6 = 66%
	27	Pitch and interval perception test (same/different)	4/6=66%
	28	Find the sound	2/5= 40%
Music Production	**18**	Gross motor score	31/40=77%
	4	Rhythm imitation (TRaCoL)	7/14=50%
	22a	Pitch matching: vocal glides on 'la'	2/4=50%
	22b	Pitch matching: Sustained /i/, /u/, and /a/ vowels at three pitches	0/9 = 0%
Reading ability	**15**	words correctly produced ('Swimming' passage) incorrect words : yell = yet; cool = cold words omitted: 'I can' x4, 'and', 'after', 'with' rate: 2 syllables/second	78% n=2 n=21; 20%

performed at less than chance on word-discrimination tasks (test 7). Results of the music perception tasks do suggest a reduced ability in melodic perception tests but an ability to discriminate between pairs of words on the basis of rhythm (e.g. between 'beans on toast' and 'spaghetti').

During speech and music perception tasks that required discrimination of 'same/different' (Table 4.1: tests 7, 14 and 27), Andrew's responses were inconsistent when items were repeated. He also alternated 'same' and 'different' responses to nonword discrimination tasks (test 7b). He was able to determine 'same/different' in pre-test items, and was accurate in stating 'same/different' in rhythms of contrasting duration and six beats or fewer: however, above this level of items, he became uncertain and changed his mind.

4.1.2 Motor-music abilities

Observation of Andrew's movements against a checklist (Appendix 2, p. 324) showed that Andrew was able to copy and create large motor movements to pop music, and that he could copy marching movements closely to the beat. However, he exhibited difficulties in timing different types of gross-motor movement to the beat. His marching movements, requiring alternating steps on the spot, were slower than those that were demonstrated, and varied in regularity of movement. His arm movements matched those of the demonstrator, as did his upper torso twists. However, these large motor movements were timed at half the speed of tapping feet and marching, and the timing of his movements became irregular at the faster rate. When asked to clap 'in time' to the music, Andrew clapped at a much slower rate than the beat (4/4) and was unable to synchronise to the music. Following a demonstration, he clapped at a faster pace that was more appropriate to the music, but decreased in tempo once the visual support stopped.

4.1.3 Oro-motor and phonological abilities

According to Andrew's SaLT report he has a high palatal arch and reduced tongue mobility, and he tends to use his jaw to compensate for muscular difficulties in lateral and upwards tongue movements. He was unable to complete the DDK test (Table 4.1: test 9).

Andrew's PPC score on DEAP is 78% (Table 4.1: test 5) when adjusted for the habitual error of gliding (Table 4.2). Andrew had articulatory difficulties and produced some sounds with incomplete closure. This led to frication of some stops: e.g. unvoiced and voiced bilabial stops /p/ and /b/ were produced as fricatives [ɸ] or [β]. Error patterns (n=5+ occurrences)

Table 4.2: Error patterns produced by Andrew in the DEAP task. The table shows the number of errors made in the task and the error patterns (n= 5 or more; except WSD* n=3), expressed as a percentage of the total number of errors made. The errors are coded according to the conventions outlined in DEAP, which defines the following categories: D indicates delayed error patterns; U indicates unusual error patterns.

DEAP		
Percentage of error pattern/total errors		
Gliding	**D**	34.0
Cluster Reduction	**D**	13.0
Fronting of velars	**D**	13.0
Weak syllable deletion*	**D**	7.8
Other	**U**	7.8: e.g. /ʃ/ = [t]; /θ/= [w]; /ʃ/= [nt]
Number of errors		38.0
Percentage of errors accounted for by error patterns		75.6

are shown in Table 4.2. Other errors included deaffrication, stopping and devoicing: for example, the production of [tɹwæf] for 'giraffe'. The mean percentage of phoneme errors in all tasks was within a range of 22-30%, except for imitated words of one syllable (Table 4.3). Consonant errors were highest in CNREP (test 6) and lowest in reading (test 15); however the percentage of vowel errors (PVE) and omitted sounds (PEO) were similar in both tests (Table 4.3). Only reading and spontaneous comments were produced with overt dysfluency which included repetition of words and syllables: only one word was produced with initial blocking of the sound. Andrew was able to read confidently but he omitted several repeating words (*I*, *can*) and he processed words semantically (e.g. he mis-read 'cool' as *cold*).

Table 4.3: The percentage of speech errors produced by Andrew in core speech tasks. The table shows the percentage of errors in phonemes (PPE), consonants with habitual errors removed (PCE-R), vowel errors (PVE), the percentage of sounds omitted (PEO) and the percentage of dysfluency.

	Task	**n=**	**PPE**	**PCE-R**	**PVE**	**PEO**	**Dysfluency**
5	**DEAP**	**45**	21.74	29.23	3.70	4.85	0.00
11	**Syllables**	**27**	**24.32**	**32.56**	**14.81**	**5.30**	0.00
	1-syllable words	**9**	12.87	18.39	0.00	0.00	0.00
	2-syllable words	**9**	30.85	42.45	22.22	6.67	0.00
	3-syllable words	**9**	29.23	34.21	22.22	9.23	0.00
6	**CNREP**	**10**	51.06	58.62	24.24	21.67	0.00
10	**Phrase imitation (Newton (1999) CSP repetition)**	**25**	46.05	40.91	40.54	32.89	0.00
15	**Reading**	**24**	29.02	29.55	21.43	16.34	22.5% repetition of word/syllable; 1 stutter
n/a	**Spontaneous speech: comment 1**	**16**	18.52	31.50	0.00	18.52	11syll /20 repeated

4.1.4 Voice range and measures of voice quality

Andrew was capable of producing a vocal range of an octave when imitating glides on vowels (Table 4.4: test 22), but his voice range in connected speech was less than half an octave (Table 4.4: test 10). He produced modal phonation in the central portion of vowels (Table 4.5), but onset and offset of vowels were produced with high values of DUV and jitter. All sustained vowels were produced with atypical phonation, as indicated by the presence of interharmonics on the spectrogram (Table 4.5). The mean perturbation measures for sustained vowels are considerably higher than for all vowels in speech. Although the data were collected in different sessions (except for *), the voice data in speech tasks are similar for most measures. However, not all vowels were produced with sufficient periodic phonation to generate measures (Table 4.5).

Table 4.4: The pitch range of Andrew's vocal glides and connected speech. For each task, the table shows the mean, minimum and maximum values of fundamental frequency in Hz, and the SD of F0 around the mean, in Hz and in ST. The mean range is also shown in semitones (ST). Mean F0 and SD of F0 is not given for glides, whose purpose was to examine vocal range.

Test no:	Test name	mean F0, Hz	Min F0, Hz	Max F0, Hz	SD of F0, Hz; ST	Pitch range, (ST)
22	**Pitch glides on 'la'**		108.94	238.84		13.59
10	**Picture description**	118.57	53.71	221.21	30.17; 4.43	18.96
n/a	**Spontaneous speech (comment for K)**	97.05	50.29	193.71	28.37; 5.08	n/a - ventricular

Table 4.5: Mean voice perturbation measures for Andrew in sustained vowels and vowels in words. For sustained vowels, mean measures are given and separate values are given for the initial 100 ms (onset), medial duration (Total duration - (onset+offset)), and the final 100 ms (offset). Measures are given for the vowels in each syllable of words produced in each task. The tasks marked with an asterisk were recorded within the same session, in that order. The DEAP task (shaded) was of inferior recording quality: data are provided for information only. The table shows the mean measures of voice breaks, unvoiced frames, maximum phonation time, harmonics-to-noise ratio, jitter (%), shimmer (%). The table also shows the percentage of vowels that were produced with interharmonics.

Test no:	Test name	No. of vowels produced, n=	No. of valid vowels for voice measures, n=	DVB, %	DUV, %	HNR, dB	Jitter, %	Shimmer, %	CPP, dB	No. and percentage of vowels produced with interharmonics
19	**Sustained vowels: Total duration (T)**	**9**	**9**	3.01	3.67	23.21	0.67	3.25	6.16	9/9=100%
	Onset (t=100 ms)	**9**	**9**	0.00	22.90	14.45	1.59	8.52	3.61	9/9=100%
	Medial duration (T- [onset+offset]) ms)	**9**	**9**	4.02	4.12	21.23	0.60	3.77	6.24	7/9=78%
	Offset (t=100 ms)	**9**	**9**	4.57	43.29	13.10	4.86	14.09	3.59	9/9=100%
5	**DEAP**	**65**	**27**	1.29	45.01	4.77	3.02	18.83	1.96	65/65=100%
11	**Syllables**	**56**	**54**	4.71	30.25	8.67	3.65	16.39	2.58	50/56=89%
6	**CNREP***	**34**	**34**	0.78	30.48	4.78	4.85	18.13	2.09	32/34=94%
10	**Newton (1999)***	**38**	**22**	2.43	35.55	4.81	3.33	17.34	2.08	38/38=100%
15	**Reading***	**117**	**24**	8.10	32.45	4.64	4.48	13.87	2.46	116/117=99%

4.1.5 Learning

In addition to group targets (Appendix 3, p. 340), Andrew's targets were to:

- listen and think before repeating songs/rhythms;
- sustain phonation for three seconds at three pitches using a kazoo;
- match pitch for notes that are within speaking range; and
- improve accuracy in singing melodic intervals.

Andrew did not always listen and wait before engaging with a task and he sometimes refused to participate (see Appendix 3: p. 345 [15][1]; p. 350 [33]; p. 353 [8-11]). He developed in his ability to produce continuous phonation using a kazoo, and to sustain steady exhalation for three seconds. This is below the typical adult MPT and below that of DS adults (Lee et al., 2009). Andrew habitually tensed his shoulders and upper torso when preparing to hum. He required reminders to 'relax' in order to activate the kazoo and to blow pinwheels (Appendix 3, p. 348 [1]). His development in pitching will be discussed in Section 4.3.

4.2 Motor-music entrainment and imitation

4.2.1 Beat entrainment and extraction

Andrew's accuracy in synchronising to a beat was variable and changed according to tempo (AT, ms; AT, %: Table 4.6). At 104 bpm, Andrew initially played at half the tempo. After a visual demonstration, he doubled his speed: the results at the target speed are given in Table 4.6. At 104 bpm his timing to the beat (AT, ms; AT, %) was on average 13% out; at 132 bpm this increased to 42%; and at 160 bpm his AT(%) averaged 35%. The accuracy of mean timing between his beats (ITI, ms; ITI.%: Table 4.6) ranged between 102% (104 bpm) and 94% (160 bpm). The instability of his timing (SD of ITI, ms; SD of AT, ms) increased as the target tempo increased. However, his performance at 160 bpm improved following a visual demonstration (not shown). Andrew did not produce any consistent changes in the intensity of his playing to mark meter.

Andrew's timing when he played the drum to the *acapella* song, *Can You Tap The Beat?* was variable in the first phrase of the song at each tempo. Table 4.6 shows Andrew's performance during the second phrase only. His performance was most accurate at 100 bpm (ITI, %, AT, %) and the least variable (SD of ITI, ms; SD of AT, ms: Table 4.6).

[1] numbers in square brackets refer to the line numbers on each page in Appendix 3

Table 4.6: The accuracy of Andrew's timing in beat entrainment and extraction tasks. The table shows mean ITI and AT values in milliseconds and as percentage of the target interval, and the SD in milliseconds (ms). Mean results are given for entrainment tasks and for extraction tasks.

	Entrainment			Extraction		
Tempo, bpm	**104**	**132**	**160**	**87**	**100**	**132**
Interval time, ms	**576**	**455**	**375**	**686**	**597**	**454**
ITI, ms	545	460	470	1210	597	569
ITI, %	102	97	94	176	100	125
SD of ITI, ms	30	80	180	222	49	194
Mean range of ITI, %	3.7			33		
AT, ms	-76	-193	-132	-251	-12	-64
AT, %	13	42	35	37	2	14
SD of AT, ms	220	680	560	151	59	164
Mean AT, %	30			17.7		

4.2.2 Rhythm entrainment

Andrew did not spontaneously entrain to either rhythm: instead, he played at an even tempo but neither synchronised to the beat nor to the pattern. After three bars he began to sing and drum to *Twinkle Twinkle Little Star* and when prompted to re-focus he continued to play at an even pace with no rhythmic pattern. However, after a visual demonstration before the onset of Rhythm 2, Andrew played with a clustering of the first four beats, similar to that seen in the target rhythm, although these were not synchronised to it (Figure 4.1: shaded sections). The onset of his pattern synchronised more closely with the onset of each new bar (e.g. Andrew's tap number 18, Bar 3, R2). He repeated the pattern most accurately in Bar 2 (Table 4.7), but the intensity of his taps were variable (Figure 4.1).

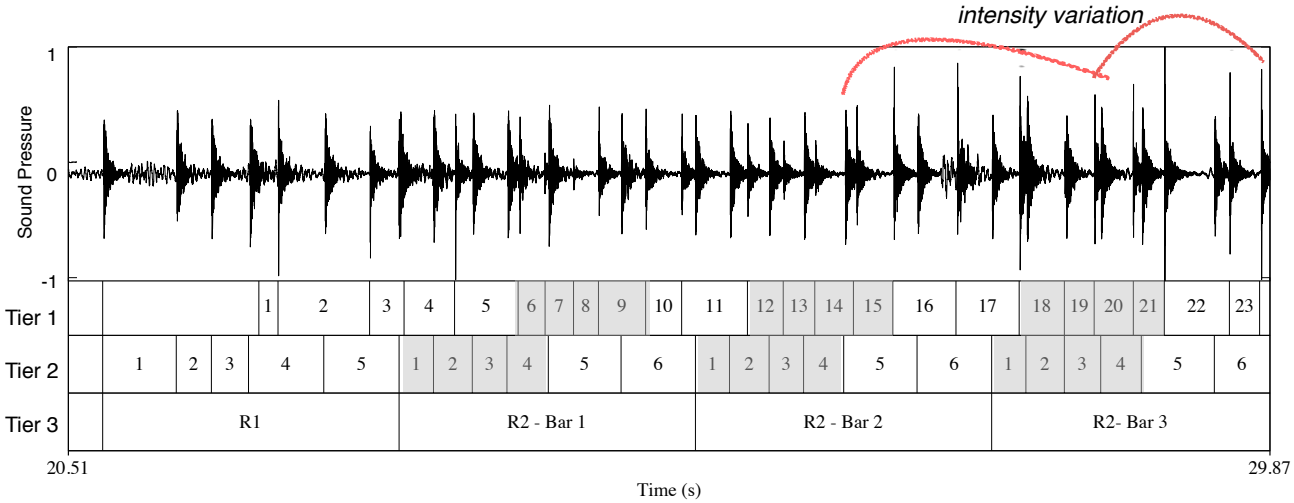


Figure 4.1: Image showing Andrew's rhythmic production in the Rhythm Entrainment task – *tee-tee tee-tee ta ta*. The image shows the composite waveform of the stimulus and Andrew's production of the final bar of Rhythm 1 and first three bars of Rhythm 2. Andrew's beats are shown in Tier 1, numbered 1-23. The target rhythms are marked in Tier 2: the short notes are shaded in the demonstration and in Andrew's production. Andrew was supported with a visual demonstration during Bar 1 of Rhythm 2. The relative timing of his taps is comparable to the stimulus during beats 12-17 (Bar 2). Note the variation in the intensity of Andrew's playing in Bars 2 and 3.

Table 4.7: The accuracy of Andrew's timing in three bars of rhythm entrainment, following a visual demonstration. The table shows the mean value and range of inter-tap intervals (ITI) between the short rhythm (tee-tee) and the contrasting long rhythms (ta) (see Figure 4.1). The target values are shown and Andrew's means are expressed as a percentage of this.

Target: note name and numbers	**ITI, ms**	**Bar 1**	**mean accuracy, %**	**Bar 2: 12-17**	**mean accuracy, %**	**Bar 3: 18-23**	**mean accuracy, %**
tee-tee (1-4)	**289**	247	85	282	97	281	97
range	***268-312***	*199-360*		*245-306*		*240-351*	
ta (5-6)	**575**	396	69	492	86	371	64
range	***n/a***	*280-512*		*571-580*		*507-235*	
relative values of tee: ta	**1: 2= 50%**	62		57		76	

4.2.3 Rhythm imitation

In terms of number and relative duration, Andrew reproduced both two-clap rhythms, two of the four three-clap rhythms, and four of the five four-clap rhythms (Table 4.8). He partially reproduced two rhythms of three claps (numbers 3 and 4), but reproduced too many claps in rhythms of four target claps (numbers 8 and 12). He produced too few claps in rhythms of six or seven target claps (numbers 13 and 14).

Table 4.8: Andrew's accuracy in imitating clapped rhythms in the TRaCol task (Treharne, 1999). The table shows the accuracy of Andrew's responses in terms of reproducing rhythms in number and duration (N&D), and partial replication of the number (N) or relative duration (D). The notes produced in error are indicated in red text.

Target number of claps	**Target rhythms, with notation**	**Andrew: responses and number of correct responses**	
		N+D	**Partial**
2	**1.** ♩ ♩ **2.** ♫	1. ♩ ♩ 2. ♫	
3	**3.** ♩ ♫ **4** ♫ ♩ ; **5.** ♩ ♩ ♩ **6.** ♪ ♩ ♩	5. ♩ ♩ ♩ 6 ♪ ♩ ♩	3. ♫ (D) 4. ♫ (D)
4	**7.** ♫ ♫ **8.** ♫ ♩ ♩ **9.** ♩ ♩ ♫ **10.** ♩ ♫ ♩ **11.** ♩ ♩ ♩ ♩	7. ♫ ♫ 9. ♩ ♩ ♫ 11. ♩ ♩ ♩ ♩	8. ♫ ♩ ♩ ♩ (D) 10. ♩ ♫ ♩ ♩ (D)
5	**12.** ♩ ♩ ♫ ♩		♫ ♫ ♩ (N, D)
6	**13.** ♫ ♩ ♩ ♫		♩ ♫ ♩ ♩ (D)
7	**14.** ♩ ♫ ♫ ♩ ♩		♫ ♫ ♩ ♩ (D)
TOTAL	**14**	**7**	7
Percent correct		**50%**	**50%**

4.2.4 Syllable segmentation

Andrew was able to segment the syllables to three words independently. He produced the rhythm of caterpillar without speaking the word on three occasions, whilst producing four claps (twice) and five claps on the third production. He was able to produce a rhythm on a drum from the word sequence, prompted by a visual representation (Figure 4.2; Appendix 2, p. 326) and with a verbal prompt for the words and timing of the beat.

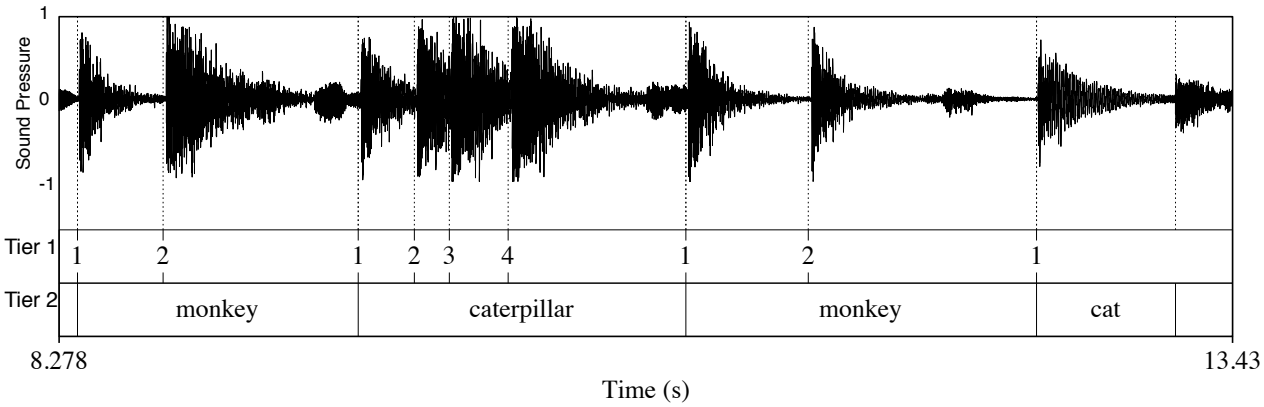


Figure 4.2: Annotated waveform showing Andrew's production of a rhythmic sequence from visual notation. The waveform and Tier 1 shows the onset of Andrew's taps. He did not say the words, but these were provided as the researcher pointed to each image in the sequence.

4.3 Voice and prosody in imitation tasks

4.3.1 Sustained vowels

Eight of the nine sustained vowels were produced with few perturbations in the central portion of phonation (Table 4.9): only /a/1 was atypical in perceived quality and in the degree of voice perturbation. The vowels /u/3 (Figure 4.3) and /a/1 (Figure 4.4) were chosen as examples. The sample for /u/3 is close to the mean values (Table 4.9) and can be considered typical of his production. In contrast, the sample for /a/1 is atypical: it was produced with audible diplophonia throughout and was atypically low in mean F0 (96.97 Hz), high in DVB (29.17%), and was produced with interharmonics throughout (Table 4.9).

Table 4.9: The mean fundamental frequency and voice perturbation measures for Andrew's sustained vowels. The table shows the mean data for the entire production of nine sustained vowels. Measures given are for mean F0 (Hz), SD of F0 (Hz), DVB (%), DUV (%), HNR (dB), jitter (%), shimmer (%) and CPP (dB). The duration (milliseconds) of single or complex (2+) interharmonics is shown, and their position in the vowel is indicated, as follows: onset (i; t=100 ms), medial (m: the period after onset, but excluding the final 100 ms), and final (t= 100 ms). Interharmonics that are medial but continuous with onset or offset are indicated with +. Bold type indicates values that are close to norms for jitter and shimmer for sustained vowels (Goy et al., 2013), and diplophonic production.

Vowel	Duration, ms	Mean F0, Hz	SD of F0, Hz	DVB, %	DUV,%	HNR, dB	Jitter, %	Shimmer, %	CPP, dB	1 Inter-H: Duration (ms) and position	2+ Inter-H: Duration (ms) and position
/i/1 -D	**1040**	208.64	2.53	0.00	1.13	24.96	**0.56**	**2.70**	6.45	0.00	39 ms(i); 90 ms (f)
/i/2-C	**1450**	177.47	2.03	0.00	0.00	24.51	**0.60**	**2.04**	5.58	0.00	93 ms (i); 100 ms (f)+49 ms (m)
/i/3 -D	**950**	188.34	2.34	0.00	0.00	24.64	**0.51**	**1.79**	6.67	0.00	97 ms (i); 100 ms (f)+10 ms (m);
/u/1 -D	**688**	214.52	2.86	0.00	1.96	28.35	**0.44**	**2.11**	5.24	0.00	72 ms (i); 100 ms (f)+ 7 ms (m)
/u/2 -C	**870**	161.91	1.28	0.00	0.00	28.81	**0.82**	**2.30**	6.67	0.00	100 ms (i)+16 ms (m); 91 ms (f);
/u/3 -C	**915**	183.13	1.59	0.00	1.50	26.27	**1.04**	**2.15**	6.20	0.00	93 ms (i); 91 ms (f)
/a/1 -C	**876**	96.97	1.81	27.08	29.17	5.14	1.97	8.68	4.19	**593 ms** (m)	**100 ms (i) +59 ms (m); 100 ms (f)+24 ms (m)**
/a/2- C	**1405**	155.95	3.59	0.00	0.00	22.64	**0.35**	3.26	7.85	0.00	49 ms (i); 100 ms (f)+16 ms (m)
/a/3 D	**1896**	177.59	0.72	0.00	0.00	24.05	**0.26**	**0.05**	7.47	0.00	100 ms (i) + 131 ms (m); 100 ms (f) + 32 ms (m)
Mean	**1121**	**173.84**	**2.08**	**3.01**	**3.53**	**23.26**	**0.73**	**2.79**	**6.26**		

Example: /u/3

The vowel /u/3 was sustained with little variation in F0, with unbroken phonation and high HNR and CPP values (Table 4.9). Phonation was perceptually modal throughout. The waveform (Figure 4.3) shows regularity in the period with little noise, except at onset and offset. There is a change in the contour of Formant 2 (in red: top) during Andrew's production which coincides with some changes in amplitude of the waveform: e.g. with a decrease in the waveform's amplitude during section 2, and a gradual increase in section 3. Offset of the vowel is abrupt and the spectrogram and waveform show a break in phonation (section 5).

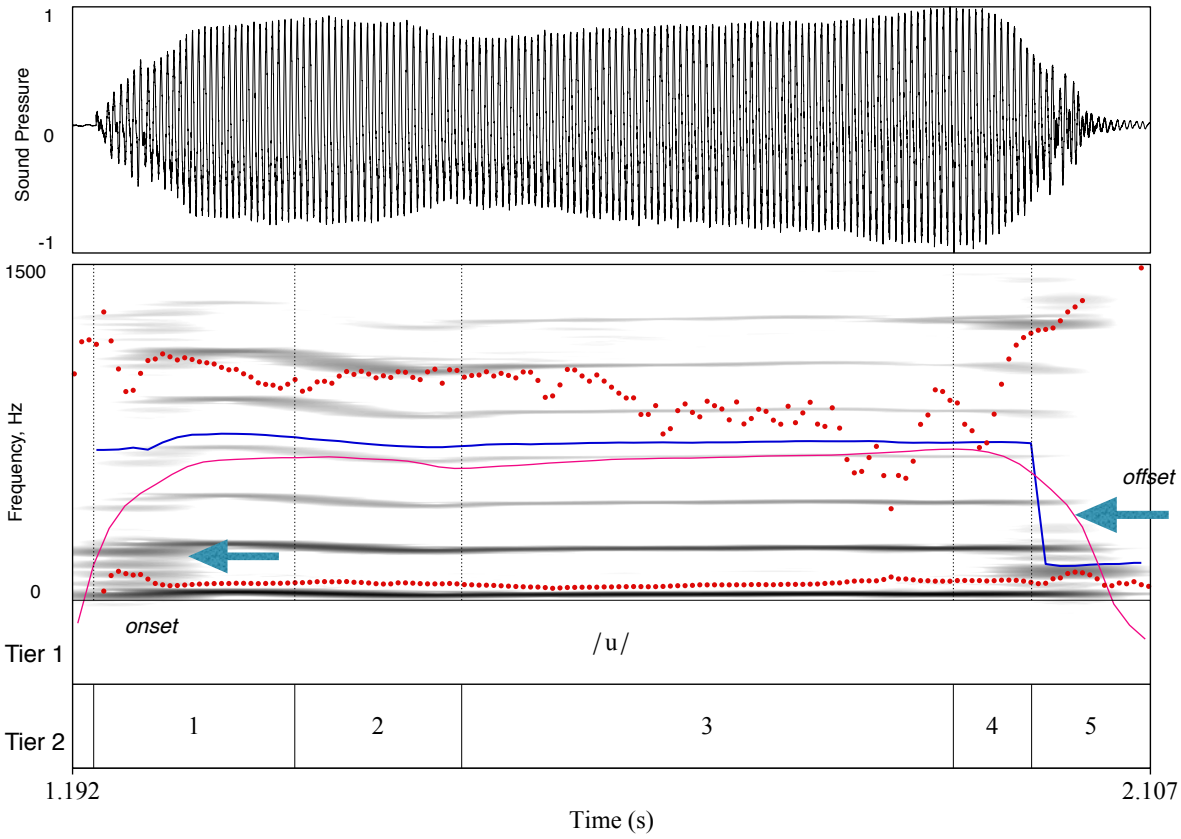
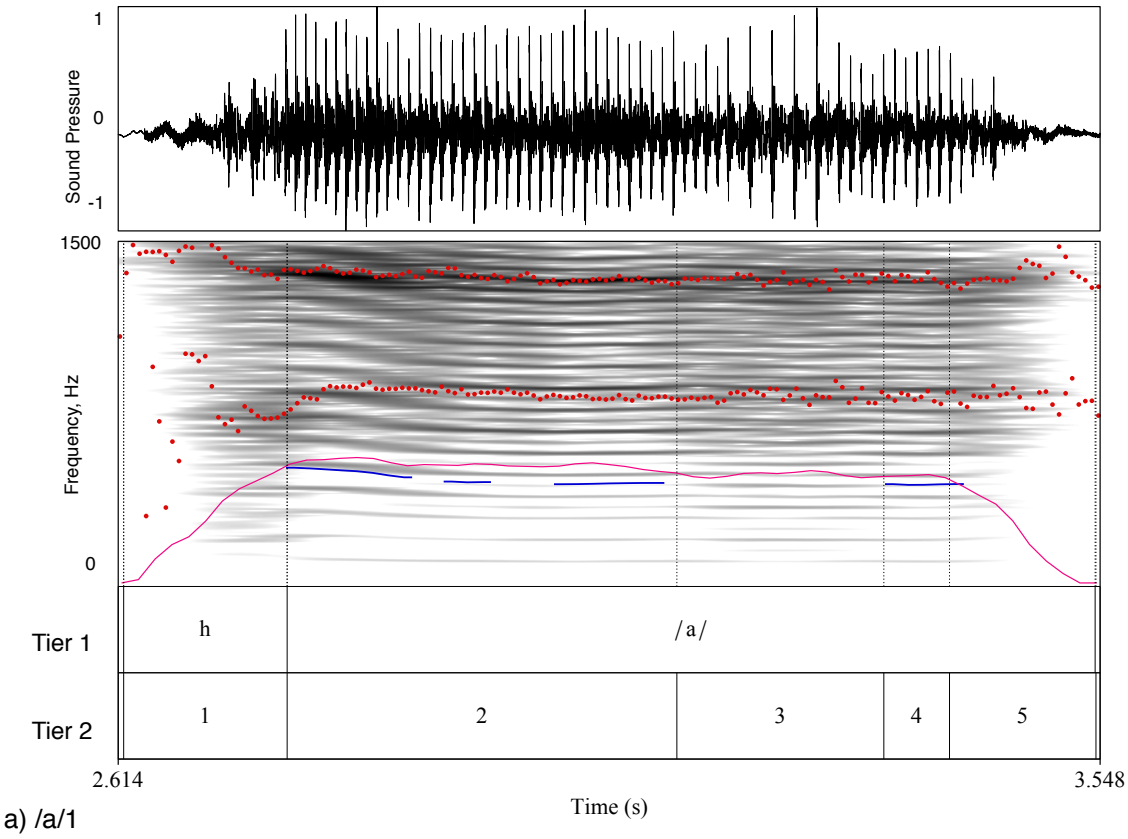


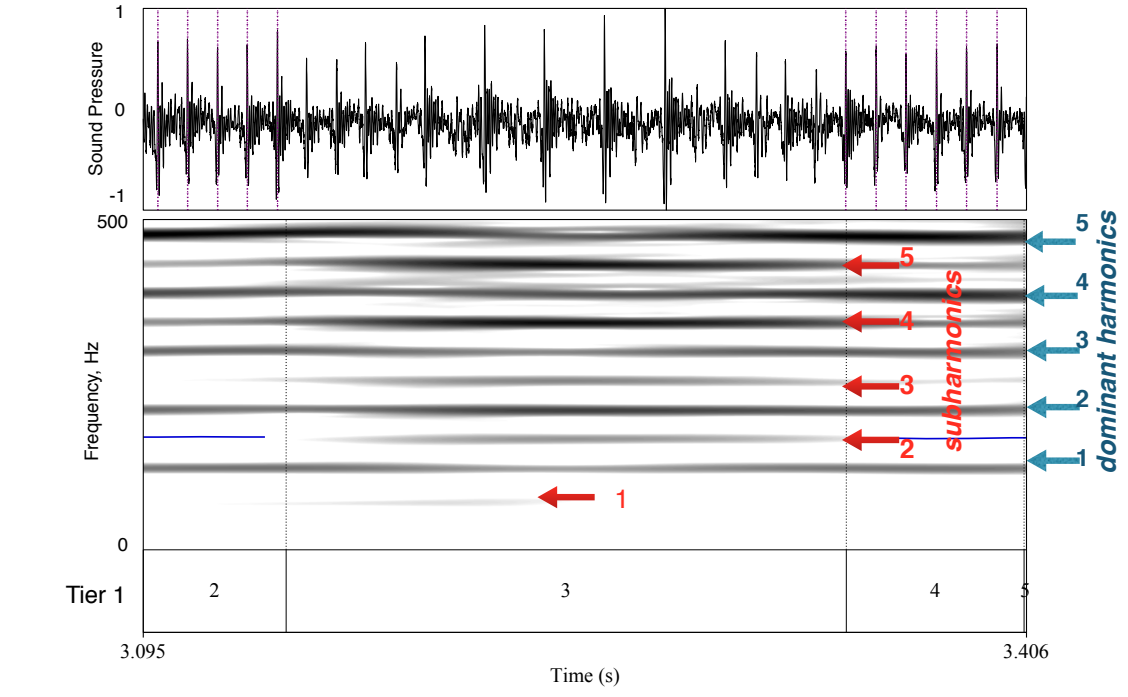
Figure 4.3: Waveform and annotated spectrogram showing Andrew's sustained vowel, /u/3. The image shows the waveform (top) and narrow band spectrogram (0-1500 Hz; window length 150 ms; 40 dB) of the sustained vowel. Pitch settings are 50-500 Hz, and intensity settings are 40-100 dB. The spectrogram is overlaid with the contours of the first and second formant frequencies (in red dots), the pitch contour in blue, and intensity contour in pink. Section numbers in Tier 2 relate to points of interest that are discussed in the text. The arrows mark interharmonics at onset and offset of the vowel.

Example: /a/1

Andrew's production of /a/1 was the shortest in duration. The data show very low HNR and CPP values and a high degree of unvoiced frames (DUV, %)(Table 4.9). The pitch contour is broken (Figure 4.4a), which illustrates the degree of unvoiced sections. Onset was aspirated (Figure 4.4a: section 1). The waveform is characteristic of 'pulse' phonation, and there are multiple interharmonics throughout, and clearly defined subharmonics. The dominant harmonics (Figure 4.4b) occur at multiples of about 96 Hz (Table 4.9), and sub-harmonics at about 48 Hz. Subharmonics also occur when voicing is absent, as indicated by breaks in the pitch contour. The production was audibly diplophonic when voicing occurred, and perceptually harsh and ventricular in quality throughout. The spectrogram shows an increase in the intensity of the dominant harmonic 2 and subharmonics 4 and 5 (Figure 4.4b: section 3).



a) /a/1



b) Detail

Figure 4.4: Waveform and annotated spectrogram showing Andrew's sustained vowel /a/1 at 97 Hz. The image shows a) the whole vowel; and b) detail of Sections 2, 3 and 4. The images show the waveform (top) and narrow band spectrogram (0-1500 Hz; window length 150 ms; 40 dB) of the sustained vowel. Pitch settings are 50-500 Hz, and intensity settings are 40-100 dB.The spectrogram is overlaid with the contours of the first and second formant frequencies (in red dots), the pitch contour in blue, and intensity contour in pink. Section numbers in Tier 2 (a) and Tier 1 (b) relate to points of interest that are discussed in the text. The harmonics relating to the fundamental frequency are obscured by subharmonics of differing intensity. An increase in the number of these on the spectrogram coincides with a break in the pitch contour. Figure b) shows this in the detail of section 3. For the spectrogram, the frequency range was reduced (0-500 Hz) and the dynamic range was reduced to 25 dB to show the contrast between the number of subharmonics.

4.3.2 Pitch and interval matching

Andrew was unable to match single pitches in initial or final assessments (Figure 4.5: A1). In his second assessment (A2), he produced notes over a wider pitch range and produced a rise in pitch between G3 and D4. His productions of G4 and C5 were closer to an octave below the target (Figure 4.5).

Figure 4.5: Graph showing Andrew's accuracy when matching melodic intervals during his first and final assessments. The fundamental frequency of the target pitch (in Hz) is shown in blue and Andrew's mean F0 is shown for productions in his first assessment (A1) and final assessment (A2), which were six weeks apart. The target pitch is shown in blue and the same note is shown at an octave below, in orange.

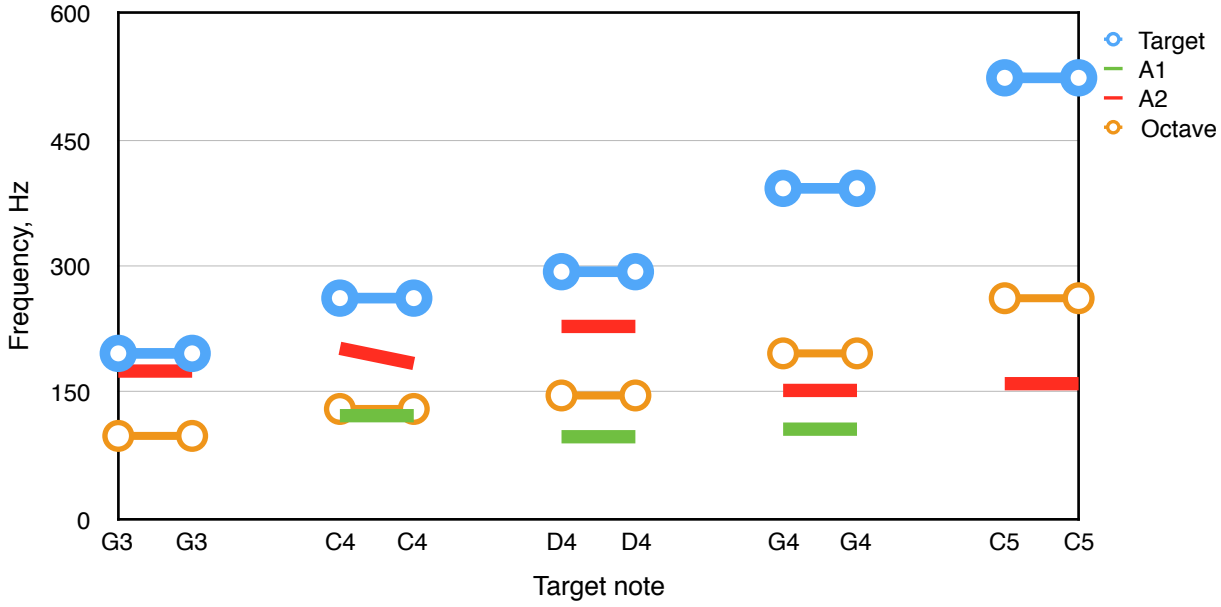
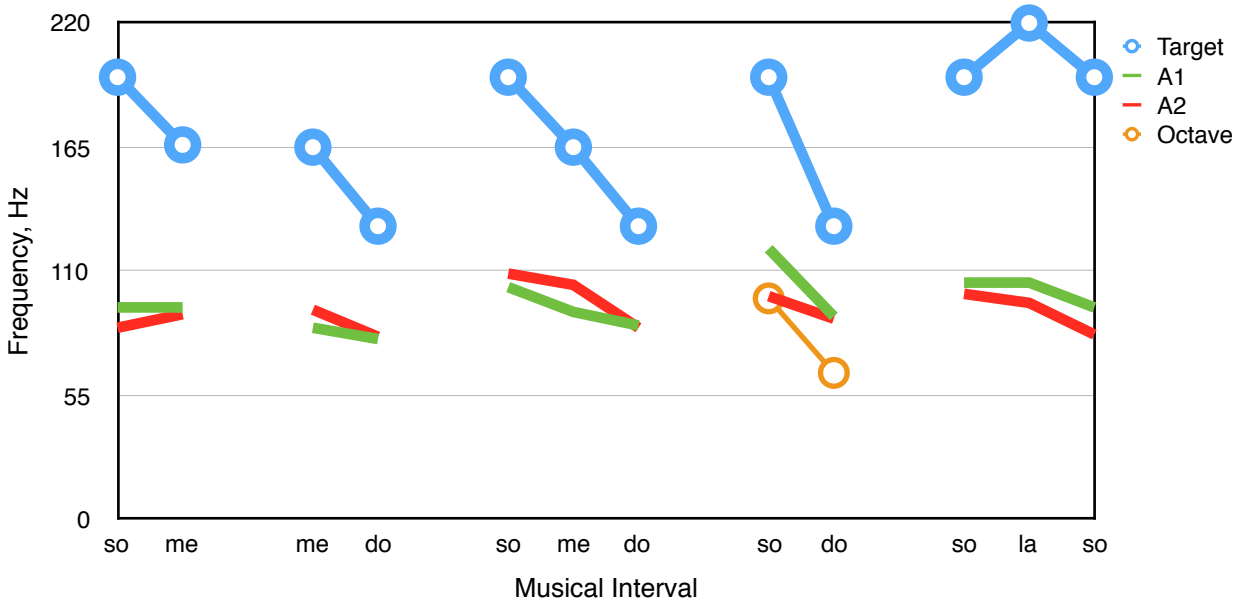


Figure 4.6: Graph showing Andrew's accuracy when matching melodic intervals during his first and final assessments. The fundamental frequency of the target pitch (in Hz) is shown in blue and Andrew's mean F0 is shown for productions in his first assessment (A1) and final assessment (A2), which were six weeks apart. The target pitch is shown in blue and the same note is shown at an octave below, in orange.



Andrew did not match intervals on either assessment (Figure 4.6: A1 and A2). However, he demonstrated a conceptual understanding of pitch direction when reading an ascending and descending scale from visual notation (Figure 4.7). The upward scale had been demonstrated and he was then asked to read the full scale independently. Despite no pitch movement on the upward scale between notes 2 and 4, Andrew produced a descending scale from the fifth note (Figure 4.7, Tier 1: notes 5-9).

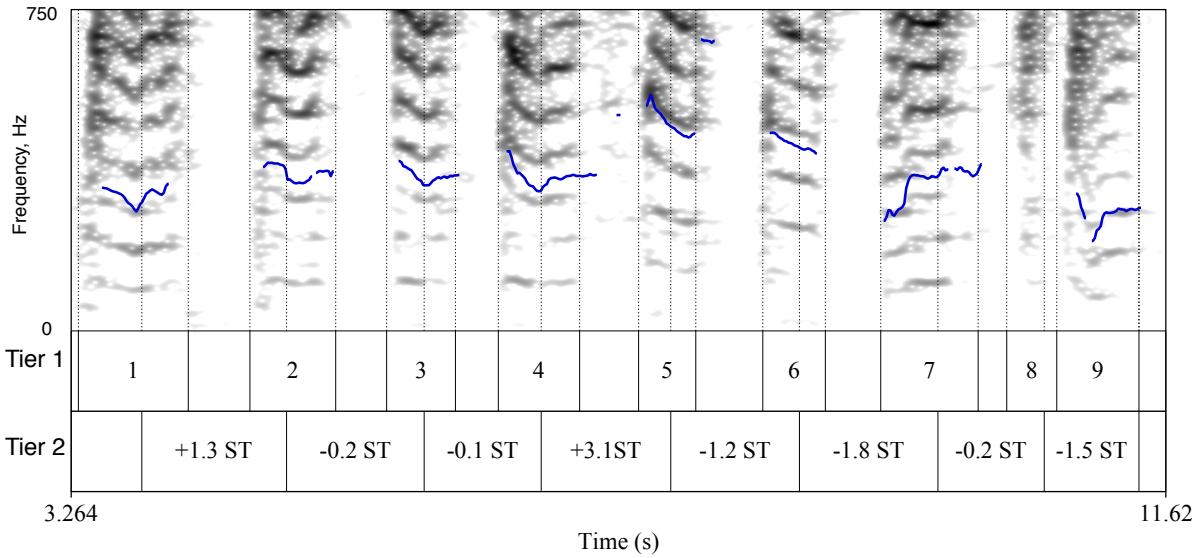


Figure 4.7: Annotated spectrogram showing Andrew singing a musical scale on 'la' from a visual score. The image shows the spectrogram (0-750 Hz; window-length 150 ms, 40 dB) and pitch contour (in blue: settings at 50 dB-250 dB) as Andrew sang from an image that represented a nine-note scale that ascended (notes 1-5) then descended (notes 5-9). The notes are numbered (Tier 1) and Andrew's mean pitch change and direction of change (ascending is +; descending is -) is shown in Tier 2, in semitones. Andrew did not produce phonation in note 8 (Tier 1).

4.3.3 Words of increasing syllable length

Table 4.10 shows prosodic and articulatory features of Andrew's imitations of words of increasing target syllable length. In comparison to the demonstration, as the target syllable length increased from one to three, fewer words were reproduced with the same pitch contour (n= 9, 8, 6) and relative intensity pattern (n= 9, 6, 7). In addition, the number of syllables produced with interharmonics also increased as the target word length increased (n= 3, 6, 18). The intensity of the interharmonics (visible at 15 dB) was greatest in three-syllable words. The number of consonants Andrew produced reduced slightly in longer words but remained similar to the target number of sounds (90%, 89%, 86%), irrespective of accuracy of production, which declined in longer words (Table 4.3).

Table 4.10: A summary of the segmental, prosodic and vocal features of Andrew's imitation of words in the Syllables test, arranged by number of target syllables. The table shows the number of words matched for: pitch contour; changes in relative intensity; and the number of produced consonant sounds, as a percentage of the target. The mean speaking rate (syllables per second) is shown for the demonstration and Andrew's production. The number of syllables affected by interharmonics at 35 dB is shown, and the number is also given for those visible at 15 dB, as a measure of the intensity of interharmonics.

Target no. of syll	Mean speaking rate, syll/ sec	No. of words =	Words matched for pitch contour, no. and percentage, %	Words matched for changes in intensity contour, no. and percentage, %	Consonants produced, no. and percentage, %	Mean speaking rate, syll/ sec	Syllables affected by interharmonics, no. and percentage, % (at 35 dB)	No. and percentage of vowels produced with interharmonics still visible at 15 dB
1	**1.61**	**9**	9/9=100%	9/9=100%	18/20=90%	1.69	5/9=56%	3/9=33%
2	**3.17**	**9**	8/9=89%	6/9=67%	24/27=89%	3.70	11/18=61%	6/18=33%
3	**4.38**	**9**	6/9=67%	7/9=78%	32/37=86%	3.62	24/27=89%	18/27=77%
Mean	**3.05syll/ sec**		**85%**	**81%**	**88%**	**3syll/sec**	**74%**	**50%**

The prosodic and rhythmic features of words of two syllables are shown in Figures 4.8 and 4.9. Andrew's imitation of words of two syllables (Figure 4.8b) were similar to the demonstration (Figure 4.8a) in pitch contour, but they differed in the relative duration of syllables and changes in intensity. Andrew produced additional syllables in *jabber* and *city*, and six of the nine words were produced with different patterns of relative intensity (e.g. *harden:* marked with an arrow, Figure 4.8). Comparison of the syllable boundaries (Tier 2 of Figure 4.8a and b) show some differences in the relative duration of syllables in words: for example, in comparison to the demonstrations, Andrew's first syllable in *soften* was short and his first syllable in *harden* was long.

Andrew's imitated words of three syllables (Figure 4.9b) also differed in prosodic contour, relative intensity and speaking rate (Table 4.10). The perceptual stress of some words was different to the stimulus as a result: for example, *thickening* was produced with a weak initial intensity but greater intensity on the final syllable. Andrew's production of four words overlapped with the final syllable of the demonstration, resulting in additional peaks in the intensity waveform on the diagram (marked by asterisks). Although words of three syllables were produced more slowly than the demonstration (Table 4.10), the duration of initial syllables was reduced when initial consonants were omitted (e.g. *softening*).

Mean F0 and HNR were highest in words of two syllables and lowest in words of three syllables (Table 4.11): SD values of both measures were similar in all three conditions. Mean DUV increased in line with increasing target syllable length. The percentage of vowels

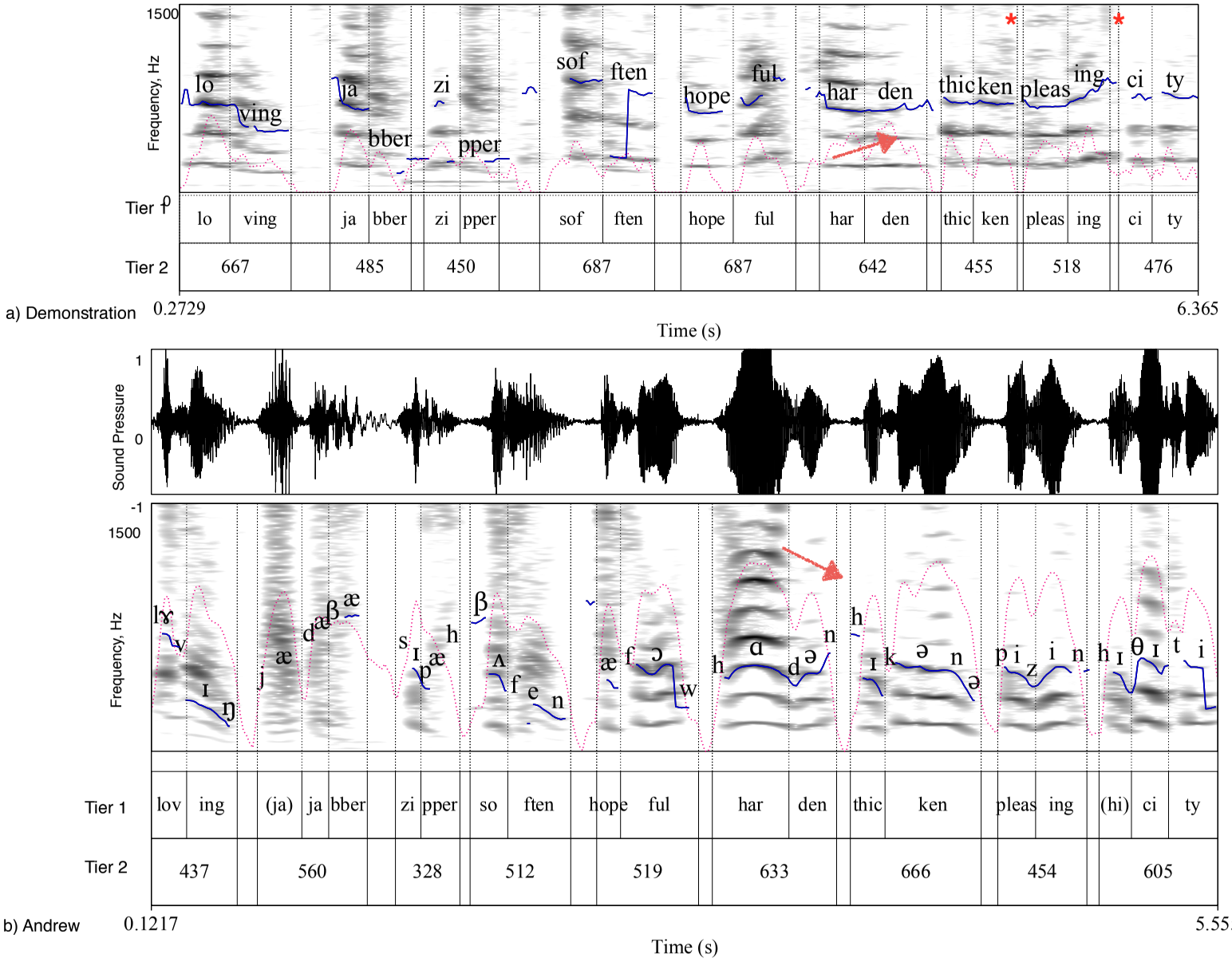


Figure 4.8: Annotated spectrograms showing the stimulus (above) and Andrew's imitation of words of two syllables. Both images show the spectrogram (0-1500 Hz; window length 150 ms, 50 dB) pitch contour (in blue) and intensity contours (in pink). The duration (milliseconds) of each word is shown in Tier 2. The red arrow alongside the intensity contours for 'harden' illustrate differences between Andrew's imitation and the demonstration.

a) Demonstration. The pitch contour is broken by environmental noise but indicates the demonstrated changes. Three red asterisks on the spectrogram indicate where Andrew's imitation overlapped with the demonstration.

b) Andrew's imitation.The image shows phonetic detail on the pitch contour. Two words (*jabber* and *city*) were produced with initial additional syllables, which are included within the total duration.

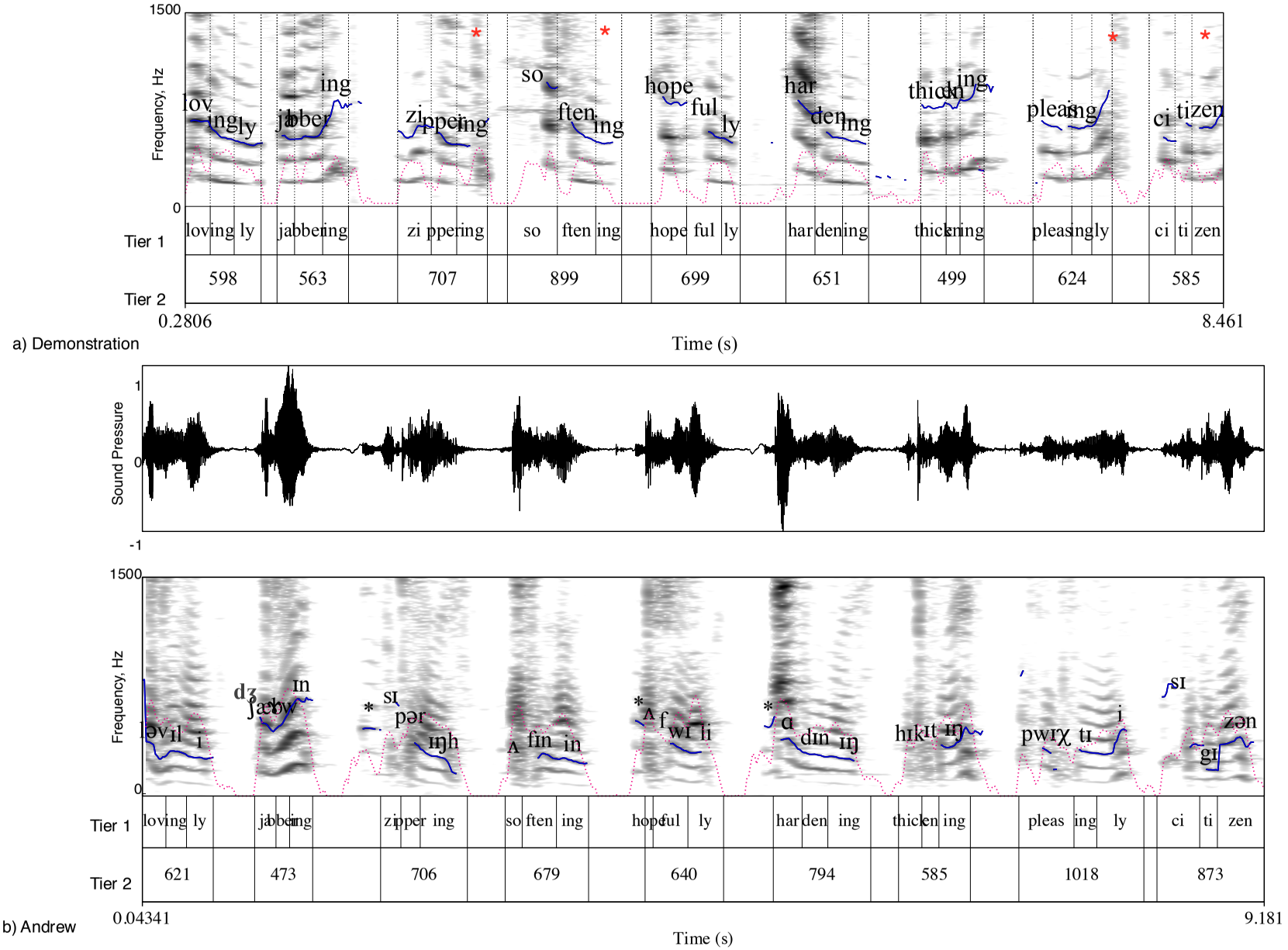


Figure 4.9: Annotated spectrogram showing the stimulus (above) and Andrew's imitation of words of three syllables. Both images show the spectrogram (0-1500 Hz; window length 150 ms, 50 dB), pitch contour (in blue) and intensity contours (in pink). The duration (milliseconds) of each word is shown in Tier 2.

a) Demonstration. The pitch contour indicates the demonstrated changes. Four red asterisks on the spectrogram indicate where Andrew's imitation overlapped with the demonstration.

b) Andrew's imitation. The image shows the waveform and phonetic detail on the pitch contour. Asterisks alongside the transcription show contamination from the demonstration. Andrew's waveforms and the presence of interharmonics indicate ventricular production (e.g. during *zippering and thickening*).

produced with voicing was highest for words of two syllables, which were produced with 33% atypical phonation overall. The percentage of voicing and mean CPP (dB) were lowest in words of three target syllables.

Table 4.11: Voice data for the Syllables test, arranged by the target number of syllables. The table shows mean and SD values for voiced vowels only for F0 (Hz), mean DUV (%) and mean HNR, (dB). CPP (dB) is shown for all vowels.

No. of target syllables	No. of target vowels	No. and percentage of vowels produced with voicing	Mean F0, Hz (SD)	Mean DUV, % (SD)	Mean HNR, dB (SD)	No. of valid vowels	CPP, dB (SD)
1	**9**	5 (55%)	126.72 (34.64)	23.28 (26.14)	8.19 (5.39)	9	3.10 (1.73)
2	**18**	12 (67%)	132.87 (35.58)	29.51 (34.57)	10.97 (6.04)	18	3.35 (1.96)
3	**27**	8 (29 %)	105.47 (29.82)	33.08 (36.35)	7.02 (4.51)	27	1.90 (1.21)
Mean		**n/a**	**121.7**	**28.62**	**8.73**	**n/a**	**2.78**

The voice data are shown according to the position of the vowel in the word (Table 4.12). This confirms that three-syllable words were produced with descending pitch and two-syllable words with ascending pitch. It also shows that DUV, HNR and CPP values change according to the position of the vowel and target length of the word.

Table 4.12: Mean voice data for Andrew's Syllables test, arranged by the position of the vowel in the word (vertical), and according to the number of target syllables in the word (horizontal). The table shows mean data for DUV (%), DVB (%), HNR (dB), and CPP (dB).

Position of vowel in the word (vertical) against number of target syllables n=		Mean F0, Hz			Mean DUV, %			Mean HNR, dB			CPP, dB			
		1	2	3	1	2	3	1	2	3	n=	1	2	3
1st	**25**	126.72	123.59	120.88	23.28	31.16	37.43	8.19	8.58	6.05	27	3.10	3.19	2.13
2nd	**12**		144.79	102.46		27.85	45.75		14.03	8.17	18		3.50	1.88
3rd	**8**			96.18			16.05			6.88	9			1.68

Example: 'Jab, jabber, jabbering'

This set was chosen because Andrew produced no modal phonation in any word. Accordingly, the set represents his most atypical production of the task, and cannot be represented by periodic voice measures. CPP values were 1.15, 1.17, and 1.09 dB for words of one, two and three syllables, respectively. Perceptually, all three words were creaky and breathy in quality. Figures 4.10-4.12 show that the waveforms of all vowels in all

productions were produced with turbulence and contained aperiodic patterns. Andrew's segmental production was inaccurate in all productions and variable: for example, the voiced bilabial stop was unvoiced in 'jab', was produced as a glide in 'jabber', and was produced accurately in 'jabbering'.

Jab: [ʒdæb]. Narrow transcription: [{ ʒ̬!! ʒdæb ʒ̬!!}]

Andrew's production of the vowel was perceptually harsh, unvoiced and whispered in quality (Figure 4.10). The waveform is aperiodic but there is a high intensity periodic production in section 2 and low-frequency periodicity in section 3 (Figure 4.10), consistent with ventricular production. The initial segment is a devoiced postalveolar fricative [ʒ̥] which is followed by a devoiced alveolar plosive [d̥] (section 1). The final bilabial plosive [b̥] (section 4) is also devoiced.

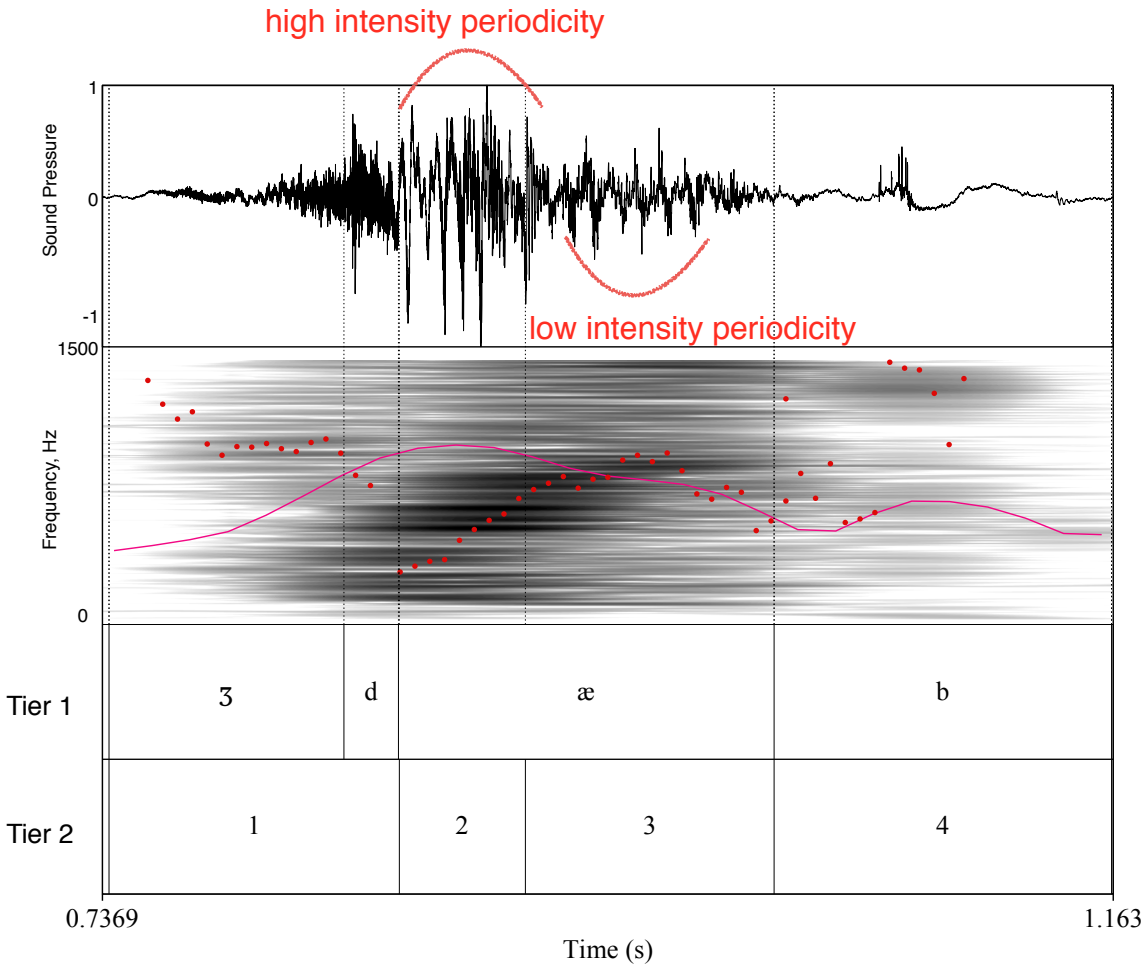


Figure 4.10: Waveform and annotated spectrogram showing Andrew's imitation of 'jab'. The image shows the waveform, spectrogram (0-1500 Hz; window length 150 ms; 40 dB), and phonetic transcription. The image is overlaid with the intensity contour (in pink: 40-100 dB) and the contour of the first formant frequency (red dots). There is no pitch contour as Andrew did not produce phonation. However, there are two types of periodic vocal production evident on the waveform (as marked). Tier 1 shows the phonetic production and Tier 2 defines sections that are discussed in the text.

Jabber: [**ja̱**d̲æβæh]. Narrow transcription: { Ɣ̊!! ja̱ dæβæh Ɣ̊!!}]

The word was produced without phonation but with a perceptually harsh and whispered quality, and with three syllables, the first of which [ja̱] (section 1) is possibly dysfluent production. The first syllable was perceptually stressed and produced with higher intensity, as shown in the waveform, spectrogram and intensity contour (in pink) (Figure 4.11: section 1). The waveform shows a low-frequency periodic pulse in section 1 (as marked). A lower intensity pulse is also apparent in the vowel of the second syllable, which is partially voiced (section 3).

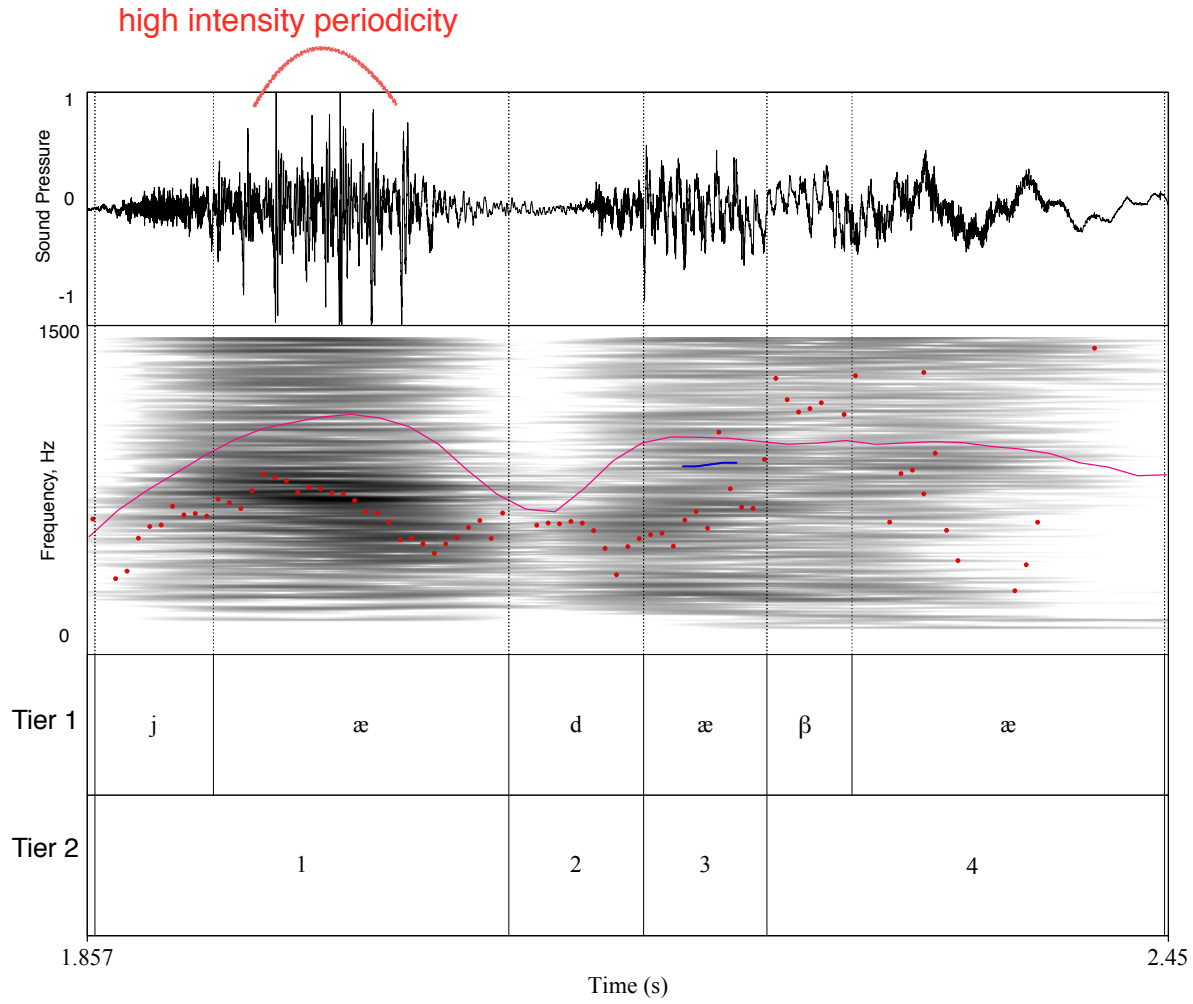


Figure 4.11: Waveform and annotated spectrogram showing Andrew's imitation of 'jabber'. The image shows the waveform, spectrogram (0-1500 Hz; window length 150 ms, 40 dB) and broad phonetic transcription of the word 'jabber' (Tier 1). The intensity contour (40-100 dB) is marked in pink, and the contour of the first formant frequency is shown in red dots. Points of interest that are discussed in the text are marked in Tier 3.

Jabbering: [**dʒæ**bəwɪŋ̊]. Narrow transcription: [{ Ɣ̊!! dʒæbəwɪŋ̊ Ɣ̊!!}]

The first syllable was perceptually stressed: the first vowel is relatively longer in duration than the second and was produced with increased intensity (Figure 4.12; section 1). The

second and third syllables were produced with some periodicity at a low frequency: examples are indicated in section 3 and at the start of section 4 (Figure 4.12). Darker striation lines are visible on the spectrogram at these points.

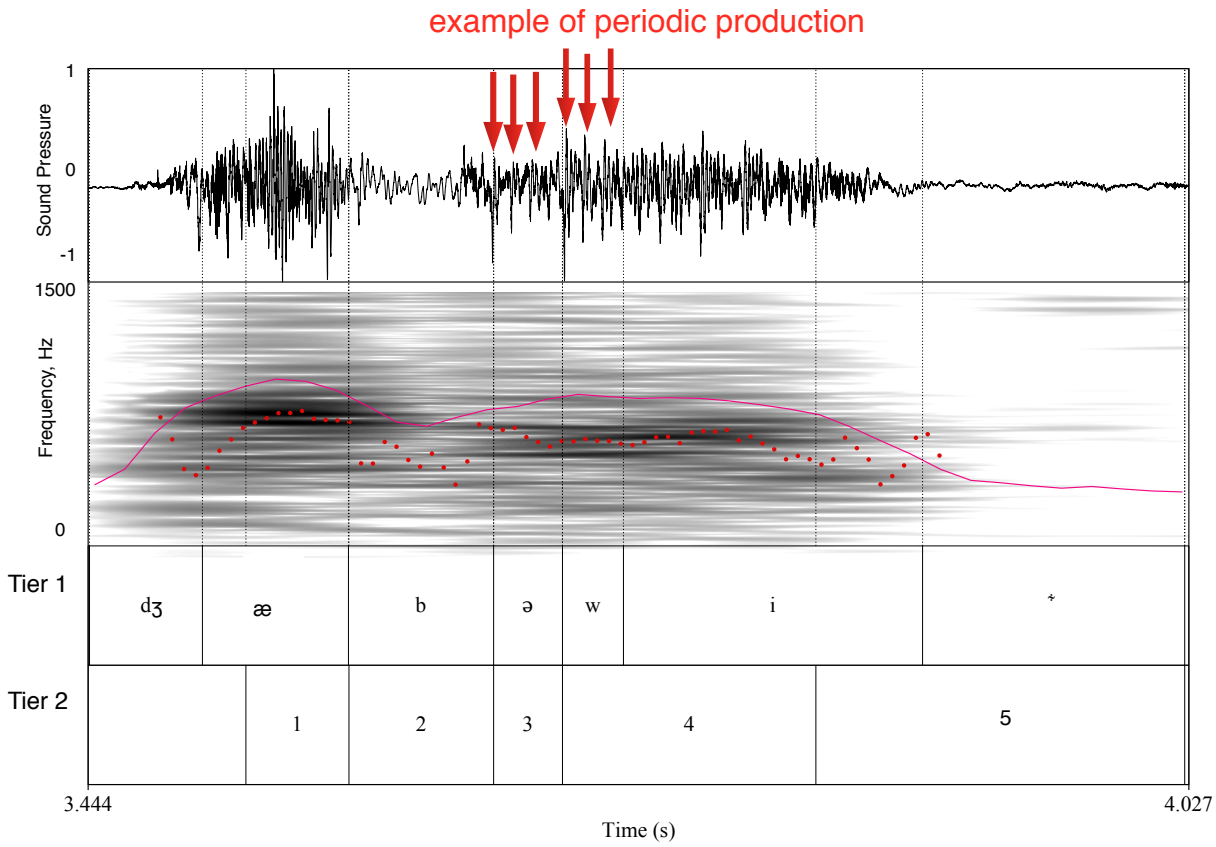


Figure 4.12: Waveform and annotated spectrogram showing Andrew's imitation of 'jabbering'. The image shows the waveform, spectrogram (0-1500 Hz; window length 150 ms, 40 dB) and broad phonetic transcription of the word 'jabbering' (Tier 1). The intensity contour (40-100 dB) is marked in pink, and the contour of the first formant frequency is shown in red dots. There is no pitch contour, but periodic production is visible in the waveform in sections 3 and 4: examples are marked by arrows. Points of interest that are discussed in the text are marked in Tier 3.

4.3.4 Nonword imitation

Andrew imitated nonsense words with appropriate pitch contours and reproduced the number of syllables in all but three items (Table 4.13; *blonterstaping*, *defermication*, *reutterpation*). He produced 79 of a possible 98 sounds (81%: Table 4.13) but the accuracy of these was just 46% (see Table 4.3). His speaking rate was faster (5.0 syllables/second) than the demonstration and his production of *blonterstaping* interrupted the demonstration, which affected the number of syllables he was able to reproduce. His pitch range was almost half that of the demonstration, but he reproduced the pitch contours of eight words. Andrew produced all items with multiple interharmonics (e.g. Figure 4.10-4.12): Table 4.13 shows the percentage of these that still remained visible when the intensity of the signal was reduced to 15 dB.

Table 4.13: Prosodic features of the demonstration and of Andrew's imitation of non words (CNREP). The table shows the perceived stress, the pitch range (semitones) and the duration (milliseconds) of the demonstrated word. Andrew's production of the each target word is shown with a phonetic transcription, and the number of words produced with the same (S) or different (D) pitch contour is indicated. The table also shows mean number of sounds and syllables reproduced, the mean ST range, and the speaking rates. All words were produced with subharmonics but the percentage of subharmonics that remain visible at a threshold of 15 dB is also shown, as an indicator of intensity of atypical phonation. ** Andrew interrupted the demonstration of blonterstaping.

Demonstration			Accuracy of imitation						Percentage of vowel produced with interharmonics visible at 15 dB
Target word and stress	Pitch range, ST	Duration, ms	Production (IPA transcription)	No. of syllables produced	No. of sounds produced	Pitch range, ST	Pitch contour	Duration, ms	
Hampent	11	641	**æ**bənt	2/2	5/7	5	S	501	50%
Pennel	7.8	589	**pɛ**nəl	2/2	5/5	11	S	622	50%
Doppelate	8	437	**dʌ**bəwɛt	3/3	7/7	2.5	S	521	60%
Glistering	10	1065	**gwɪ**tjəwiŋ	3/3	8/9	3.7	S	903	100%
Blonterstaping	12	1049	**bwə**sti**	2/4	6/13	5.5	S	667	0%
Cont**ram**ponist	12	783	**gə**jæɸɪmɪs	4/4	9/13	3	D	486	0%
Per**plis**teronk	15	987	hɪb**wɪz**dɪwʌnt	4/4	12/12	5.1	S	716	0%
Defer**mica**tion	8	918	**su**wɪxeɪʃən	4/5	9/11	3	S	671	85%
Reutter**pat**ion	9	701	**ʌd**ipætɪn	4/5	8/10	8	D	670	67%
Sepretennial	7	1052	**se**wɤ**te**wiew	5/5	10/11	4	S	826	85%
Mean	**10**	**4.25 syll/ sec**	**47/60 speech sounds**	**33/37**	**79/98**	**5.1**	**8/10**	**5.01 syll/ sec**	**55%**

Mean F0 and DUV were similar for nonsense words of two, three and five syllables but those of four syllables were produced with a higher mean F0 and fewer unvoiced (DUV) sections (Table 4.14). HNR was highest in the longest words and lowest in words of three syllables; CPP was also lowest in words of three syllables. Within words, vowels in final syllables were produced with the highest number of perturbations and lowest mean F0 (Table 4.15). Examples were chosen for illustration that demonstrated the most accurate and least accurate imitations of longer nonwords (Figure 4.13 and 4.14, respectively). The perceived prosody of Andrew's production of *perplisteronk,* the relative duration of syllables, and the

Table 4.14: Mean voice perturbation measures of Andrew's vowels in CNREP, arranged according to the number of syllables in the target word. The table shows the mean measures of F0 (Hz), DUV (%), HNR (dB), CPP (dB), and the percentage of subharmonics that remain visible at 15 dB.

No. of syllables	No. of words: n=	Mean F0, Hz	DUV, %	HNR, dB	CPP, dB	Percentage of vowel with Subharmonics visible at 15 dB, %
2	**2**	92.56	31.22	4.83	2.08	50.00
3	**2**	87.87	37.21	3.24	1.63	80.00
4	**3**	155.15	26.91	4.61	2.23	0.00
5	**3**	98.29	30.18	5.55	2.20	79.00

pitch contour were similar to the demonstration (Figure 4.13: Tier 2). In contrast, *contramponist* was produced with a different pitch contour, intensity contour, rhythm (Figure 4.14) and perceived stress (Table 4.13).

Table 4.15: Mean voice perturbation measures of Andrew's vowels in CNREP, arranged according to the position of the syllable in the word. The table shows the mean measures and SD of F0 (Hz), DUV (%), HNR (dB), and CPP (dB). Nb. mean data are from voiced vowels only for all values except CPP, which was available for all vowels.

Position of syllable	Voiced vowels, n=	Unvoiced vowels, n=	Mean F0, Hz (SD)	Mean DUV, % (SD)	Mean HNR, dB (SD)	CPP, dB (SD)
First	**7**	**3**	132.36 (92.12)	27.32 (34.16)	7.38 (2.48)	2.54 (1.04)
Medial	**7**	**3**	132.21 (52.44)	29.47 (22.32)	4.97 (2.37)	2.58 (1.54)
Final	**10**	**4**	91.47 (51.56)	33.48 (28.65)	2.97 (1.26)	1.50 (0.83)

4.3.5 Imitation of phrases

Andrew omitted syllables and words from each phrase, affecting their segmental accuracy and rhythm. The rhythm of phrases 1 and 2 (Table 4.16) were affected by dysfluency. Andrew substituted words in four phrases, adding *all* to phrase 4, replacing 'Mary' with *she*, and adding *him* to phrase 5. Prosodic imitation was affected by voice difficulties and differences in pitch contours: phrases 1, 3 and 4 were produced with ventricular phonation; and phrases 1-4 were produced with a rising intonation, in contrast to the descending intonation of the stimuli. Phrase 5 was chosen as an example. The phrase was produced with the widest pitch range (10 ST) and near-modal sounding voice quality and without hesitation (Table 4.16). As such, it may represent Andrew's prosodic ability when his performance is unaffected by segmental or phonatory difficulties.

Table 4.16: Prosodic features of the demonstration and of Andrew's imitation of phrases (Newton, 1999). The table shows the target phrase, which is marked with syllabic stress, and the pitch range (ST), duration (s) and rate (syllables/second) of the demonstration. Primary stress is indicated in bold type: secondary stress is underlined. The similarity of the rhythm, pitch range (ST) and pitch contour is indicated by S (same) and D (different). The perceived voice quality is indicated by EXT IPA symbols for ventricular phonation (V!!).

Demonstration: words and stress		Pitch range, ST	Duration, s	Rate, syll/sec	Andrew: words and stress (IPA transcription)	Pitch range, ST	Pitch contour	Duration, s	Rate, syll/sec
1	You must clean your **teeth**	10.50	1.36	3.62	{V!! hm m **mʌst**ə kɬi jɪ tif V!!}	4	D-rise	2.06	1.94
2	Mary's shoes are **clean**	8.60	1.26	2.70	ʃɪz ʃɤz ʃʊz ɑ **klin**	2*	D - rise	1.65	3.03
3	I gave the elephant a ba**na**na	13.20	1.68	4.88	{V!! ge **eə**fən θʌ wɪn**ɑn**ə V!!}	8	D - rise	1.34	5.97
4	**Clare** ate all her lunch	7.80	1.57	3.11	{V!! **ki**e **ʌ** ɔ hə **lʌntʃ** V!!}	3	D- rise	1.14	4.39
5	My uncle is a **far**mer	10.01	1.38	5.30	ɛχ **ɪm b**i ə fɑmə	10	D	0.85	7.06
Mean		**10.02**	**1.45**	**3.92**	**84% of words**	**5.4**	**n/a**	**1.41**	**4.48**

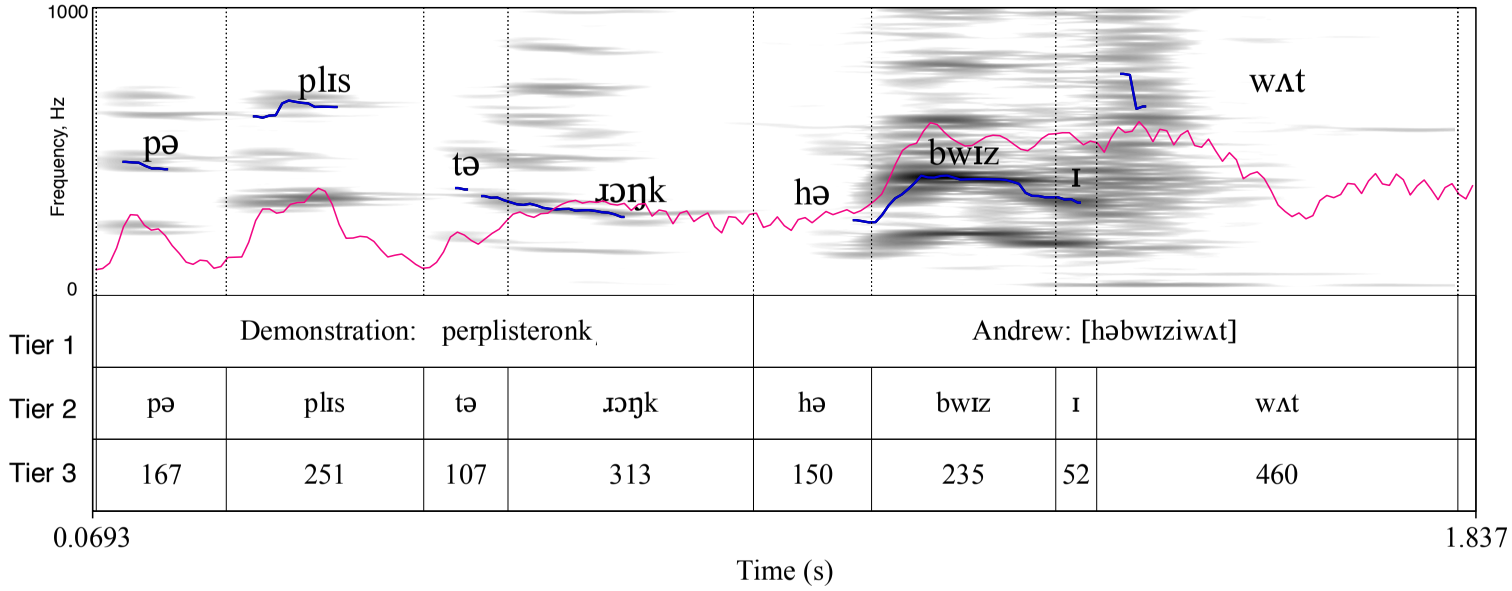


Figure 4.13: Annotated spectrogram showing Andrew's imitation of perplisteronk' (CNREP). The image shows the spectrogram (0-1000 Hz; window length 150 ms; 40 dB) pitch contour (in blue: 90-400 Hz)) the intensity contour (in pink: 40-100 dB). The phonetic productions of the stimulus word (on the left) and Andrew's imitation (right) are shown alongside the pitch contour and in Tiers 1 and 2. The division of syllables is shown in Tier 2 and the duration of these (in milliseconds) are shown in Tier 2.

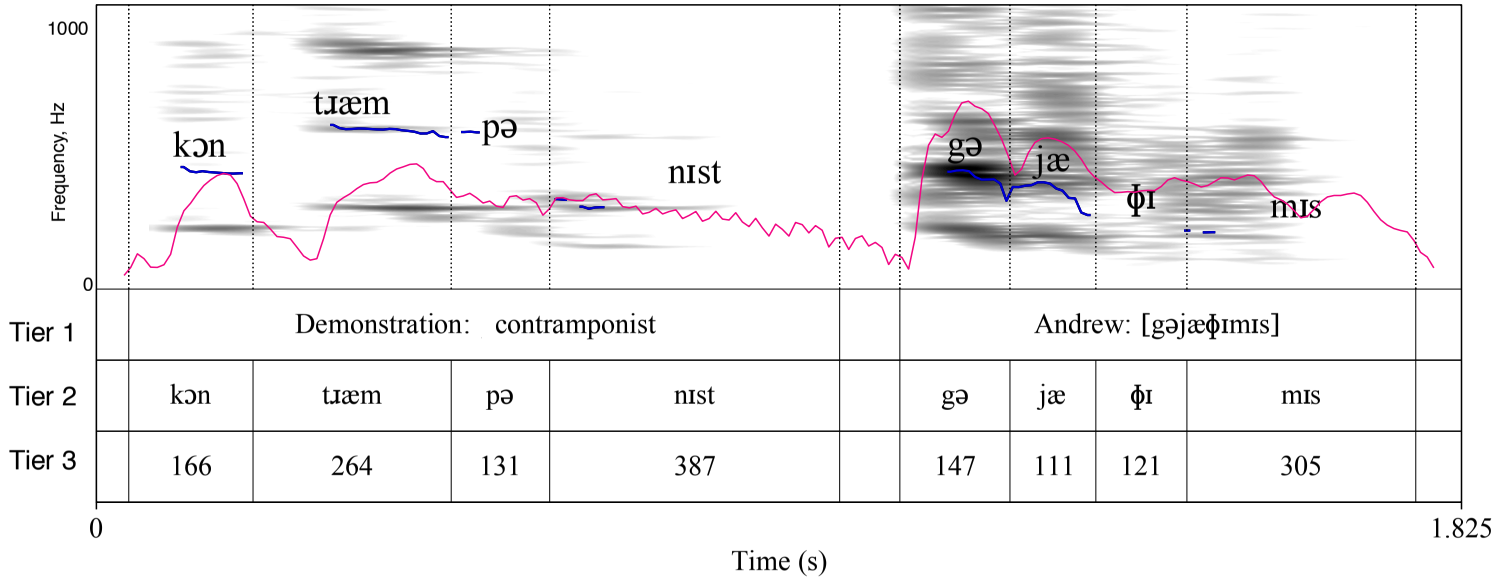


Figure 4.14: Annotated spectrogram showing Andrew's imitation of 'contramponist' (CNREP). The image shows the spectrogram (0-1000 Hz; window length 150 ms; 40 dB) pitch contour (in blue: 90-400 Hz)) the intensity contour (in pink: 40-100 dB). The phonetic productions of the stimulus word (on the left) and Andrew's imitation (right) are shown alongside the pitch contour and in Tiers 1 and 2. The division of syllables is shown in Tier 2 and the duration of these (in milliseconds) are shown in Tier 2.

Example: My uncle is a **far**mer: [ɛχ **ɪm bi** ə fɑ̰mə̰]

The intonation and perceptual stress of Andrew's imitation is different from the demonstration. There are differences in the changes of pitch and relative intensity at the boundary between 'a' and 'farmer' (Figure 4.15). However, Andrew reproduced the relative rhythm of the word *farmer,* and the descending pitch contour, despite Andrew's loss of phonation on the last syllable.

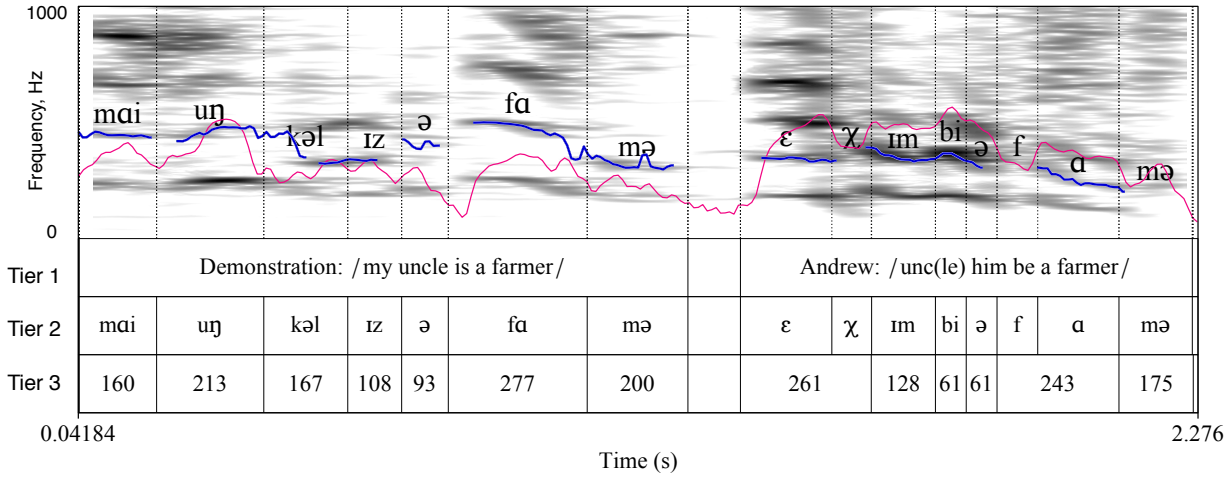


Figure 4.15: Annotated spectrograms showing Andrew's imitation of 'My uncle is a farmer'. The image shows the spectrogram (0-1000 Hz; window length 150 ms; 40 dB) pitch contour (in blue: 50-400 Hz) the intensity contour (in pink: 40-100 dB). The phonetic production of the stimulus phrase (on the left) and Andrew's imitation (right) is shown alongside the pitch contour and in Tier 2. The syllables duration of syllables (in milliseconds) are shown in Tier 3. Note: the spectrogram does not align with the pitch and intensity contours as a result of the narrow-band spectrogram and timeframe (Boersma and Weenink, 2011).

4.3.6 Song imitation: Magic Book

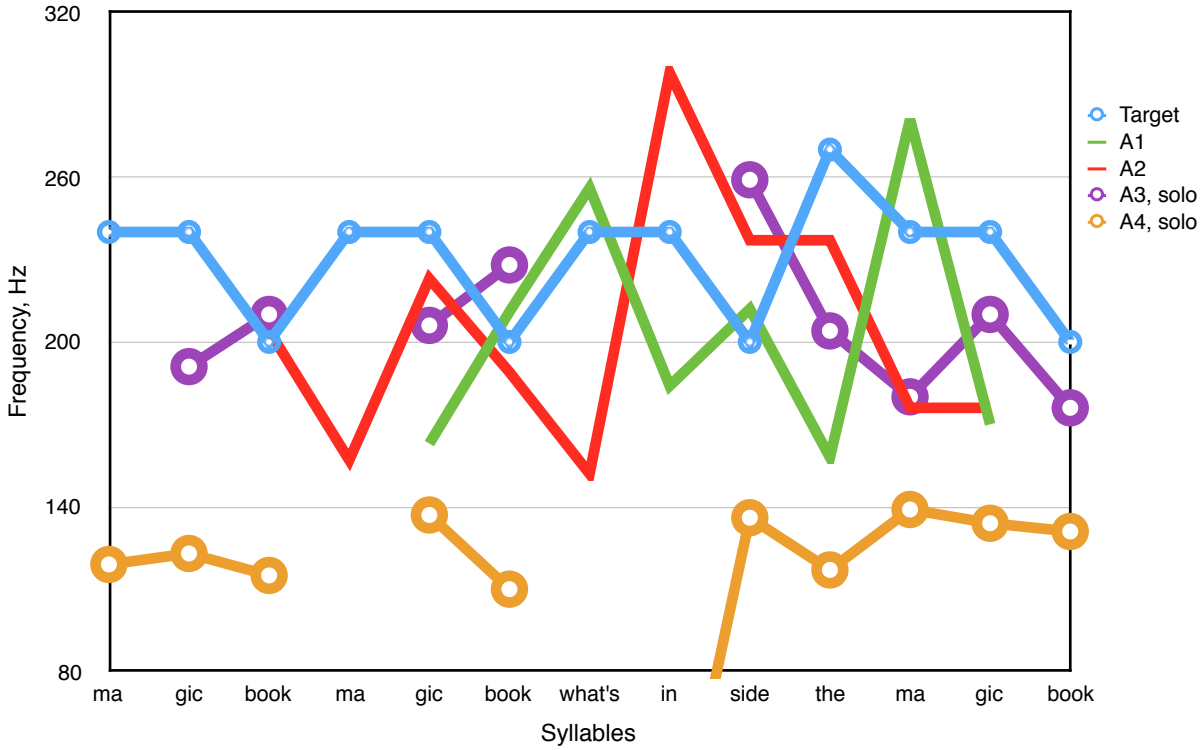
Andrew required prompts to produce the words in his first solo production of the song. His production of syllable stress was different to the target (Figure 4.16). In his final production (week 5), he was able to produce all the words without prompts but he omitted syllables (lines 2 and 5). Some syllables were different to the target production in perceived stress. Andrew's voice quality was ventricular in both productions, but his final production was also strained in quality and he displayed symptoms of stress (teeth grinding).

Target	Andrew: week 1(prompted)	Andrew: week 5 (solo)
1. **ma**gic **boo**k	1. (*magic*) {V!! bɤk	1. {V!! **bæzɪ bɤk**
2. **ma**gic **book**	2. (*ma*) dʒ**ɪ** pɤk	2. **tɪ** bɤ
3. **what**'s in**side** the	3. (*wh*)**ɔ i**sai də	3. **ɪsai ðæ**
4. **ma**gic **book**	4. mædɪ bɤ	4. mæð**ɪ bɤk**
5. **ab**raca**da**bra	5. æ**β**ə**d**ɪ**dæ**bwə V!!}	5. **æ**bwa**dæ**bwæ V!!}

Figure 4.16: Phonetic transcription and syllabic stress of Andrew's first and final production of 'Magic Book', in week one (prompted) and week five (solo). The figure shows the phonetic detail of Andrew's productions in the first and final production of the song. Prompts are indicated in brackets. Stressed syllables are highlighted in bold text. His vocal production was ventricular in quality throughout as indicated by the VoQs transcription (V!!).

Andrew's melodic production improved as he recalled more of the words (Figure 4.17: compare A2 with A1). In his second production (A2), he recalled most words and his melodic shape (red contour) is similar to the target (blue contour). Andrew did not recall all words in his final, unprompted production (A4) and his melodic production was least accurate: however, the mean F0 of this production was low and the initial syllable of *inside* was produced with whispered phonation, which affects the contour.

Figure 4.17: Graph showing the development in Andrew's recall of the words and melody of 'Magic Book'. The figure shows the melodic contour of Andrew's performance of the song, over a five-week period. A1 was his production in week 1, and his final production was A4 (week 5). The target contour is shown in blue.



The articulatory, rhythmic and melodic changes measured between Andrew's first and final production of the first phrase of *Magic Book* is shown in Table 4.17. The number of sounds he produced was higher in his final production and most sounds were close to the target production. Andrew's articulatory speed increased in the final production and his C:V ratio increased. The melodic features showed no improvement, but the relative duration of syllables was closer to target in the final production.

Table 4.17: Articulatory and musical features of Andrew's production of the first phrase of 'Magic Book'. The table shows segmental, articulatory, melodic and rhythmic features of the demonstration, and Andrew's initial and final productions of the first phrase in the song ('magic book, magic book'). Segmental production details the number of consonant sounds (C) Andrew produced and their relevance to the target sound in terms of manner and place. Articulatory features measure the rate of production of consonants and vowels and the relative duration of these. The mean change in semitones between syllables is given (melodic features). Direction is indicated (downwards (-), upwards (+): the pitch change in bold text highlights pitch reproduction that was accurate within 0.5 ST. The table also shows the mean duration and relative duration of syllables in the words (rhythmic features).

Demonstration: key features			**First production**	**Final production**
Segmental production	**Number of C sounds**	**10**	7	7
	Number of relevant sounds	**10**	7	7
Articulatory features	**Total duration C. ms**	**856**	822	688
	Articulatory rate: sounds per second	**11.8**	8.5	10
	Total duration V, ms	**1577**	1033	1051
	Average duration of C:V	**100:184**	100:126	100:153
Melodic features	***Magic*: Mean ST change**	**0**	n/a	0
	***Book*: Mean ST change**	**-3**	**-2.6**	-1.2
	***Magic*: Mean ST change**	**0**	0.8	n/a
	***Book*: Mean ST change**	**3**	-2	-3.9
Rhythmic features	**Mean duration of *magic*, s**	**398**	685	587
	Mean duration of *book*, ms	**797**	338	497
	Relative duration of syllables *ma:gic*	**1:1**	4:5	1:1
	Relative duration of the words *magic:book*	**1:2**	2:1	4:5

4.4 Prosody and voice in production of speech and song from LTM

4.4.1 Comments

Short, spontaneous phrases are shown in Figure 4.18: these were produced between tasks during teaching sessions. Andrew's comment on Kerry's (see Chapter 5) singing (Comment 1) is shown in Figure 4.19 and points of interest are marked in Tier 3. With the exception of

Comment 1: I think it is the **swee**test song, **ev**er heard
[{V̰‼ æ̰ fɪn̰ˇ s..z.. zɪ.. B.. ðə.. ðɪ.. ðəi...(0.39s).. ɪ ɪ sui ...(0.86s)..... ʃuitɪ‼s {V̰‼ sɔŋə s(ɔn sɔŋ V̰‼} {V̰‼ sɔ̰ŋ V̰‼} ...(1.56s).....əuə (ɛ:d]

Comment 2: **Well** that's **bri**lliant that **is.** [{V‼ wæ̰ θʔæ̰ bwɪnə (æ) ɪz V‼}]

Comment 3: **O**k, I **will**: here **goes.** [uge‼ ʌwə jeəʊ‼z]

Figure 4.18: Phonetic transcriptions and syllabic stress of Andrew's spontaneous comments, taken from session recordings. The interpretation indicates the perceived stress and voice quality: bold type indicates syllables that were produced with primary stress; syllables produced with secondary stress are underlined. The VoQs indicate harsh production (!!), ventricular production (V!!) with and without voice (V̰‼), and diplophonic production (V̰!!).

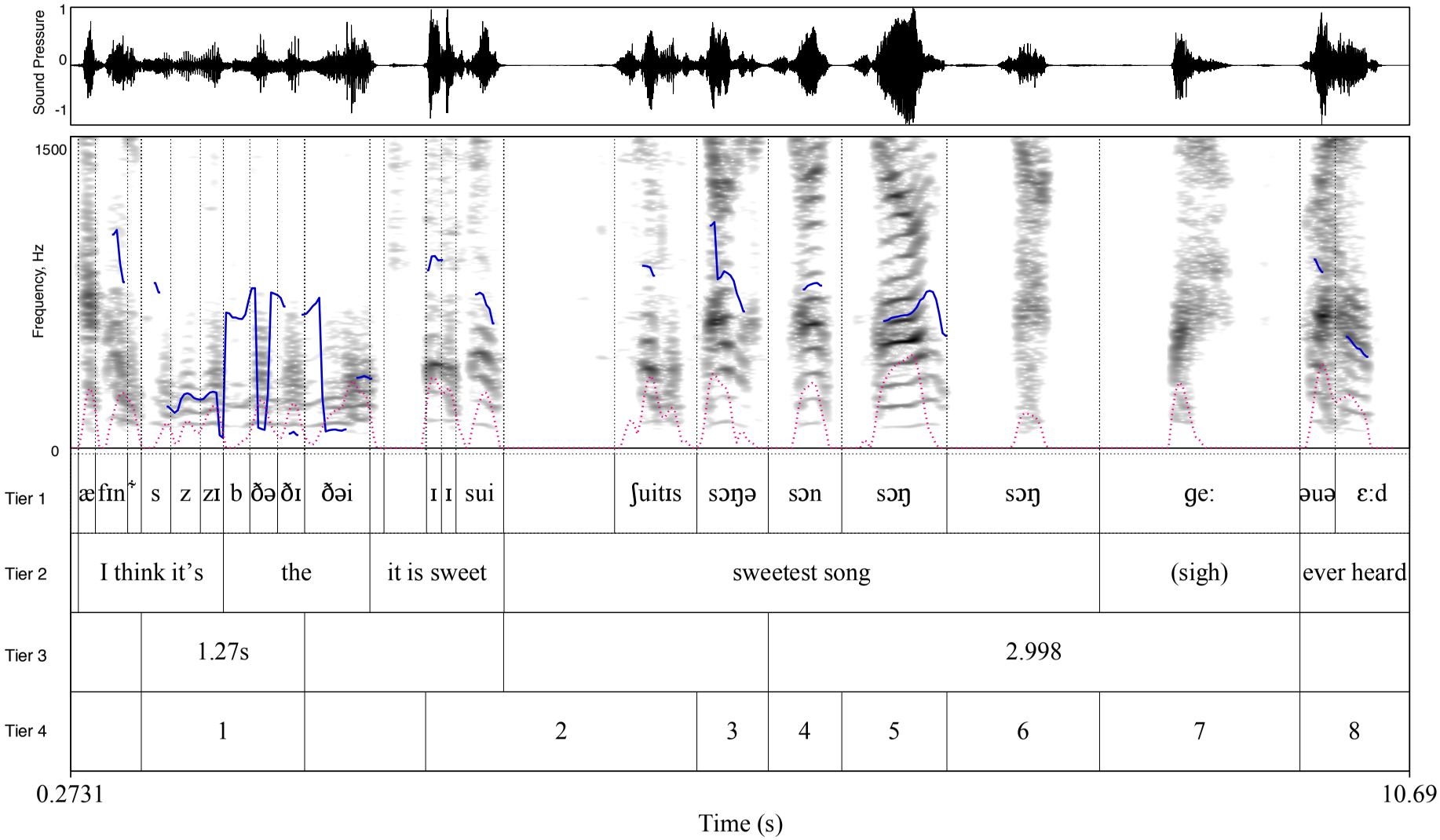


Figure 4.19: Waveform and annotated spectrogram showing Andrew's comment to Kerry (Comment 1). The image shows the narrow band spectrogram (0-1500 Hz; window length 150 ms; 50 dB) pitch contour (in blue) the intensity contour (in pink) and the phonetic production of Andrew's comment for Kerry. The words are shown in Tier 2. There are two instances of stuttering in Section 1 and sections 4-7 (Tier 4): the duration of stuttering is shown (in seconds) on Tier 3. Points of interest (tier 4) are discussed in the text.

the first and final pairs of words, his production was dysfluent (sections 1, 2, and 3-6). Each production of the word 'song' was perceptually different and his phonation was ventricular and diplophonic in three of these (sections 3, 4 and 5). The initial dysfluency (section 1) and the vowels in most of the words were produced with audible diplophonia and ventricular phonation. From section 2 to section 6, the perceived pitch declined (Figure 4.19). No pitch contour is visible as a result of his whispered production (Figure 4.18).

Comments 2 and 3 were produced without dysfluency but with omitted consonants (Figure 4.18). Comment 2 was produced at 4.76 syllables per second, and Comment 3 at 5.50 syllables per second. Despite atypical phonation and some voice breaks, rises in pitch coincided with increased loudness, giving perceptual stress to syllables. The semantic content was inferred from the context and from intonational cues.

4.4.2 Happy Birthday in sung and spoken conditions

Table 4.18 shows the relative duration of syllables and pitch changes in Andrew's sung and spoken productions of the first phrase, *happy birthday to you*. Andrew's production of the phrases in both conditions was slow and he required prompts of the first word and starting pitch in singing. His spoken production contained lengthy pauses between phrases. Both productions echoed the direction of pitch changes of the sung version, but he did not produce the absolute intervals. The pitch change and the relative value of syllable duration of Andrew's spoken version was closer to the expected values when singing.

Table 4.18: The relative duration and pitch of Andrew's syllables in 'Happy Birthday', in sung and spoken conditions. Andrew did not produce the word 'happy' in the sung version. The table shows the mean pitch change in semitones between syllables (ST), and the relative duration of syllables (RD) in comparison to the shortest syllable duration in the phrase. The (+) indicates a rising melodic interval the (-) indicates a falling melodic interval: those in bold text highlight pitch reproduction that was accurate within 0.5 ST.

Target			**Singing**				**Spoken**			
Phrase 1			**Phrase 1**		**Phrase 2**		**Phrase 1**		**Phrase 4**	
Syllable	**Pitch change, ST**	**RD**	**Pitch change, ST**	**RD 1=200ms**	**Pitch change, ST**	**RD 1=200ms**	**Pitch change, ST**	**RD 1=200ms**	**Pitch change, ST**	**RD 1=300ms**
ha	**n/a**	**1**					n/a	1	n/a	1
py	**0**	**1**					n/a	2	**0**	2
birth	**+4**	**2**	n/a	4	n/a	2	+1.3	2	+7.9	2
day	**-4**	**2**	-0.8	5	**-3.9**	2	-2.2	2	-3.3	2
to	**+5**	**2**	+2.2	5	+1.3	3	**+5**	2	0	3
you	**-2**	**4**	-4.6	5	-1.3	5	-4.6	4	-4.3	5
Mean speaking rate (syllables per second)			**0.98**		**0.94**		**1.43**		**0.81**	
Pitch range (max-min ST)			**5.3**				**7.3**			

A transcription with voice symbols is shown for Andrew's sung and spoken productions of the words to the song (Figure 4.20). The perceived stress of syllables in singing was similar to target, but Andrew produced fewer words and phrases. Andrew's voice quality when singing was perceptually harsh (Figure 4.20) and interharmonics were present throughout. Andrew recalled more words when speaking, and he produced some periods of modal phonation (Figure 4.20).

Target	Sung	Spoken
1: happy **birth**day **to you**	1: {V! **bəh**dei **tʃu** ju	2: æ̥**pi̱ bəf**dei **tu ju**
2: happy **birth**day **to you**	2: **pəf**dei tu **ju** V!}	3: hæ**pi bə̱f**dei… ti**ə:**
3: happy **birth**day dear…		**4.**ɦ {V!!æ̥V!!} {V!**pi**V!!}{V!!
4: happy **birth**day **to you**	*interharmonics 100%*	**bə**V!!**}f**dei **tʃu̱ ju̱u**
		interharmonics: 73%

Figure 4.20: Phonetic transcription and syllabic stress of Andrew's sung and spoken productions of 'Happy Birthday'. The phonetic detail includes an indication of the perceived syllable stress and of voice characteristics. The interpretation indicates the perceived stress and voice quality: bold type indicates syllables that were produced with primary stress; syllables produced with secondary stress are underlined; and the extIPA symbols indicate voiceless production [æ̥] and creaky production [u̱]. The VoQs symbols show harsh production (!!) and ventricular production (V!!). The atypical voicing as indicated by the presence of interharmonics is also shown, expressed as a percentage of the total vowel duration.

Detail of his spoken production of phrase 4 is shown in Figure 4.21, and further detail is given in Figure 4.22. The fourth phrase followed a silent period in which Andrew was grinding his teeth. As the quality of the waveform and the breaks in the pitch contour show (Figure 4.21), Andrew's phonation was changeable throughout the phrase. The points marked on Tier 3 show the points at which subharmonics are clearly visible (see also Figure 4.21) and which coincide with perceptual changes in voice quality. Figure 4.22a) shows that the offset of syllable two in *happy* (point 1) is voiced but diplophonic: a single interharmonic is visible between the two lower main harmonics. The second word *birthday* is produced with ventricular phonation and multiple interharmonics initially (Figure 4.22b), but the offset is produced with periodic phonation and is perceptually modal. The word *to* has two instances of diplophonic production (sections 4 and 5), which coincide with subharmonics (Figure 4.22c). The word *you* is initially diplophonic (Figure 4.21, section 6): subharmonics at the onset of Andrew's diphthong [iu] that are almost equal in intensity to the main harmonics (Figure 4.22c). Andrew's intensity contours stay mostly stable during these changes in phonation, but increases in the modal sections.

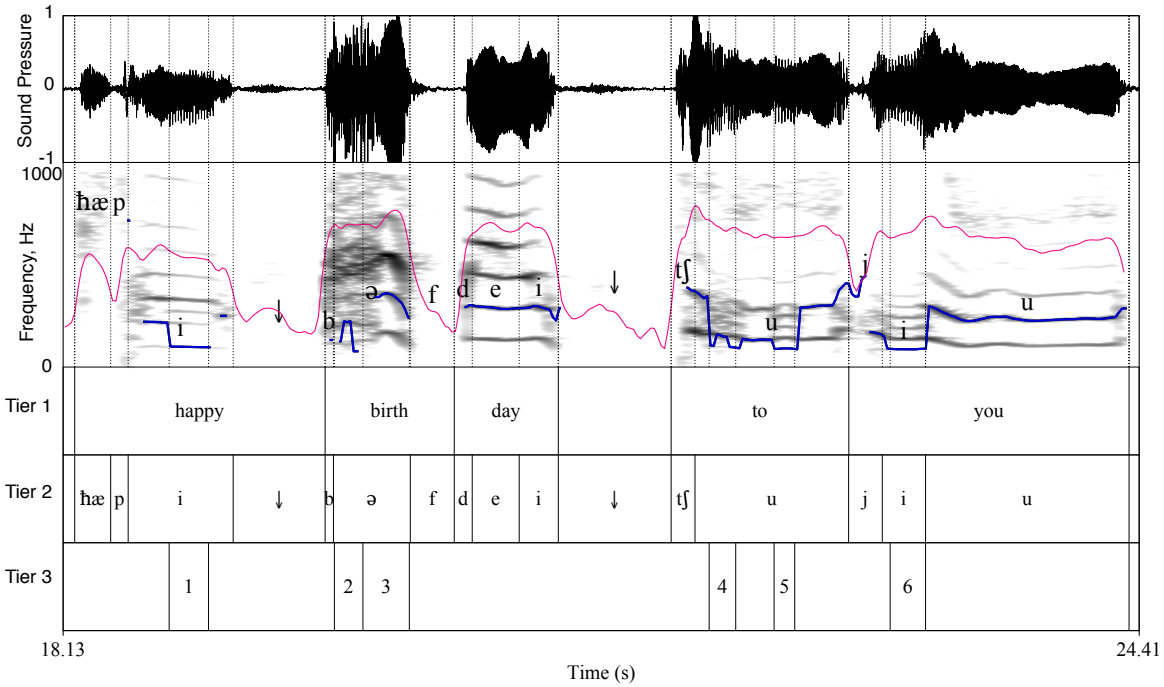


Figure 4.21: Waveform and annotated spectrogram showing phrase 4 of Andrew's spoken production of the song 'Happy Birthday'. The image shows the waveform (above) and narrow-band spectrogram (window length 0.15 ms, 45 dB, 0-1000 Hz) which is marked with the pitch contour (blue: 40-400 Hz) and intensity contour (pink: 40-100 dB). The tiers below show the words. Andrew produced (1), his phonetic production (2) and points where subharmonics are visible on the spectrogram (3): these are discussed in the text, and details of sections 4-6 are given in Figure 4.21. The VoQs symbol [↓] indicates where Andrew produced an audible, 'noisy' inhalation.

4.4.3 Twinkle Twinkle Little Star

Andrew claimed not to know *Happy Birthday* and preferred to sing *Twinkle Twinkle Little Star* (TTLS), to a melodic accompaniment. His phonetic production is given below (Figure 4.23). In comparison to the target production, Andrew's production differed in terms of syllabic stress in lines two and three (Figure 4.23), and in terms of rhythm and melody (Figure 4.24). Andrew produced short sections of song with pauses between each section, and therefore did not align with the expected beat (indicated in blue). Although he did not match the exact intervals, all the notes he sang were contained within the target melody, as indicated by the alignment between Andrew's pitches (in green) and the target pitches (blue) on Figure 4.24. However, the melodic shape is different and is affected by the inconsistent rhythmic production of the melody.

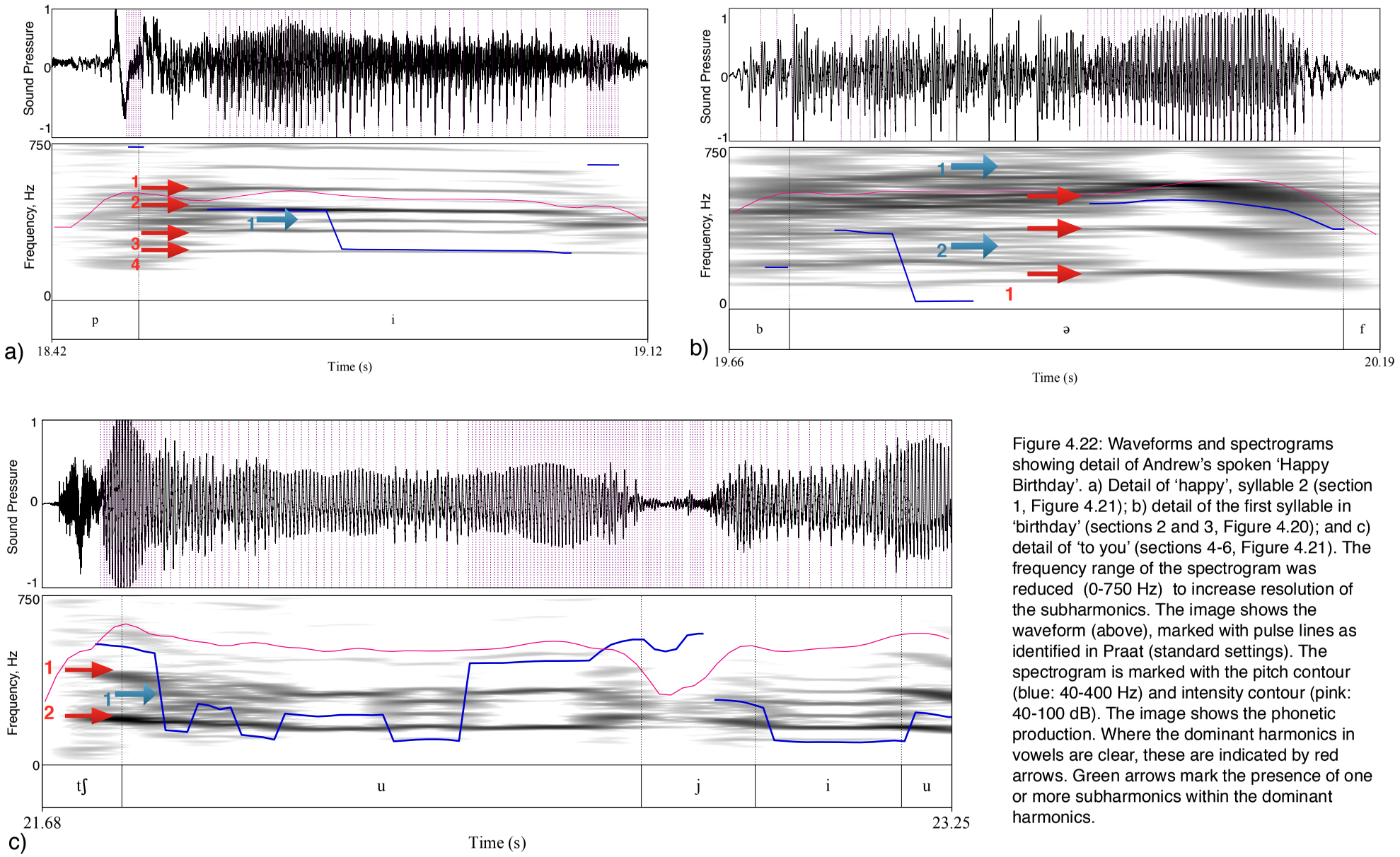
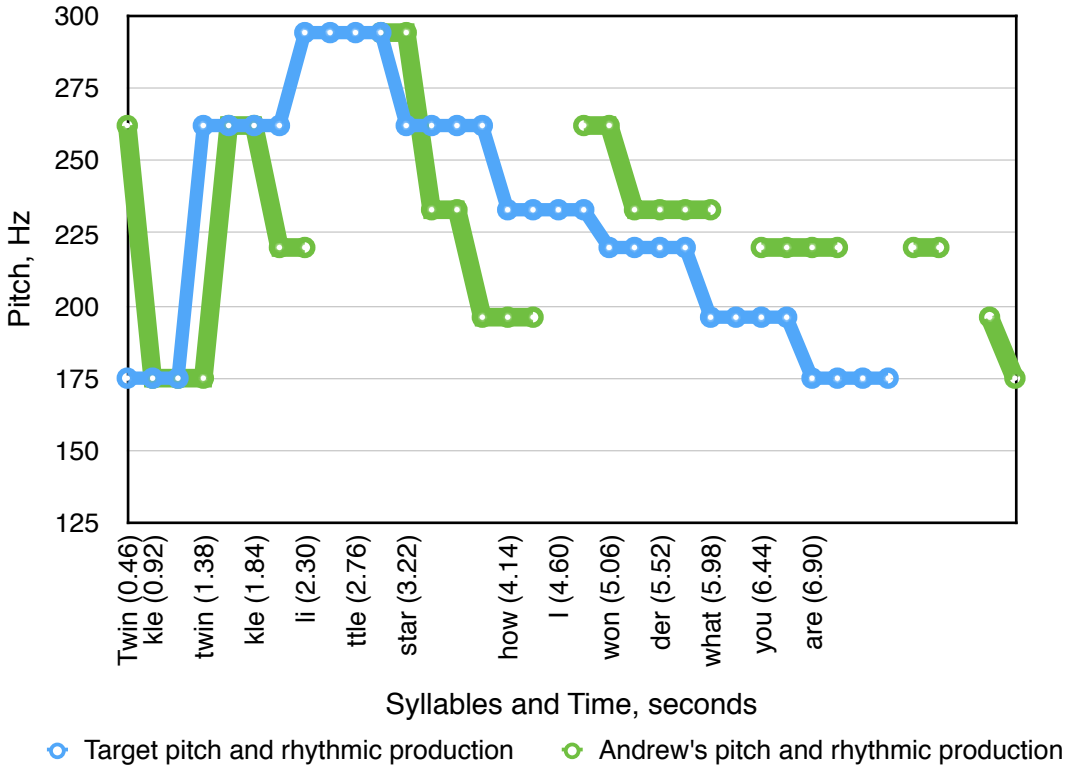


Figure 4.22: Waveforms and spectrograms showing detail of Andrew's spoken 'Happy Birthday'. a) Detail of 'happy', syllable 2 (section 1, Figure 4.21); b) detail of the first syllable in 'birthday' (sections 2 and 3, Figure 4.20); and c) detail of 'to you' (sections 4-6, Figure 4.21). The frequency range of the spectrogram was reduced (0-750 Hz) to increase resolution of the subharmonics. The image shows the waveform (above), marked with pulse lines as identified in Praat (standard settings). The spectrogram is marked with the pitch contour (blue: 40-400 Hz) and intensity contour (pink: 40-100 dB). The image shows the phonetic production. Where the dominant harmonics in vowels are clear, these are indicated by red arrows. Green arrows mark the presence of one or more subharmonics within the dominant harmonics.

Target: TTLS	Andrew
1: **twi**nkle t**wi**nkle	1: {V!! **twɪ**ŋkəʊ **twɪ**ŋkəʊ
2: **li**ttle s**tar**	2: **lei**təʊ sta
3: **how** I **wo**nder	3: hæʊ **a wʌ**ndə
4: **what** you **are**	4: **wʌ** ju **ɑ** V!! }

Figure 4.23: Phonetic transcriptions and syllabic stress of Andrew's production of 'Twinkle Twinkle Little Star'. The table shows the target text with accent marks on syllables and Andrew's production of words and sounds in TTLS. The phonetic transcription indicates the perceived stress: bold type indicates syllables that were produced with primary stress; no syllables were perceived as having secondary stress. His voice quality was perceived as ventricular throughout the production, as indicated by the VoQs {V!!}. It was not possible to calculate the percentage of interharmonics as a result of the piano accompaniment that contaminated the recording.

Figure 4.24: Graph showing Andrew's accuracy in terms of melodic shape when singing 'Twinkle Twinkle Little Star' to an accompaniment. The image shows a plot of the target notes (in blue) and the mean F0 (Hz) of Andrew's produced notes (in green) against the syllables. These are plotted against their target timing to indicate rhythm (see Methods, Chapter 3).



Discussion

The first section (4.5) will discuss Andrew's perception and production of rhythmic, vocal melodic and vocal abilities in motor-music tasks, speech and song. This will be followed with a general discussion in Section 4.6 on the links between the domains before considering the implications for using music to support his speech (Section 4.7).

4.5 Discussion of rhythm, voice and melody in speech and song

4.5.1 Temporal perception and production in music

The evidence indicates that Andrew can perceive key temporal features of beat and rhythm. For example, he changed tempo appropriately in beat entrainment tasks (Table 4.6), and in TRaCoL (Table 4.8), he imitated seven rhythms accurately in terms of number and relative duration of beats. However, features of his performance suggest a motoric difficulty. When asked to clap and drum word rhythms, he was inconsistent in his timing and production of number despite being able to count the number of rhythms to clap (Section 4.2.4). Likewise, Andrew's rhythmic motor movements to a beat were inconsistent and poorly synchronised (Table 4.6). Such features are characteristic of people with DS in a range of motor tasks (Rigoldi et al., 2011; Silverman, 2007; Weeks, Chua and Elliott, 2000), especially as the complexity of tasks increase (Palisano et al., 2001). Historically, research has linked these difficulties in DS to hypotonia, poor response time, and to reduced motor co-ordination (e.g. Latash, 2007; Vicari, 2006). In beat entrainment and beat extraction tasks, his timing was most accurate at about 100 bpm. He initially responded at half the target tempo when clapping to music (Section 4.1.2) and tapping the beat at 104 bpm (Table 4.6). This suggests Andrew may have a preferred optimum tempo, a characteristic that is determined by physiology and the central nervous system (Drake et al., 2000; Krumhansl, 2000; Repp and Su, 2013; Thaut, 2008). Therefore, Andrew may be better able to precisely time his motor movements when there is regular timing between beats and a relatively longer response time.

However, in DS, performance in motor tasks is also linked to cognitive ability (Chen et al., 2015; Malak et al., 2013; Sourtiji et al., 2010), mode of instruction (Bunn et al., 2007; Ringenbach et al., 2006) and the opportunity to practise (Latash, 2007; Latash et al., 2002, 2008). Likewise, the ability to produce accurate motor movements in musical tasks depends upon more than physiological or neurophysiological responses (Bispham, 2006). Accuracy in beat and rhythm entrainment tasks is also dependent upon which aspect of the rhythmic

structure is the focus of attention (Drake et al., 2000; Gérard and Drake, 1990), and upon memory span (Flaugnacco, Lopez, Terribili, Zoia, Buda, Tilli et al., 2014; Grahn and Schuit, 2012; Phillips-Silver, 2009; Repp and Su, 2013; Saito, 2001; Snyder, 2001, 2009). Andrew's poorer performance at tempi above 100 bpm (Table 4.6) may reflect any of these factors. His ability to momentarily produce more finely co-ordinated movements in rhythm entrainment and rhythm imitation tasks (Table 4.7, Table 4.8, Figure 4.1) suggests a difficulty in sustaining attention to the task, a characteristic that he displays frequently, in common with other DS adults (Will et al., 2014). Likewise, although his inability to tap with metrical accent (Figure 4.1) could result from motor immaturity or differences in motor development (Drake et al., 2000; Drake and Gérard, 1989), it could also result from an inability to divide attention between temporal structure and intensity (Gérard and Drake, 1990).

An inability to attend to simultaneous temporal features may also explain why Andrew was able to reproduce the number of taps in some rhythm imitation tasks, but not the relative duration (Table 4.8). However, the role of memory in rhythm imitation and in entrainment tasks is significant. Memory for rhythm is dictated by the number of pulses, not by the number of beats (Drake and Gérard, 1989): furthermore, rhythmic grouping can support serial recall (Hall and Gathercole, 2011). Rhythmic grouping based on pulse could account for Andrew's ability in the TRaCoL task (Table 4.8) to recall a sequence of six claps (number 7), double his digit span of 3 (Table 4.1). As storage of temporal patterns in beat perception is also linked to memory span (Grahn and Schuit, 2012), it is possible that Andrew's reduced STM may have affected his ability to entrain to rhythm. He had no difficulty in entraining to beat entrainment tasks or in extracting the beat at a tempo of about 104 bpm (Table 4.6), which suggests that he has adequate auditory memory to store isochronous patterns. However, one bar of the rhythm entrainment task (Figure 4.1), which was also at 104 bpm, spanned 2.3 seconds. The duration required to store the rhythmic sequence may have exceeded Andrew's auditory memory capacity. His improvement when given visual support may have supported his ability to perceive the stimulus, potentially by negating the need to draw on the phonological loop, a component of STM that is utilised in auditory rhythmic tasks (Saito, 2001).

4.5.2 Temporal perception and production in speech and song

Andrew can read (Table 4.1), but his dysfluent production indicates that this may be cognitively challenging; and he can segment words (e.g. Figure 4.2). Both abilities indicate that he can encode temporal qualities of speech sounds and syllables, and retrieve and manipulate this information with some speed. Andrew scored 100% on minimal pairs

discrimination (Table 4.1) and his performance in speech tasks does not indicate any perceptual difficulty for speech sounds in words. Although he passed the 'same/different' pre-test of the rhythm task (Table 4.1: test 26), he was unable to discriminate whether pairs of rhythms were the same or different. He also performed less well on discrimination of words and nonwords. These tasks were more demanding in terms of phonological working memory than minimal pairs tasks. Furthermore, the 'same/different' test items lasted on average three seconds. It is likely that his memory span affected his performance on this type of task, and that this is further affected by demands on phonological memory.

Andrew's results in perceptual tasks indicate that his hearing does not affect his speech perception for consonant or vowel contrasts (Table 4.1: test 13). The data in speech imitation tasks indicate that Andrew can perceive speech sounds in words. However, data from CNREP (Table 4.15) indicate that Andrew may have difficulty perceiving some weakly-stressed syllables, especially in longer words. Both words of five target syllables that began with weakly-stressed syllables were produced without the weak syllable; by contrast, Andrew reproduced all syllables and most speech-sounds in the strong-initial nonword of five syllables (*sepretennial*) and in all other nonwords with the exception of *blonterstaping* (Table 4.15). Reduced perception of weak syllables for nonsense words of four syllables was reported in DS children and adolescents (Pettinato and Verhoeven, 2008). This only affected weak syllables in specific positions within words and was taken as an indicator of poor memory for the rhythmic structure, rather than memory for the sounds. For Andrew, only weak-initial words of five syllables were affected but productive difficulties are also implicated in nonwords of four syllables. For example, *perplisteronk* (Figure 4.13) and *contramponist* (Figure 4.14) had the same target prosodic structure, but the rhythmic accuracy of Andrew's productions differed. Although he replicated the rhythm of *perplisteronk,* his rhythmic production of *contramponist* was affected by the placement of stress on the initial weak syllable.

Andrew's ability to recall longer nonwords (4-5 syllables: Table 4.13) indicates a phonological capacity that is superior to many DS subjects (Laws, 1998). However, a limited capacity for auditory information may have influenced the speed of Andrew's responses in imitation tasks. Andrew interrupted the demonstration of *blonterstaping* (Table 4.13) and his imitation of *zippering*, *hopefully* and *hardening* overlapped with the final syllable of the demonstration (Figure 4.9, marked by asterisks). His speaking rate of two-syllable words, nonwords and phrases was also faster than the demonstration (Table 4.10, Table 4.13, Table 4.16) but his imitation of words of one and three syllables was close to or slower than the demonstration

(Table 4.10). The speed of articulatory rehearsal is linked to recall in subvocal tasks and faster overt rehearsal or presentation effectively increases verbal memory span (Jarrold, Baddeley and Hewes, 2000; Purser and Jarrold, 2005). Therefore it is possible that for some words Andrew accelerated his speech production as a strategy to increase the capacity of his phonological memory. This may also account for his relatively short initial syllables in most imitated words of two and three syllables, which were within his digit span (Figure 4.8, Figure 4.9). However, in some words, omission or errors of speech sounds affected Andrew's articulation time, and his stress production was affected by an increase in vocal intensity, especially relative to the target (e.g. Figure 4.8: *harden, city*). The changes in intensity did not always coincide with changes in pitch or relative duration, leading to atypical prosodic production (e.g. Figure 4.8: *soften*).

Andrew's ability to produce speech sounds and words in phrases with appropriate rhythm was affected by articulatory and habitual speech errors (Table 4.3), by incomplete attention to demonstrations (Table 4.13) and by reduced memory for lexical content (Table 4.16). However, the data also indicate that motor timing difficulties affected his consistency in syllable-timing (e.g. Table 4.18: RD) and his speech fluency (Table 4.3). For example, Andrew demonstrated overt stuttering during reading (Table 4.3: test 15) and when planning a complex sentence (Figure 4.18: Comment 1). In tasks that require speech planning, it is believed that dysfluency occurs when there is a mismatch in timing between phonation and the intended linguistic plan (Bosshardt, 2006; Howell, 2010). However, hesitations and repetitions of segments occurred in other tasks which may indicate an underlying condition associated with motor planning difficulties (Bray, 2008). Andrew also demonstrated grunting and teeth grinding (see Section 4.4.2, p.112), which are symptoms that are associated with stuttering (Alm, 2004; Bray, 2008; Howell, 2010), and task avoidance. Teeth grinding and dysfluency both worsen under stress (Alm, 2004) and increased cognitive load (Yap, 2012). Andrew stated that he did not 'know' the words of *Happy Birthday* and *Magic Book* and it is probable that his effort to recall these affected his rhythmic production and that the stress of recall also affected his vocal quality.

Dysfluency may indicate a specific difficulty in laryngeal adjustment (Alm, 2004; Bosshardt, 2006; Sebastien et al., 2013). This may stem from a muscular weakness, from a motoric difficulty associated with stuttering, or from over-compensatory strategies to activate phonation (Alm, 2004; Sebastien, Benedict, and Balra, 2013). It is documented that Andrew uses compensatory muscles for speech (see Section 4.1.3) and it is possible that the same similar overcompensatory actions underlie his ventricular phonation and dysfluency. The

underlying difficulties may be physiological, cognitive or habitual in origin, but the data indicate that his difficulties in fluency and voice are both affected by high cognitive load or emotional response.

4.5.3 Voice quality and phonation

Andrew's voice quality is consistent with accounts of harsh, low-pitched voicing in DS (e.g. Montague and Hollein,1973; Moran and Gilbert, 1978; Moran, 1986). His production was perceptually effortful in most tasks, and during teaching and assessment activities he habitually tensed his shoulders and diaphragm to blow pinwheels, use kazoos or sing (e.g. see Appendix 3, p. 342 [38]). His symptoms indicate a form of MTD in which excessive tension is applied to the external laryngeal muscles in order to activate the voice, which affects the muscles within the larynx (Aronson and Bless, 2009; Roy, 2003; Salatoff, 2014). In MTD, the hyperadduction of laryngeal muscles and supraglottal muscles result in a perceptually harsh, strained and 'strangulated' quality (Laver, 1980; Stager, 2011). MTD can result in ventricular dysphonia in which the false vocal folds (FVF) are engaged along with or instead of the true vocal folds (TVF) (Aronson and Bless, 2009; Stager, 2011). Both MTD and the activation of the false vocal folds can also result in diplophonia (Cavalli and Hirson, 1999; Dejonckere and Lebacq, 1983; Fuks, 1999; Lee and Son, 2005) which has been identified in people with DS (Beckman, Wold and Montague, 1983). Diplophonia was present in one of Andrew's sustained vowels (Table 4.9: /a/1) and in his continuous speech samples (Figure 4.18: Comment 1, Figure 4.21, Figure 4.22).

Novak (1972) proposed that ventricular phonation is the cause of the perceptually harsh voice in DS subjects, and Andrew's symptoms are consistent with this. MTD and ventricular dysphonia may be intermittent and stem from several causes, including habitual, emotional or organic factors (Aronson and Bless, 2009; Mathieson, 2001; Van Houtte, Van Lierde and Claeys, 2011). Andrew was able to relax the muscles in his torso in order to produce phonation when instructed (Appendix 3, p. 348 [1]), which indicates that his tension is a habitual response. A minimum Phonatory Threshold Pressure (Titze, 1994) is required to initiate oscillation of the vocal folds (Braunschweig, Flaschka, Schelhorn-Neise and Döllinger, 2008; Titze, 1994, 2008, 2014); and additional subglottal pressure is required at onset (Koenig, Fuchs and Lucero, 2011; Lucero, 2005; Zhang, Neubauer and Berry, 2006). Weak expiratory pressure has been identified in DS subjects (da Silva, 2010; Salguerinho et al., 2015) and is believed to affect phonation in DS subjects (Pryce, 1994). Therefore, in subjects with DS, an excess of energy may be needed to initiate the vocal folds (Pryce, 1994). Furthermore, people with dysphonia may overcompensate for weak or poorly-toned

vocal muscles by using excessive muscular force and excessive sub-glottal pressure (Braunschwieg et al., 2008). It seems probable that Andrew's phonatory difficulties stem from a behavioural response to a physiological characteristic that is common to DS subjects.

Once voice is initiated, difficulties in co-ordinating respiratory muscles for exhalation may further affect production in DS subjects (Pryce, 1994). Heselwood et al. (1995) identified a link between the position of words in a phrase and voice quality in the production of a man with DS: they argued that the breathy quality they observed in initial pre-stressed units may result from reduced respiratory control. Data from Andrew's Syllables test show that in words of two and three syllables, second position vowels were produced with highest HNR (Table 4.12). However, CPP values were lower overall for three syllable words (Table 4.11) and for vowels in second and third position (Table 4.12). High HNR indicates that second syllables were less 'noisy', but low CPP values are typical of phonation that is irregular and affected by low-amplitude harmonics (Garrett, 2013; Fraile and Godino-Larente, 2014; Hehman-Ackah et al., 2003). Three-syllable words were also produced with a higher proportion of interharmonics that remained visible on spectrograms at 15 dB (Table 4.10). This indicates a more intense production that is consistent with the reduced CPP values. In word-initial syllables, the lower HNR (Table 4.12) could stem from a reduction in breathiness or from more periodic production (Mathieson, 2001). A reduction in word-final vowels of both HNR and CPP indicate that Andrew may have greater difficulty in maintaining efficient phonation towards the end of words.

In contrast to results from the Syllables test, initial syllables in CNREP (Table 4.15) were produced with fewer perturbations than those in subsequent syllables: however, with the exception of four-syllable words, all vowels in CNREP were produced with high perturbations (low HNR, high DUV) and a high proportion of ventricular production and high intensity interharmonics (Table 4.13). It is proposed that this would lead to air wastage, and impact upon Andrew's phonation. This effect is illustrated by the difference in duration between Andrew's production of sustained vowels with modal phonation (up to 1896 ms (/a/3): Table 4.9) and his ventricular production (876 ms (/a/1): Table 4.9, Figure 4.7). A delicate balance is required between laryngeal tension and subglottal pressure (Baken and Orlikoff, 2000; Mathieson, 2001; Titze, 1994; Titze, 2014) and excessive muscular force at initiation can disrupt the periodicity of vocal fold oscillations (Beck, 1988; Braunschweig et al., 2008). Andrew's apparent application of extra muscular tension to initiate phonation would affect the balance between airflow and muscle tension in the vocal folds, which can further

destabilise phonation (Cavalli and Hirson, 1999; Lucero, 2005; Lucero et al., 2012; Tao and Jiang, 2008; Wilden, Peters and Tembrock, 1998).

It is not clear why all three nonwords of four syllables seemed to be produced with less atypical voicing as indicated by lower DUV (%), higher CPP (dB) and less intense interharmonics (Table 4.14). In comparison with nonwords of two, three and five syllables, these productions were also produced with a higher mean F0 (Table 4.14). An increase in fundamental frequency could be reflective of a stress-response, which has been associated with reduced jitter in some instances (Giddens, Barron, Byrd-Craven, Clark and Winter, 2013). However, Andrew's increased speaking pitch may be in response to the wider pitch range of the demonstrations (Table 4.13), which used a higher pitch to emphasise the stressed syllables. Furthermore, nonwords of four syllables were less accurate in terms of segmental production than nonwords of two and three syllables (Table 4.14), which may reflect a change in the focus of his attention. He reproduced most speech sounds in the shorter nonwords, which consisted of up to eight sounds; but his productions were less precise and less appropriate to the target word in words of four syllables, which consisted of twelve or thirteen sounds. The four-syllable stimuli used in the present study are deemed high in articulatory complexity, a factor which affected accuracy in production of children with SLI (Archibald and Gathercole, 2006). The increased complexity of the target word would place greater demands on Andrew's phonological memory, and on the processing required for articulatory planning. Potentially, Andrew switched his attention from segmental to suprasegmental qualities, resulting in a reduction in planning load, with consequences for his vocal production.

Voice is a recognised indicator of an individual's response to stress (Aronson and Bless, 2009; Roy, 2003) and the results from the range of tests indicate that Andrew's voice quality may be sensitive to stress or cognitive load. Emotions such as anxiety can exacerbate tension in the larynx (Johnstone and Scherer, 1999; Patel, Scherer, Björkner and Sundberg, 2011), and increase the incidence of muscle-tension dysphonia (Aronson and Bless, 2009; Pinho, Pontes, Gadelha, and Biasi, 1999; Roy, 2003; Salatoff, 2014). The data indicate that Andrew's voice quality deteriorated when increased demands were placed on his auditory memory. In word and nonword imitation tasks, the percentage of words produced without phonation increased as the target number of syllables increased (Table 4.11, Table 4.15). Furthermore, in words of three syllables, CPP values in all vowel positions were low (Table 4.11, Table 4.12). These examples indicate that planning for longer words affected Andrew's

phonation from the outset. The effect of cognitive load on phonation is noted in TD populations, as increasing processing demands can affect the incidence of 'creakiness' (Yap, 2012, 2015) and functional dysphonia (Aronson and Bless, 2009; Giddens et al., 2013; Patel et al., 2011). The data suggest that Andrew habitually increases the muscular tension in his intrinsic and/or extrinsic laryngeal muscles in response to an increase in cognitive load. It is not clear whether this is the direct effect of an increased cognitive load, or an emotional response to increased cognitive demands of the tasks.

4.5.4 Pitch perception and production in song and speech

Andrew's melodic production in speech and song was affected by phonatory difficulties that resulted in dysphonia, diplophonia and intermittent voicing. It is therefore difficult to gauge his conceptual understanding or perception of melodic detail from his production. Additionally, his results in the 'same/different' pitch perception task were unreliable (Table 4.1: test 27), and were probably affected by memory limitations. In the pitch perception task, he correctly matched over half of the pitches and intervals to images. In the test items that he did not correctly match, his responses indicate a difficulty in discriminating a continuous change in pitch of a semitone (Appendix 2 p. 335: pitch bend), rather than in discriminating between discrete pitches; and a difficulty in identifying the difference between a pair of notes in unison and a pair an octave apart (see Appendix 2, p. 335). The latter may stem from a perception that a note an octave apart is the 'same', but the inability to discriminate pitch bends may indicate a perceptual limit for small pitch changes. The data from Andrew's production tasks indicate that he is able to perceive intonation contours in words (Figure 4.8, Figure 4.9, Table 4.10), even though he does not reliably reproduce them.

Andrew improved in his ability to match pitches (Figure 4.5) but his pitch remained limited in range. His melodic production when learning *Magic Book* developed over the course of five weeks, but was most accurate when prompted with starting notes and words (Figure 4.17: A2). This indicates that Andrew is able to produce melodic detail more accurately when his memory is supported, an effect that is apparent in self-reported 'tone-deaf' singers (Wise and Sloboda, 2008). His accuracy in independently producing intervals did not improve (Figure 4.6, Figure 4.17) and his production of intervals in familiar songs was also inaccurate, but within a consistent key (Figure 4.24). Andrew's ability to 'read' a scale from a visual chart (Figure 4.7) demonstrates knowledge of tonality and a conceptual understanding of pitch direction that may have underpinned his production of TTLS (Dowling, 1978; Deutsch and Feroe, 1981; Deutsch, 2013). However, his production was partly spoken and partly sung. It is possible he did not recall the absolute intervals, or that his voicing

difficulties affected his accuracy, but the data also suggest that Andrew is at an early stage of singing development. His performance is consistent with Rutkowski's (1997) description of an 'inconsistent limited range singer', which is the fourth phase of nine in a developmental model.

Andrew produced intonation sufficiently well to convey meaning (Figure 4.19) and to reproduce the pitch contours of some words (Figure 4.8, Figure 4.9) but he incorrectly produced most spoken phrases with a rising intonation (Table 4.16). The results are comparable to those of Stojanovik (2011) and Setter and Stojanovik (2011) who reported inconsistencies in the ability of DS children to signal differences in statements and questions. Although Stojanovik (2011) and Setter and Stojanovik (2011) argue that this signals productive prosody deficits, Andrew demonstrated no difficulties in reproducing melodic contours in any other speech or music task, with the exception of his first imitation of *Magic Book* (Figure 4.17). The rising intonation in phrases may signal questioning or uncertainty in his production, rather than a prosodic deficit. This may in itself result from poor memory for the lexical content.

Phonatory instabilities affected the perceived melody of Andrew's speech. As such, these difficulties limit Andrew's ability to produce the 'fine-grained' changes in pitch required for accurate singing. As previously observed, Andrew produced a higher percentage of atypical phonation (loss of phonation, increased interharmonics) in tasks requiring increased cognitive load. The data indicate that Andrew's expressive difficulties in prosody are primarily motoric and are partly due to difficulty initiating voice, but that this is susceptible to increased memory load.

4.6 Links between musicality and speech

4.6.1 Difficulties common to speech and music domains

Andrew showed consistent difficulties in producing consistent rhythmic motor movements in music, speech and song. His results are consistent with his difficulties in gross motor movements and oro-motor movements (Section 4.1) and with research into motor difficulties in DS (Latash et al., 2002, 2008; Ringenbach et al., 2006; Sacks and Buckley, 2003). This may indicate a general muscle laxity or immaturity that affects his limbs and his speech. However, during rhythm entrainment tasks, Andrew was better able to reproduce the number of items than the relative duration (Figure 4.1), and he was inconsistent in clapping syllables of words despite knowing how many claps to produce (Section 4.2.4). The variability in

relative duration of syllables in repeated words in songs (Table 4.17, Table 4.18), dysfluencies, phonatory difficulties and poor timing to a beat may all be symptoms of a motor timing difficulty. It is possible that Andrew has a central timing difficulty that is common to his manual motor production and oro-motor production in speech and song. However, the data suggest that Andrew's voice production and attention are affected by task demand, such as memory limitations and the ability to focus attention. These difficulties may have further consequences for articulation, for rhythm in speech, song and motor tasks, and for melody and prosody.

Andrew's digit span affected his accuracy in all imitation tasks and may also have affected his ability to perceive rhythms and the rate of his response and production in speech tasks. Andrew's performance in rhythmic motor tasks improved when attention was drawn to the salient beat with visual prompts (e.g. Figure 4.1), in line with other DS individuals (Ringenbach et al., 2006; Chen et al., 2015). It is possible that the visual representation supports or supplements the demands on his phonological memory, which is required for rhythmic processing in musical tasks as well as speech tasks (Saito, 2001; Williamson, Baddeley and Hitch, 2010). Similarly, his voice quality in all tasks was less atypical in tasks that required little memory (e.g. sustained vowels), which indicates the effect of auditory memory demands. However, his improved motor-rhythm performance in response to visual stimuli also supports evidence that the accuracy of his rhythmic production may be affected by a difficulty in attending to the salient feature of rhythms (Gérard and Drake, 1990). Andrew's attentional and behavioural characteristics are characteristic of impaired executive function which results in a difficulty in sustaining attention, reduced inhibition of impulses and difficulties in performing more than one task (Lanfranchi et al., 2009, 2010, 2012; Will et al., 2014). Therefore, Andrew may not have enough 'spare' cognitive capacity for tasks that require processing on several levels — e.g. phonological, temporal, melodic — which may lead to his observed behaviour and indications of stress in his phonation and fluency. A discussion of Andrew's production of *Happy Birthday* in sung and spoken conditions illustrates how the combined effects of Andrew's cognitive difficulties may have affected his production of speech and song.

4.6.2 Happy Birthday: evidence of transfer potential

Andrew expressed difficulty in recalling the words when singing and speaking *Happy Birthday.* The time required to recall the words and phrases slowed his production, resulting in the observed difficulties in producing the words rhythmically (Table 4.18, Figure 4.21). Slowed speech performance is an indicator of stress (Buchanan, Laures-Gore and Duff,

2014). However, despite ventricular and diplophonic production in part of the spoken task, Andrew's vocal production was less overtly effortful than in other speaking tasks, with the exception of word imitation tasks, and contained fewer interharmonics (e.g. compare the visible 'noise' and number of spectrograms in Figure 4.21 to Figures 4.10, 4.12, 4.15 and 4.19). It is likely that a combination of factors led to this. Andrew's voice and his speech production appear more susceptible to ventricular phonation in tasks requiring increased auditory memory (Table 4.11, Table 4.14) or linguistic planning (Figure 4.18, Figure 4.19). The production of pre-learned sequences of speech sounds and words may have resulted in a reduction of tension as a result of previous articulatory practice and through a reduction of required linguistic planning. A reduction in physical tension and stress would promote less effortful phonation (Aronson and Bless, 2009; Lee and Son, 2005). In addition, Andrew's ventricular production of /a/1 (Table 5.9, Figure 4.4) suggests that he has underlying difficulties in producing modal phonation even when cognitive load is low. Such difficulties are exacerbated in speech, partly in response to rapid articulatory demands (Fourcin and Abberton, 2008), and because phonation may be slow to stabilise, a feature observed in people with dysphonia (Schaeffler et al., 2015). The slow production of words and relatively long duration of Andrew's vowels in the spoken condition of *Happy Birthday* (Figure 4.21) may therefore have allowed time for his vocal system to stabilise, resulting in a higher percentage of stable phonation.

Although Andrew's production of melodic intervals and the absolute relative timing of subsequent syllables differed to the target, his sung and spoken phrases were similar to each other in terms of the relative duration of syllables and changes in melodic contour (Table 4.18). Whilst there is evidence that these melodic and rhythmic features of the song transferred to his spoken production, he recalled more words and phrases in the spoken condition (Figure 4.20). There are two possible interpretations of Andrew's improved recall in the spoken condition: his response to the tasks; and the different demands associated with speaking and singing. When asked to sing the song, Andrew's statement of 'I don't know the words' may have reflected a reluctance, rather than explicitly a lack of knowledge, which affected his production. In contrast, his response when asked to speak the words was 'yes, of course I can'. It is recognised that singing songs can place additional demands on attention and memory during the learning process and that the simultaneous processing of different tasks (e.g. melodic, rhythmic, lexical) can increase cognitive load (Deutsch, 2013; Racette and Peretz, 2007). It is also recognised that DS adolescents have difficulty in processing two types of verbal-based information simultaneously (Lanfranchi et al., 2010) and in switching attention in verbal tasks (Carney et al., 2013; Costanzo et al., 2013). For

Andrew, the dual demands of recalling words alongside musical detail may have affected his response. However, despite Andrew's difficulty in recalling words, the previous learning of the text as a song may have assisted his production of speech sounds and voice in a similar way to that observed in MIT. Research indicates MIT can assist speech production as a result of slower production, increased processing time, melodic and rhythmic chunking effects, and longer duration in which phonation may stabilise (Schlaug et al., 2008; Wan et al., 2010).

4.7 Implications

The data suggest that motor difficulties may partly account for Andrew's musical, speech and singing abilities and that these stem primarily from muscle immaturities, poor co-ordination or reduced muscle tone. Training can improve accuracy and efficiency in a range of limb motor-movements in DS subjects (Sacks and Buckley, 2003; Latash, 2007; Latash et al., 2008) and appropriate sensory feedback can support musical motor movements (Rigenbach et al., 2006) and articulatory movements (Hamilton, 1993; Wood, 2010). Potentially, Andrew could develop his motor skills in all domains through rhythmic activities and singing. Additionally, Andrew's positive response to visual feedback in the rhythmic task (Figure 4.1) and in 'singing a scale' (Figure 4.7) suggest that the use of visual feedback would be beneficial in developing motor-rhythm and singing skills. Simple gestural feedback is beneficial to children as they develop accuracy in pitching (Tsang et al., 2011; Welch, 1985), and gestural, visual, and computer-assisted feedback has been successful in supporting singing accuracy in hearing-impaired children (Welch et al., 2015). However, in singing, he would first need to overcome his habitual responses that can result in physical tension. The physical effort he uses for phonation results in ventricular production. This reduces the stability of co-ordination between vocal folds and respiratory system and makes it more difficult for Andrew to control the relative duration in longer words or phrases as he runs out of air. Using wind instruments could help him to achieve controlled, sustained exhalation. Activities that promote the relaxed onset of phonation should be a priority for Andrew. These could include humming or the use of kazoos, which can support respiratory control and co-ordination between respiration and phonation (Thaut, 2008). The practice of humming is recommended in singing pedagogy to support efficient phonation (e.g. Phillips, 1996) and has successfully reduced symptoms of vocal tension in people with MTD (Lee and Son, 2005; Ogawa et al., 2014).

Inconsistencies in timing to a beat, and in clapping or speaking rhythms of familiar words suggest that Andrew has difficulties in rapid motor timing. If this is the case, rhythmic motor

activities to support muscle memory may not benefit his speech imitation or production at the syllabic level unless he can also develop the accuracy and consistency of his motor timing. Thaut (2008) argues that within neurological constraints, motor timing may also be supported through rhythmic activities, especially if these are based on the individuals' preferred tapping rate and physical abilities. Andrew might therefore benefit from drumming and rhythmic activities. These can support large motor movements, and would provide Andrew with a more stable base from which to develop timing in finer motor movements. He may also benefit from activities such as rhythmic speech cueing, which encourages speech production at a pre-determined pace with equal syllables (Thaut, 2008). This could support the timing and co-ordination of his speech at a syllabic level (Thaut, 2008; Wan et al., 2010). Additionally, a rhythmic structure may support the timing of phonation (Davidow, Bothe and Ye, 2011; Thaut, 2008), which would help promote more efficient vocalisation and reduce associated physical tension.

Andrew demonstrated an ability to learn songs in terms of word content and rhythmic and melodic shape, but he made little progress in his vocal production or accuracy in singing pitches or intervals. This is not unexpected, given his age and the short period of teaching. However, even in a lengthier, or earlier, intervention his progress would be hampered by his ability to listen and attend. The ability to hear notes and rhythms before producing them is required for accurate singing (Phillips, 1996). Andrew did not leave himself time to do this, which would affect the accuracy of his performance and slow his learning. In common with the DS population, this suggests difficulties in executive functions (Costanzo et al., 2013; Lanfranchi et al., 2010; Will et al., 2014). As with young children, he may also have difficulties in focussing on more than one aspect of a task in motor-rhythm tasks (Gérard and Drake, 1990). Similarly, his apparent difficulty in reproducing segmental and suprasegmental aspects in speech imitation tasks may indicate a difficulty in dividing attention when imitating speech, which has been observed in children with learning disabilities (Jackson, Treharne and Boucher, 1997). Therefore, Andrew would need to develop his ability to focus and sustain attention in order to benefit from learning. When learning new words or songs, he might benefit from a specific teaching approach in which his attention is directed to specific aspects of the task, one at a time, before combining segmental and suprasegmental aspects.

Andrew's results indicate that his phonation has fewer perturbations in singing than in speech, but that fewer perturbations are produced when segmental content is low (Table 4.10) and when words are drawn from long term memory (Figure 4.21). Stress-recognition

and management techniques may be necessary to alleviate any emotional responses to cognitively demanding or stressful tasks. Additionally, in learning songs, his voice production would need to be supported by strategies to reduce the cognitive load. A reduction in auditory memory demands can support vocal skills in singers (Berkowska and Dalla Bella, 2009a; Christiner and Reiterer, 2013; Dalla Bella, Tremblay-Champoux, Berkowska and Peretz, 2012; Welch, 2016). Strategies such as singing with Andrew or using visual support could help reduce auditory memory demands and support his vocal production when singing. Potentially, singing activities may help Andrew to overcome a habitual response to producing voice. This would be required in order to further assess his potential in learning from musical activities, and to fully consider the potential use of music and singing to develop aspects of his voice or speech.

Chapter 5: Kerry

Section 5.1 will summarise information from Kerry's college records regarding her speech, voice, behaviour and musical experience. The first section will also summarise data from the battery of assessments (Chapter 3, Section 3.2) and her learning in sessions. Subsequent sections will examine specific aspects of musical and vocal abilities in more detail in order to consider the range and influence of different factors on her musical and vocal abilities. Section 5.2 will discuss her motor-music skills in moving to music, and in imitating and creating clapped rhythmic patterns. Section 5.3 will focus on her ability to imitate melody and rhythm in musical patterns, words, phrases and songs, and on her voice quality. Section 5.4 will also examine Kerry's ability to produce melody, rhythm and voice, but it will focus upon spontaneous speech and extracts of speech and song drawn from long-term memory.

Kerry's ability to produce rhythm, melody and to sustain phonation in these tasks will be discussed in Section 5.5. The section will highlight her key strengths and difficulties in producing these elements in speech and in song. Section 5.6 will examine which abilities and difficulties are common to speech and song and how and why these may differ between domains. The chapter will conclude with a summary in Section 5.7 of Kerry's abilities, and a discussion on how music-based activities may support her speech development.

5.1 Summary of history, learning and performance in core assessments

Kerry is 23;4 years of age and she is described in her college records as having SLD, dysfluency, an unspecified degree of bilateral hearing loss, severe myopia (for which she wears glasses), and very low speech intelligibility. She has a history of behavioural difficulties, related to non-compliance, and displays frustration when asked to repeat speech. Her college file also reports that she 'fixates' on some people, objects or events, which can affect her behaviour in some activities. According to her SaLT, Kerry has no difficulties in perceiving speech sounds; she has a basic understanding of language, and cannot understand sentences of more than three words; she uses mostly single words and some Makaton to communicate; and she can retain single pieces of information in her auditory memory. Her college records state that Kerry is a keen music-lover with a strong preference for rock and pop music, and that in the three years prior to the study, she had engaged in weekly singing, drama, music and 1:1 drum tuition.

5.1.1 Cognitive and perceptual abilities

Kerry's test results (Table 5.1) suggest a verbal MA equivalent of three years (BPVS) but a non-verbal MA of seven years (Draw-a-man). Her auditory memory is measured as 3 digits.

Kerry's responses in the music perception tasks (Table 5.1: tests 21, 27 and 28) are unreliable as she consistently selected the picture that reminded her of a friend. As she was unable to complete

Table 5.1: Profile: a summary of Kerry's performance in core tasks. The table summarises the results of Kerry's core assessments in nonverbal and verbal ability, speech perception, and production and musical perception and production. Test numbers refer to the test descriptions in Chapter 2, Section 3.3.

Test Number		Test name	Results
Cognition	**1**	BPVS RAW score	26
	1	BPVS age equivalent	3;1 (2;9 - 3;7 years)
	2	Draw-a-person	7;3 years
	3	TAPS - digit span	3
Perception (speech)	**7a**	Same-different discrimination non-words (Stackhouse and Wells App C.2)	1/10 = 10%
	7b	Same-different discrimination real-words (Stackhouse and Wells App C.2)	NULL
	13a	Minimal pairs perception: discrimination of consonants	2/5 = 40%
	13b	Minimal pairs perception: discrimination of vowels	4/5 = 80%
	14	Simple speech prosody test	3/8 = 38% - but NULL
Speech production	**5**	DEAP: Percentage of Phonemes Correct (PPC)	38%
	8	S/Z	n/a
	9	DDK	n/a
Perception (music)	**17**	PMMA	n/a
	21	Pitch perception test	0/3 = 0% but NULL
	27	Rhythm perception test	0/4 = 0%
	27	Interval perception test	3/4 = 75% - but NULL
	28	Find the sound – pitch matching	1/5 = 20% - but NULL
Music production	**4**	Rhythm imitation (TRaCoL)	4/14 = 28%
	18	Gross motor score	27/40 = 67%
	22a	Pitch matching on /la/	3/4 = 75%
	22b	Pitch matching on /a/, /i/, /u/	5/11 = 45%
	23a	Pitch direction: vocal glides on /la/	4/4 = 100%
	23b	Interval matching	1/5 = 20%
Reading ability	**15**	Words correctly produced ('Swimming' passage) Incorrect words: 'float'= *floor,* 'mum' = *moon*; bubbles= *ball*; 'yell': spelled out: *y-e-l;* 'dry off with' = *dry to watch*; 'cool'= *cold*	96/103 = 93% n=7 Rate: 0.7 syll/sec

the preliminary tasks, she did not attempt the PMMA test. Her scores on speech perception tests (Table 5.1: test 13) indicate that she has difficulty in discriminating between consonants (40% accurate) and between vowels (80% accurate). However, the influence of her hearing loss on the test itself is not known and her behaviour during assessments and teaching suggests that she may not always hear instructions: for example, Kerry sometimes responded to imitation tasks and instructions with *what?* or *what's that?*.

Kerry read slowly and most words were articulated clearly (Table 5.1: test 5). She did not recognise all words (e.g. 'yell', which she spelled out) and she mistook the word 'cool', for *cold*, which may indicate that she processed the meaning of the word or phrase. She mistakenly read 'dry off with' as *dry to watch*, which suggests she initially processed 'dry' as *try*. Kerry did correct this which indicates that she was monitoring her reading for meaning.

5.1.2 Motor-music abilities

Based on a visual assessment (see Appendix 2, p. 324) Kerry was able to imitate movements in time to the beat (Table 5.1: test 18). She was able to alternate feet and arm movements and to imitate large gross motor movements such as body twists with close timing to the dominant beat. She was unable to create movements in time to the music.

5.1.3 Oro-motor and phonological abilities

Kerry produced a range of articulatory errors and errors that are deviant, as defined by DEAP (Dodd et al., 2002)(Table 5.2). Her common phonological errors included voicing and devoicing of consonants (e.g. egg = [eɪk]); cluster reduction (square = [swɛs]); and deletion of weak syllables

Table 5.2: A summary of Kerry's speech errors in the DEAP task. The table shows the number of errors made in the task and the error patterns (n= 5 or more; except WSD* n=3), expressed as a percentage of the total number of errors made. The errors are coded according to the conventions outlined in the DEAP manual, which defines the following categories: D indicates delayed error patterns; U indicates unusual error patterns.

DEAP		
Percentage of error pattern/total errors		
Cluster reduction	**D**	12.5
Voicing	**D**	19.6
Final consonant deletion	**D**	8.9
Weak syllable deletion*	**D**	5.3
Initial consonant deletion	**U**	10.7
Other	**U**	14.3: e.g. ISD;/dʒ/-[k]; f =[h]; θ =[j]; f= [s]; ŋ =ps; v= [p]; k= [ps]
Number of errors		**56**
Percentage of errors accounted for by error patterns		**71.3**

(tomato =[matoʊ]). Unusual errors included initial consonant deletion (duck = [ʌkz]); backing (frog = [fɹɒq]); and fronting of stops (book = [bəps]). A number of sounds were produced with intrusive aspiration or inaccuracies in absolute placing. For example, final stops tended to be produced as fricatives (sheep = ʃiɸ], suggesting incomplete closure of the lips; or with a degree of nasal escape, suggesting incomplete closure of the nasal cavity. The liquid /ɹ/ was substituted with a glide [w] in many instances, but not all. This was included as a habitual articulatory error (Table 5.3), rather than as a phonemic error (see the explanation of PCE-R, Chapter 3, section 3.2.1).

Much of Kerry's speech in imitation and production tasks was dysfluent (Table 5.3) and she was unable to complete the DDK task (Table 5.1). The percentages of omitted sounds and of words affected by dysfluency increased as target syllable length increased (Table 5.3).

Table 5.3: The percentage of speech errors produced by Kerry in core speech tasks. The table shows the percentage of errors in phonemes (PPE), consonants with habitual errors removed (PCE-R), vowel errors (PVE), the percentage of sounds omitted (PEO) and the percentage of dysfluency.

Test no.	Task	No. of words, n=	PPE, %	PCE-R, %	PVE, %	PEO, %	Dysfluency, %
5	**DEAP**	**45**	37.89	51.00	14.89	11.04	43.02
11	**Syllables**	**27**	42.84	70.22	12.36	46.82	67.60
	1-syllable words	**9**	40.42	62.00	0.00	11.80	22.00
	2-syllable words	**9**	43.61	70.80	6.24	13.30	55.00
	3-syllable words	**9**	54.00	75.72	22.75	24.80	55.00
6	**CNREP**	**10**	71.39	80.13	54.20	36.47	30.00
10	**Phrase imitation (Newton, 1999)**	**25**	37.73	70.72	50.30	55.40	39.00
15	**Reading**	**24**	27.87	29.15	6.00	9.10	0.00

5.1.4 Phonation and voice quality

Kerry's vocal range and highest F0 when imitating pitch glides (Table 5.4: test 22b) was similar to her range and highest mean F0 in spontaneous speech: both spanned less than an octave (Table 5.4). In the picture description task (Table 5.4: test 10), her minimum value of F0 (80.76 Hz)

Table 5.4: The mean pitch and pitch range of Kerry's vocal glides and connected speech. For each task, the table shows the mean, minimum and maximum values of fundamental frequency in Hz, and the SD of F0 when speaking, in Hz and semitones (ST). The mean change in speaking pitch is also shown in ST. Mean F0 and SD of F0 is not given for glides, whose purpose was to examine vocal range.

Test number		Mean F0, Hz	Min F0, Hz	Max F0, Hz	SD of F0, Hz; ST	Pitch range, ST
22b	**Pitch glides on /la/**		221.20	324.10		6.61
10	**Picture description**	245.55	206.08	293.27	19.52; 1.37	6.11
n/a	**Spontaneous speech (comment for R)**	281.42	213.96	358.60	43.13; 2.66	8.94

reflects the perceptually 'creaky' vocal production she sometimes produced, and her overall pitch range was lower (Table 5.4).

The mean voice perturbation data for vowels in core speech tasks and for sustained /i/, /a/ and /u/ vowels are shown in Table 5.5. The mean data show atypical values for all measures. In sustained vowels, fewer vowels were produced with interharmonics (for methods, see Chapter 3, Figure 3.11 and Figure 3.12) towards offset. All but one of the sustained vowels was produced with higher perturbation values at onset and without interharmonics. The percentage of vowels affected by interharmonics is similar in the DEAP, Syllables and CNREP tasks, but there is a difference in the voice perturbation measures in these tasks. There are also differences in the mean measurements between those tasks that were recorded within the same session (those marked with an asterisk): of these, the reading task was produced with lowest HNR (dB) and CPP (dB), and with highest DVB (%), DUB (%), jitter (%), shimmer (%) and percentage of interharmonics.

Table 5.5: Mean voice perturbation measures for Kerry in sustained vowels and vowels in words. For sustained vowels, mean measures are given and separate values are given for the initial 100 ms (onset), medial duration (Total duration - (onset+offset)), and the final 100 ms (offset). Mean MPT for sustained vowels was 0.575 seconds. Measures are given for the vowels in each syllable of words produced in DEAP, and Syllables tasks. The table shows the mean measures of voice breaks, unvoiced frames, maximum phonation time, harmonics-to-noise ratio, jitter (%), shimmer (%). The table also shows the number of sustained vowels, and vowels in DEAP, the Newton phrase imitation task (1999), Syllables, CNREP and reading tasks that were produced with interharmonics. The tests marked with an asterisk (*) were recorded within the same recording session.

Test no.	Test name	No. of valid vowels, n=	DVB, %	DUV, %	HNR, dB	jitter, %	shimmer, %	CPP, dB	Number and percentage of vowels produced with interharmonics
19	**Sustained vowels: total duration (T)**	**11**	0.11	1.95	19.31	1.95	5.74	4.55	10/11 = 91%
	Onset (t=100 ms)	*9*	*3.72*	*21.67*	*10.65*	*3.90*	*10.81*	*3.30*	*8/9 = 89%*
	Medial duration (T- (onset + offset))	*11*	*0.00*	*1.55*	*19.81*	*1.96*	*5.39*	*4.36*	*8/11 = 73%*
	Offset (t=100ms)	*11*	*0.00*	*21.25*	*16.04*	*2.52*	*10.41*	*3.95*	*2/11 = 18%*
5	**DEAP***	**65**	6.90	7.53	10.65	1.74	11.49	4.23	48 = 74%
10	**Newton (1999)**	**30**	8.23	19.76	10.03	3.11	13.07	3.29	23 = 77%
11	**Syllables**	**43**	2.57	7.62	13.58	2.20	7.92	3.98	32 = 74%
6	**CNREP***	**24**	6.50	16.43	10.28	1.93	13.43	3.70	19 = 79%
15	**Reading***	**53**	8.38	25.98	8.04	3.01	15.02	3.01	48 = 90%

5.1.5 Learning

Based on an initial, impressionistic assessment, Kerry was able to sing the melodic shape of songs, but her pitching of notes and intervals was inaccurate. Her production of words in songs was fluent but speech sounds were poorly articulated. Kerry's voice quality was sometimes

perceptually harsh when singing and she had difficulty in producing sustained phonation when humming. Kerry's individual targets were as follows:

- match pitch in 5 notes across range;
- sustain phonation for 5 seconds;
- imitate *so-me* model in *hello* and *magic book*; and
- sing key words from memory in familiar songs, with clear articulation.

Kerry tended to sing at a lower F0 than the demonstrated notes, but her accuracy in matching intervals developed after listening to and observing demonstrations (Appendix 3, p. 344 [28]). She had some difficulties in maintaining exhalation to blow pinwheels and to use a humming sound to activate the kazoo (Appendix 3, p. 342[34]), but managed to sustain phonation for up to four seconds when humming or producing vocal sirens (Appendix 3, p. 344 [27]). She became more accurate in recalling words to repeated songs, especially when words were cued with gestures, including Makaton signs (e.g. Appendix 3, p. 350 [23]; p. 351 [3]).

5.2 Motor-music entrainment and imitation

5.2.1 Beat entrainment and extraction

Table 5.6 shows the mean anticipation time (AT) time and inter-tap interval (ITI) time of Kerry's beats during entrainment and extraction tasks. Data were taken for all taps from the onset of playing until she first paused. Her mean ITI (%) was least accurate at the faster tempo (160 bpm) but her steadiness (SD of ITI, ms) and her ability to synchronise (AT, ms; AT, %) were more

Table 5.6: The accuracy of Kerry's timing in beat entrainment and extraction tasks. The table shows mean ITI and AT in milliseconds and as percentage of the target interval, and the SD in milliseconds. Mean results are given for entrainment tasks and for extraction tasks.

	Entrainment			**Extraction**		
Tempo, bpm	**104**	**132**	**160**	**102**	**124**	**211**
Interval time, ms	**576**	**455**	**375**	**589**	**485**	**285**
No. of taps produced =	**21**	**18**	**35**	**16**	**16**	**16**
ITI, ms	557	442	354	587	482	296
ITI, %	103	103	106	99	99	99
SD of ITI, ms	46	32	33	39	14	20
Mean range of ITI, %	4			1		
AT, ms	-110	-22	-82	-60	-4	-15
AT, %	19	5	22	10	8	5
SD of AT, ms	20	30	90	33	38	21
Mean AT, %	15			8		

variable at 104 bpm. When Kerry independently accompanied the *acapella* song, *Can You Tap The Beat?* at three tempi (Table 5.6), her accuracy in timing between beats (ITI, %) remained on average within 1% of the target. Her timing between beats (ITI, %) and her ability to synchronise (AT, %) were more accurate in the extraction task than in entrainment. In both tasks, her performance on both measures was most accurate at the median tempi, but was comparable to her performance at other tempi in terms of variability of timing (SD of AT, SD of ITI).

Figure 5.1 shows an extract of Kerry playing to the beat at 132 bpm. Her timing was closely synchronised in bars 2 and 4 (Tier 2). No metrical accent was perceived in her performance: the intensity of her beats as shown on the waveform confirm that she did not produce the same metrical changes to beat intensity as are present in the recorded pattern.

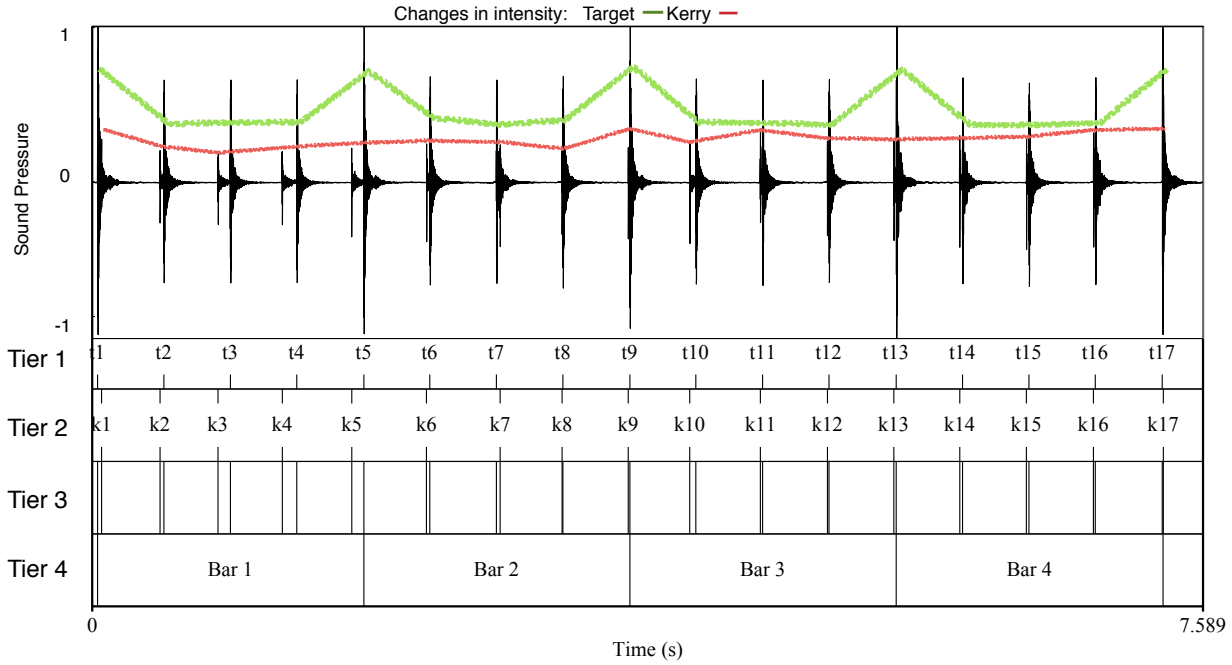


Figure 5.1: Image showing Kerry's rhythmic production during the Beat Entrainment task at 132 bpm. The image shows the target beats (Tier 1), Kerry's beats (Tier 2) and the alignment between them (Tier 3). The bars are numbered in Tier 4. The red contour on the waveform indicates the changes in the intensity of Kerry's beats. The green contour indicates the changes in intensity in the recorded pattern.

5.2.2 Rhythm entrainment

Figure 5.2 shows how quickly Kerry entrained to a simple rhythmic pattern consisting of quarter notes (*ta*) and eighth notes (*tee-tee*). During bar two she began to play the pulse for four beats (k2-k5). From point k6, she reproduced the rhythm and synchronised closely to the target pattern throughout bar three. Kerry continued to play this pattern for a further three bars before she became distracted by external events. The accuracy of Kerry's rhythm in bars three and four is shown in Table 5.7. The mean duration of her short notes (*tee-tee*) and her long notes (*ta)* were close to that of the stimulus in duration (ms) and as a percentage. Kerry was not able to synchronise to a syncopated rhythm, consisting of more complex subdivisions of five beats.

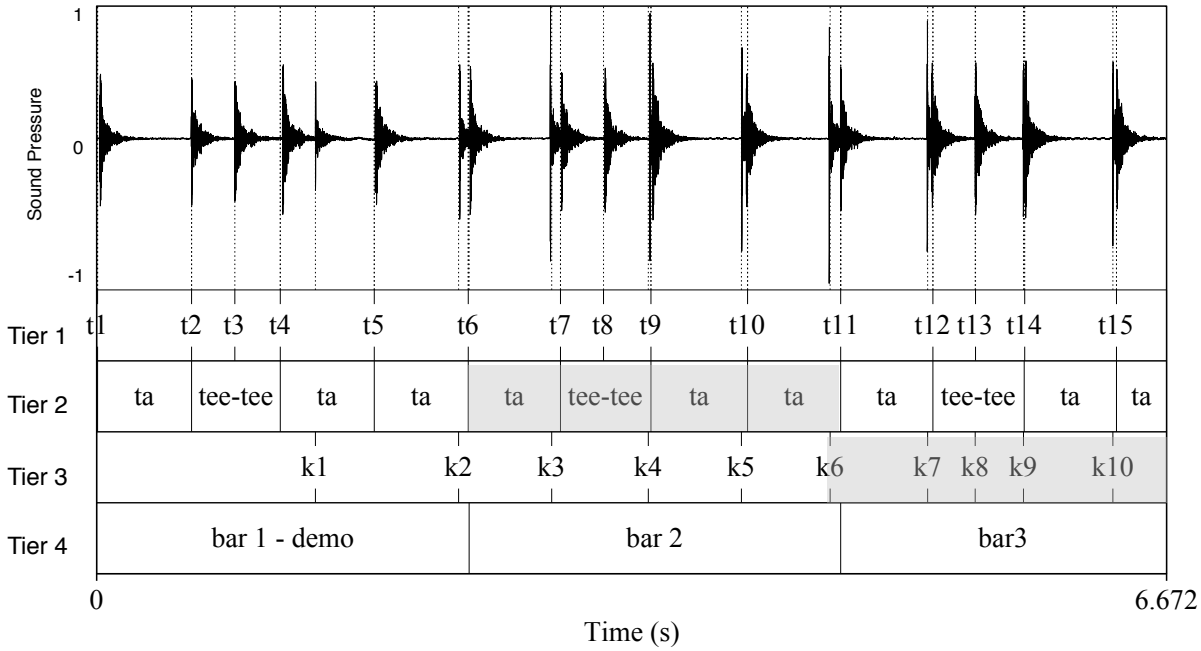


Figure 5.2. Image showing Kerry's rhythmic production in the Rhythm Entrainment task – *ta tee-tee ta ta*. The image shows the waveform of the stimulus and Kerry's imitation: the stimulus is marked in Tier 1 (target timing) and Tier 2 (rhythm names). Kerry's taps are marked in Tier 3 and the bars in Tier 4. Kerry began to play in Bar 1 and paused: she then played 4 beats of near-equal duration close to the beat (k2-k5), before playing an approximation of the rhythm (k6-k10: shaded boxes).

Table 5.7: The accuracy of Kerry's timing in two bars of rhythm entrainment. The table shows the mean value and range of inter-tap intervals (ITI) between the short rhythm (tee-tee) and the contrasting long rhythms (ta) (see Figure 5.2). The target values are shown and Kerry's means are expressed as a percentage of this.

Target: note name and numbers	Mean ITI (ms)	Range of ITI (ms) and SD	Bars 3-4: mean duration, ms	Range of ITI (ms) and SD	Mean accuracy, %
Tee-tee (1-4)	**289.2**	***268-312 (SD: 8.7)***	298.0	*246-350 (42.5)*	103.0
Ta (5-6)	**584.5**	**n/a**	579.8	n/a	99.0

5.2.3 Rhythm imitation

Kerry accurately reproduced five of the fourteen target rhythms (35.7%) in TRaCoL (Table 5.8). She also imitated two correctly in number, but with incorrect duration between claps, and nine were partially correct in duration. Kerry was able to reproduce in number all patterns of two and three claps but reversed the relative duration of number 3, and produced an extra clap for number 6. In patterns consisting of four claps she reproduced two correctly in number and duration, and one correctly (number 7) in the number of claps but with only partial replication of the rhythm. Kerry did not reproduce any rhythms consisting of five or more claps.

Table 5.8: Kerry's accuracy in imitating clapped rhythms in the TRaCoL task (Treharne, 1999). The table shows the accuracy of Kerry's responses in terms of reproducing rhythms in number and duration (N&D), and partial replication of the number (N) or relative duration (D). The notes produced in error are indicated in red text.

Target number of claps	**Target rhythms, with notation**	**Kerry: responses and number of correct responses**	
		N+D	**Partial**
2	**1.** ♩ ♩ **2.** ♫	1. ♩ ♩ 2. ♫	
3	**3.** ♩ ♫ **4.** ♫ ♩ **5.** ♩ ♩ ♩ **6.** ♪ ♩ ♩	4. ♩ ♫	3. ♫ ♩ (N, D) 5. ♩ ♩ (D) 6. ♪ ♩ ♫ (D)
4	**7.** ♫ ♫ **8.** ♫ ♩ ♩ **9.** ♩ ♩ ♫ **10.** ♩ ♫ ♩ **11.** ♩ ♩ ♩ ♩	8. ♫ ♩ ♩ 10. ♩ ♫ ♩	7. ♫ ♩ ♩ (N, D) 9. ♩ ♩ (D) 11. ♩ ♩ (D)
5	**12.** ♩ ♫ ♩ ♩		♫ ♫ (D)
6	**13.** ♫ ♩ ♩ ♫		♫ (D)
7	**14.** ♩ ♫ ♫ ♩ ♩		♫ (D)
Total	**13**	**5**	**7 (D); 2 (N and D)**
Percent correct		**35.7%**	**64.3%**

5.2.4 Syllable segmentation

Kerry was able to independently clap the rhythms of the words *monkey, caterpillar, cat* and the song *Can You Clap The Words?.* She was unable to simultaneously speak the words and clap the syllables in her initial assessment. However, during re-assessment, Kerry was able to independently clap the words to the song *Magic Book* (Figure 5.3), following a demonstration of the song's first phrase (Figure 5.3: Tier 2). Kerry joined in vocally and by clapping in the first phrase at point t3/k1 (Tier 3: shaded section), then continued independently. Her spoken production was hesitant at points k6 and k9 and she omitted a clap for the word *the*. With this exception, she synchronised her claps with the onset of each syllable.

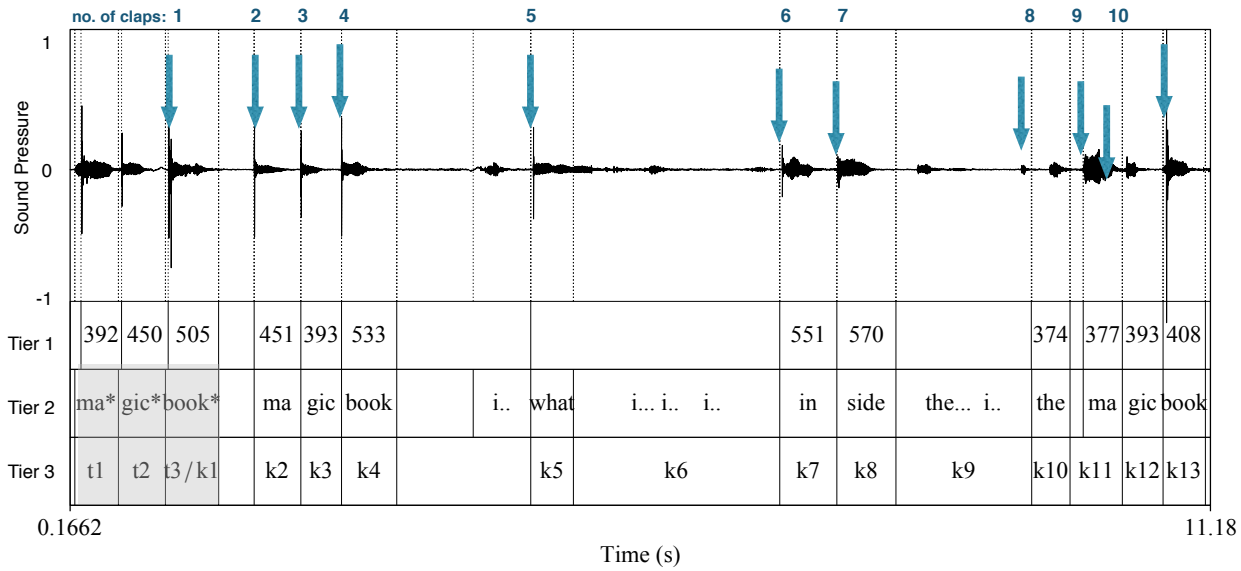


Figure 5.3: Image showing the duration and timing of Kerry's claps relative to her syllables when singing 'Magic Book'. Tier 1 shows the duration between claps (milliseconds): for Kerry, these coincide with the duration of syllables. The syllables the researcher and Kerry produced are shown in Tier 2 and numbered in Tier 3: the asterisks and shaded section indicate the words the researcher produced.

5.3 Voice and prosody in imitation tasks

5.3.1 Sustained vowels

The data for eleven sustained vowels are given in Table 5.9. Shimmer was above normative values for young females (<2.37%, Goy et al., 2013) for vowels except for /u/1. Jitter was within (<0.37%, Goy et al., 2013) or very close to normal limits for six of the vowels: /i/1, /i/2, /u/1, /u/3, /a/2 and /a/3. All vowels were produced with interharmonics (see Chapter 2). Where interharmonics occurred, the vowel was perceived as atypical in quality: harsh, creaky, or diplophonic. Five vowels were produced with audible intermittent diplophonia (/i/3, /i/4, /u/2, /a/1 and /a/3) (e.g. see Figure 5.4): typically, this occurred when there was just one clear or dominant subharmonic.

Table 5.9: The mean fundamental frequency and voice perturbation measures for Kerry's sustained vowels The table shows the mean data for the entire production of eleven sustained vowels. Measures given are for mean F0 (Hz), SD of F0 (Hz), DVB (%), DUV (%), HNR (dB), jitter (%), shimmer (%) and CPP (dB). The duration (milliseconds) of single or complex (2+) inter-harmonics is shown, and their position in the vowel is indicated, as follows: onset (i; t=100 ms), medial (m: the period after onset, but excluding the final 100 ms), and final (t= 100 ms). Where bifurcations are medial but are continuous with onset or offset, this is indicated with +. Two additional vowels (/i/4 and /a/4) were included as Kerry's initial productions were short and she was asked to repeat these and to sustain them for longer. Kerry's production of two vowels (**) overlapped with the demonstration and the contaminated portion (shown in 'duration') is excluded from these measurements. Bold type indicates values that are close to norms for jitter and shimmer, and indicates diplophonic sections.

Vowel	Duration of overlap/ Duration of vowel, ms	Mean F0, Hz	SD of F0, Hz	DVB, %	DUV, %	HNR, dB	Jitter, %	Shimmer, %	CPP, dB	1 Inter-H: Duration (ms) and position	2+ Inter-H: Duration (ms) and position
/i/1	**558**	299.38	0.25	0.00	0.00	20.28	**0.57**	4.65	4.11	67 (m) +27 (f) ms	0.00
/i/2	**574**	268.66	0.25	0.00	0.00	22.67	**0.69**	3.53	3.59	0.00	100 (i) +18 ms (m); 76 ms (m)
/i/3	**510**	181.67	18.84	0.00	12.12	11.02	5.39	7.75	3.40	**88 ms** (m)	100 (i)+122 (m) ms; 100 (f)
/i/4**	282**/842**	293.19	0.62	0.00	0.00	16.78	3.78	8.37	4.03	**284 ms** (m)	98 ms (m) +68 ms (f)
/u/1	**553**	249.45	0.15	0.00	0.00	28.48	**0.34**	**1.86**	4.95	71 ms (m)	100 (i) ms
/u/2	**561**	278.87	0.28	0.00	2.50	16.50	2.84	**2.49**	5.05	**332** ms (m)	100 (i) +18 (m) ms;
/u/3**	294**/470**	286.32	0.35	0.00	0.00	26.39	**0.39**	**2.77**	4.77	n/k: onset contaminated	
/a/1	**707**	314.32	1.37	0.00	3.39	12.99	3.94	10.87	4.95	**0**	67 ms (i);
/a/2	**529**	247.95	0.89	0.00	0.00	20.68	**0.69**	3.85	4.67	0	100 (i) + 29(m) ms
/a/3	99**/420**	233.89	0.55	0.00	0.00	21.51	**0.35**	8.21	5.61	**230** (m) ms; **100** ms (f)	78 (i) ms
/a/4	**660**	306.47	1.07	1.25	3.39	15.11	2.48	8.75	4.95	**89 (m) ms, 225 (m) ms**	**100 (i)**
Mean	**580**		**2.24**	**0.11**	**1.95**	**19.31**	**1.95**	**5.74**	**4.55**		
No. of vowels affected										**7/10**	**9/11**

Example: /a/1

The vowel /a/1 was selected as an example of Kerry's poorest production in terms of pitch stability and voice quality in sustained vowels. The poor production highlights a difficulty in both initiating and sustaining phonation. The vowel was produced with low HNR and high jitter (Table 5.9). The waveform and spectrogram of Kerry's vowel /a/1 are shown in Figure 5.4. The waveform shows subharmonics (sections 1 and 3: e.g. see red arrows), that coincide with audible diplophonia. Multiple interharmonics are also visible between the main harmonics and subharmonics in sections 1-4. The intensity of the subharmonics reduces in section 2. This coincides with an increase in pitch and a change in the contour of F1. The detail (Figure 5.5) shows a secondary waveform between glottal pulses in section 3, that coincides with a clear subharmonic.

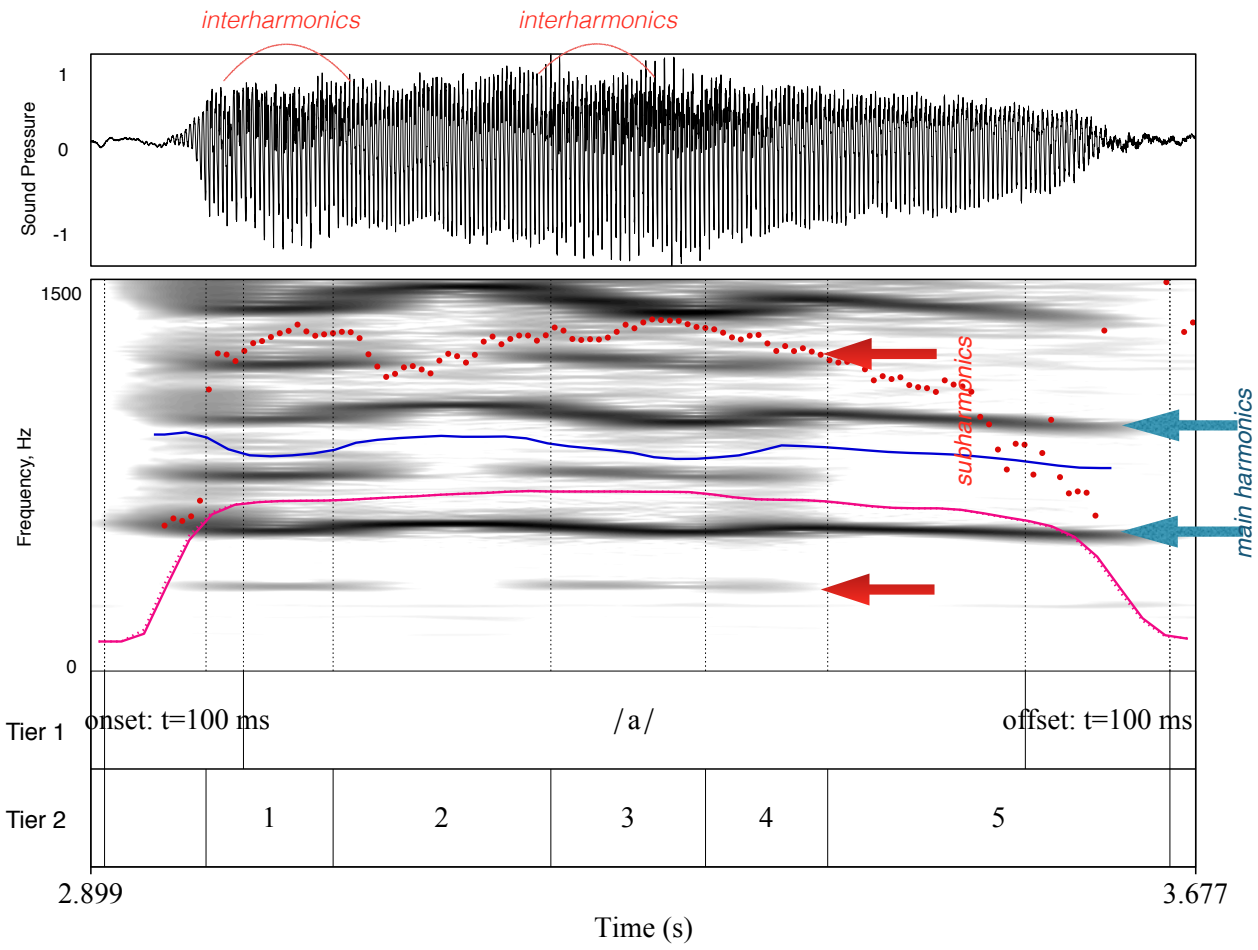


Figure 5.4: Waveform and annotated spectrogram showing Kerry's /a/ vowel at G4 (/a/1). The image above shows the waveform, narrow-band spectrogram, pitch contour (blue), intensity contour (pink) and the first formant frequency contour (red dots) of Kerry's imitated /a/ vowel. For the image, the narrow band spectrogram settings were amended to 1500 Hz and 150 ms bandwidth, dynamic range 45 dB. Pitch settings were changed to 50-250 Hz and intensity settings were 40-100 dB. Tier 1 shows the onset and offset of the vowel. Tier 2 shows the sections referred to in the text. Interharmonics and subharmonics are visible in between the main harmonics in sections 1-4 (e.g. see red and blue arrows on the spectrogram). Additional harmonics are also visible on the waveform in sections 1 and 3, as indicated by the red contours above the waveform.

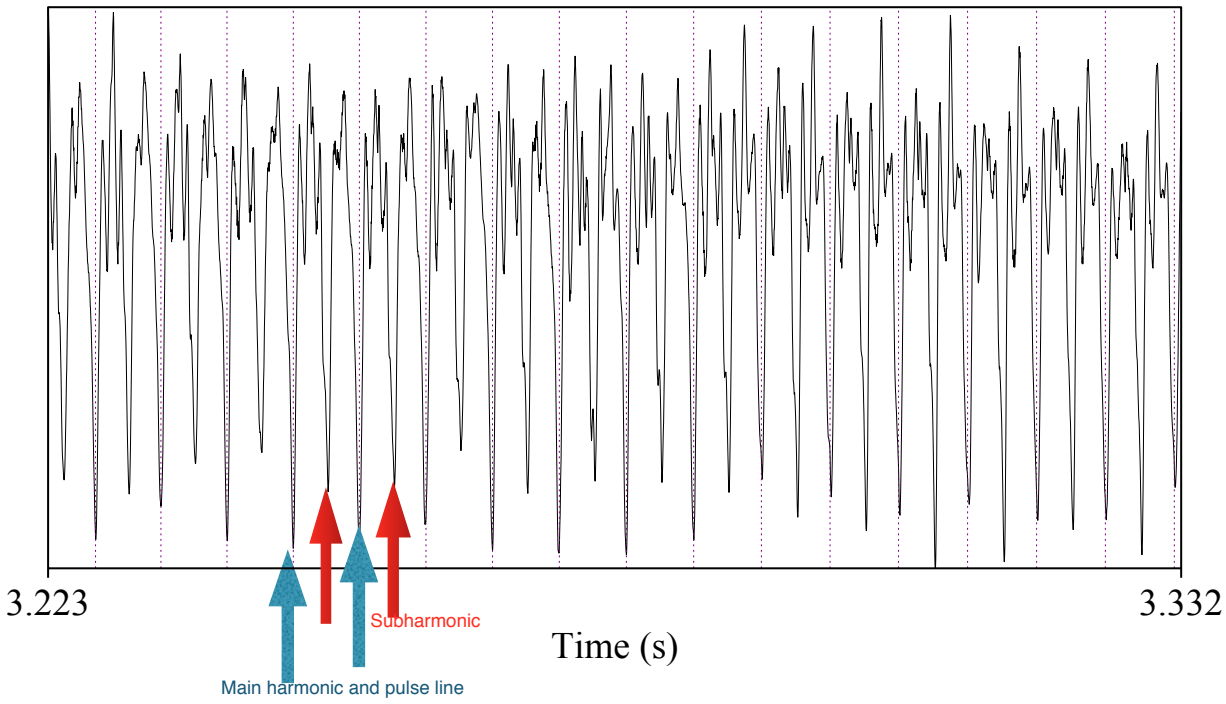


Figure 5.5: Detail of the waveform from Kerry's /a/ vowel at G4 (/a/1), section 3. The image shows a detail of the waveform of section 3 (Figure 5.4), and the pulse lines, as identified in Praat. The settings are the same as for Figure 5.4. A second harmonic (e.g. see the red arrows) is visible between the pulse lines (e.g. indicated by the blue arrows). This coincides with perceived diplophonia.

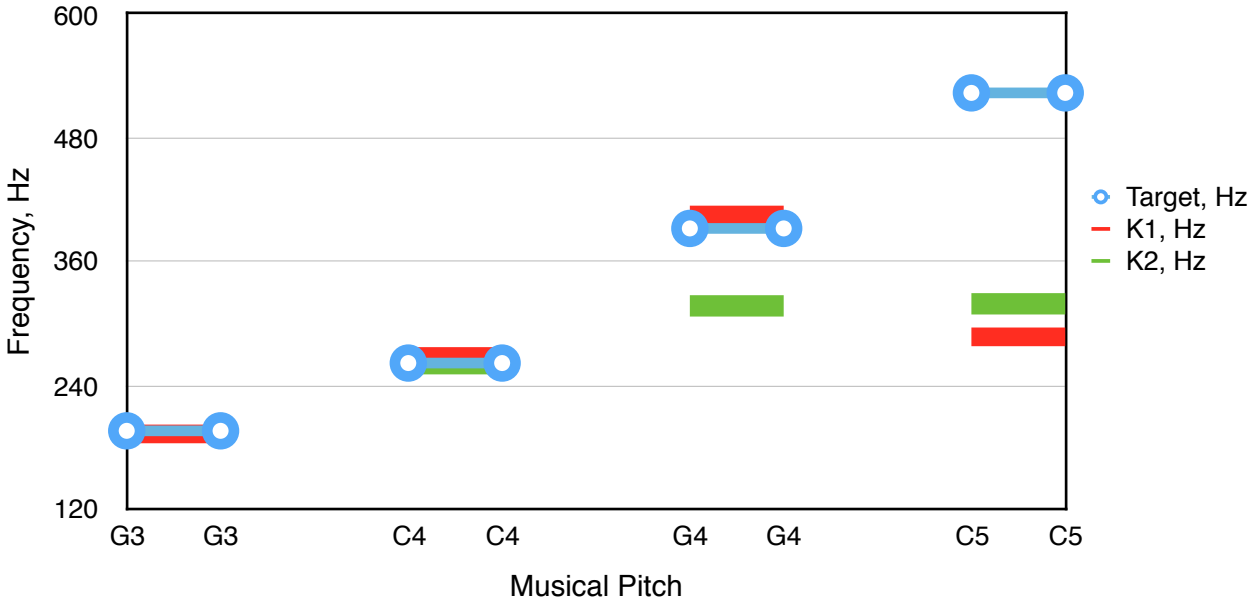
5.3.2 Pitch and interval matching

Figure 5.6 shows Kerry's accuracy in matching single pitches of two notes, over two octaves, on the sound /la/. In the first assessment (K1), the notes G3 (196 Hz) and C4 (261.63 Hz) were both within 3 Hz 0.3 ST of the target and G4 (392.00 Hz) was 13 Hz (0.6 ST) higher than target. Kerry did not attempt to match C5 (523.25 Hz) at absolute pitch, reproducing an average of 287 Hz, approximately a semitone above C4, between C♯ 4 and D4. The second assessment (K2) shows a less accurate performance. Kerry did not produce G3 (196 Hz) at all, and both G4 (392.00 Hz) and C5 (523.25 Hz) were produced at comparable pitches to each other (317 Hz, 319 Hz). The note C4 was produced precisely on pitch, at 261 Hz.

In the initial assessment of interval-matching (Figure 5.7) Kerry was unable to match the intervals or the absolute pitch of notes. However, her imitation of the s*o-me* interval of notes G-E (392-329 Hz) was too small by just 0.5 ST. Also Kerry's descending *la-so* interval in the target *so-la-so* (F♯- G- F♯) was just 0.3 ST larger than the target interval. For the first four tasks in this test, Kerry appears to have produced the last note from the demonstration as her starting pitch.

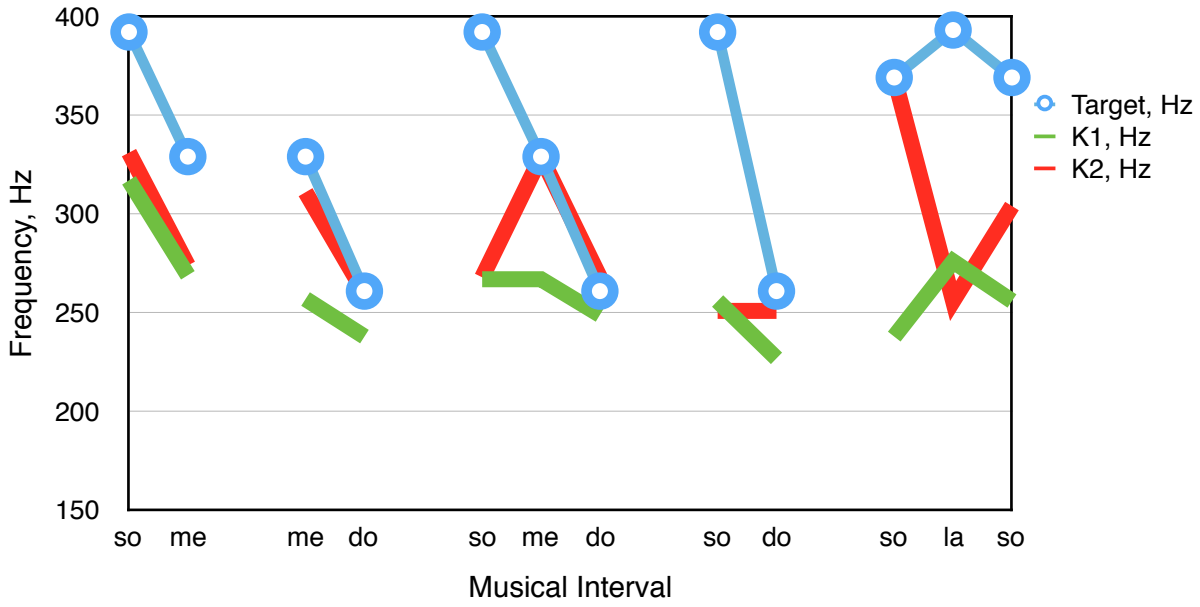
With the exception of *so-me* and *so-do*, Kerry's imitations in the second assessment (Figure 5.7: K2) were less accurate in terms of relative interval than her first attempts; however, they were

Figure 5.6: Graph showing Kerry's accuracy when matching pitches during initial and final assessments. The fundamental frequency of the target pitch (in Hz) is shown in blue and Kerry's mean F0 is shown for her first assessment (K1) and final assessment (K2), which were six weeks apart.



closer to the target in terms of absolute pitch. The *me-do* portion of *so-me-do* matches both absolute pitch and interval, as shown by the overlap of the green line on the blue target line. The pattern *me-do* matches closely in starting and finishing pitch and in interval but was one semitone below the target of 329-261 Hz (E-C), at 311-260 Hz. Kerry's ability to produce intervals did not improve with visual support, and she was not able to sing a scale from a visual image during her initial assessment.

Figure 5.7: Graph showing Kerry's accuracy when matching melodic intervals during her first and final assessments. The fundamental frequency of the target pitch (in Hz) is shown in blue and Kerry's mean F0 is shown for productions in her first assessment (K1) and final assessment (K2), which were six weeks apart. The target pitch is shown in blue.



5.3.3 Words of increasing syllable length

Twelve (44%) of Kerry's words in the syllables imitation task were dysfluent, which decreased her mean speaking rate (Table 5.10) and affected the rhythm. Kerry's mean rate of production of fluent sections was similar to the demonstration (95%) but she produced fewer consonants as the target syllable length increased, and she did not reproduce all syllables in six of the three-syllable words (67%).

Table 5.10: A summary of the segmental, prosodic and vocal features of Kerry's imitation of words in the Syllables test, arranged by number of target syllables. The table shows the number of words matched for pitch contour, relative intensity and the number of produced consonant sounds, as a percentage of the target. The mean speaking rate (syllables per second) is shown for the demonstration and for Kerry's production, which is also calculated as a percentage of the former.

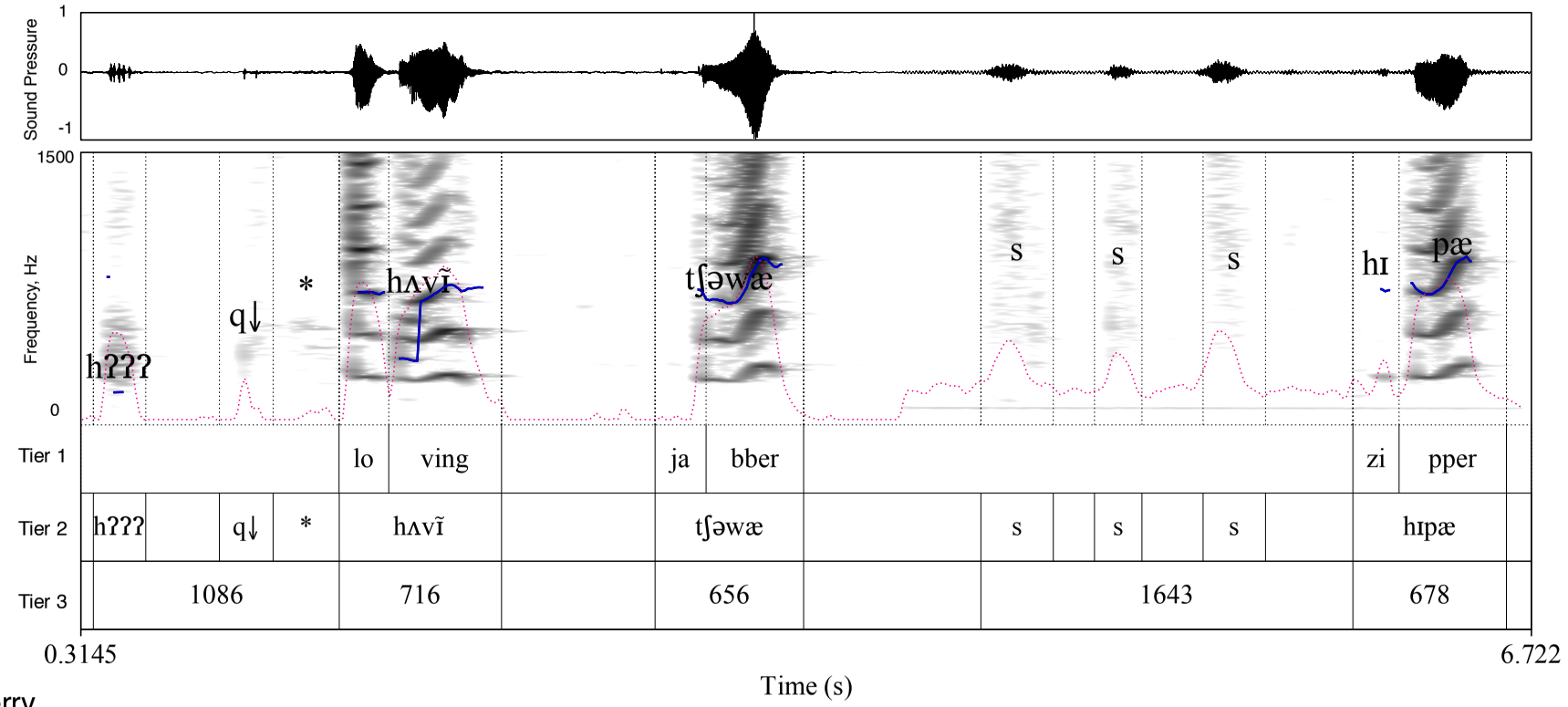
Target no. of syllables	Mean speaking rate, syll/ sec	Mean pitch range, ST	n=	No. of words matched for pitch contour	No. of words matched for intensity	No. of syllables produced	No. of words affected by dysfluency	No. and percentage of consonants produced	Mean speaking rate, syll/ sec	Mean speaking rate including dysfluency	Mean pitch range, ST
1	**1.99**	**4.75**	**9**	9	9	9/9	2	16/20 = 80%	1.66	0.92	3.0
2	**2.54**	**4.07**	**9**	8	5	18/18	5	20/26 =77%	3.02	1.46	2.6
3	**4.18**	**7.40**	**9**	7	7	20/27	5	21/37 = 57%	3.60	2.11	4.3
Mean	**2.90**	**5.41**		**7.7**	**7**		**44%**		**2.76 syll/sec**	**1.5 syll/sec**	**3.3 ST**
Mean, as a percentage of demo				**89%**	**70%**	**87%**	**44%**	**71%%**	**95%**	**52%**	**61%**

When the voice perturbation measures were arranged by both the number of syllables produced and by syllable position (Table 5.11), HNR increased as the target number of syllables increased, and CPP decreased. Words of one-syllable were produced with fewer breaks in phonation (DVB, %), lower DUV values and high CPP values (Table 5.11). In multi-syllabic productions, vowels in the initial syllable position were produced with a higher DUV values and lower CPP values than vowels in final-position vowels, but with a higher degree of HNR (Table 5.11).

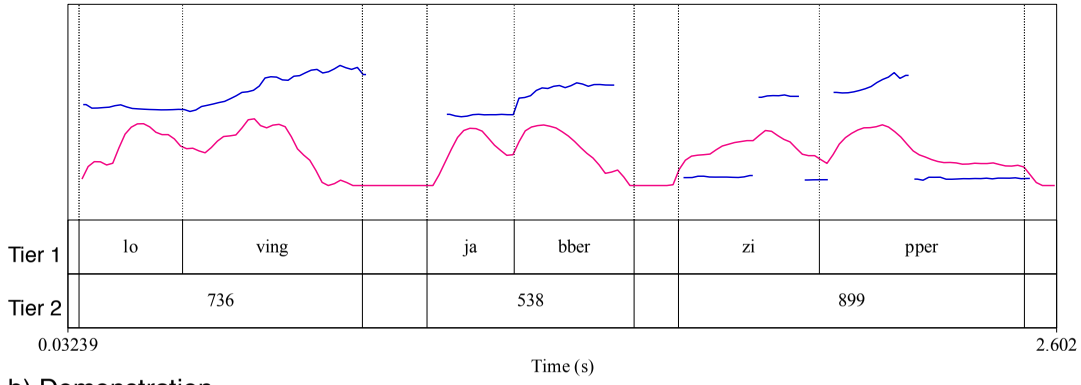
Table 5.11: Voice perturbation measures in the Syllables task according to the number of syllables produced by Kerry and the position of the vowel in the word. The table shows mean data and SD for DUV, (%), DVB (%), HNR (dB), and CPP (dB). The data are arranged by the target number of syllables, and are also separated into data produced in the first syllable and the final syllable. The words 'soften' and 'softening' were excluded from the data as the samples contained extraneous noise; of the words of 3 target syllables, one was produced with one syllable and two with three syllables.

Number of target syllables	Position of syllables	n=	DUV, % (SD)	DVB, % (SD)	HNR, dB (SD)	CPP, dB (SD)	No. of vowels produced with interharmonics
1	**all**	**9**	3.25 (4.74)	1.16 (3.31)	12.46 (2.78)	4.85 (1.04)	5/9 = 55%
2	**all**	**16**	16.51 (23.37)	3.47 (10.05)	13.84 (3.24)	4.22 (1.17)	11/16 = 68%
	1st	*8*	*28.64 (35.63)*	*8.82 (16.37)*	*14.75 (3.69)*	*3.97 (1.25)*	
	final	*8*	*4.38 (6.88)*	*3.53 (9.34)*	*13.08 (2.71)*	*4.48 (0.81)*	
3	**all**	**18**	13.9 (11.32)	0	14.22 (3.93)	4.16 (1.03)	16/18= 89%
	1st	*8*	*17.50 (13.02)*	*0*	*16.00 (2.98)*	*3.52 (0.45)*	
	final	*7*	*10.34 (14.63)*	*0*	*12.43 (5.67)*	*4.80 (1.49)*	

An example of the first three words from the task (Figure 5.8) illustrates the effect that dysfluency had on her production and allows comparison to the demonstration. The words *loving* and *zipper*



a) Kerry



b) Demonstration

Figure 5.8: Waveform and annotated spectrogram showing the stimulus (above) and Kerry's imitation of three words of two-syllables. The above image (a) shows the narrow-band spectrogram (0-1500 Hz, 150 ms, 50 dB) of Kerry's production of words of two syllables. Pitch contour is shown in blue (range 50-450 Hz), phonetic production in black, against the pitch contour. The pitch break in 'loving' coincides with a change to creaky phonation; the asterisk marks Andrew's voice (Chapter 4), in the background. The intensity contour in shown in pink (45 dB-100 dB). The duration of words is shown in Tier 3 and the duration of Kerry's dysfluency at the onset of words is marked. The researcher's demonstration of these words is shown below (b), with target syllables shown and the duration of the word. As the recording was quieter, the overall intensity contour is reduced and pitch contour is affected in places by environmental noise.

were produced with over one second of hesitation and dysfluency: the initial syllable of both words was short in comparison to the first syllable of the demonstration (Figure 5.8b). The initial syllables in all three words were produced with lower intensity than the second (Figure 5.8a): by comparison, the demonstrations (Figure 5.8b) were produced with a less distinct change in intensity between syllables.

Example: 'Love, loving, lovingly'

Figures 5.9 to 5.11 show Kerry's productions of *love*, *loving* and *lovingly*. The vowels in each syllable were produced with an increase in the number of interharmonics as the target syllable length increased. This set was used because Kerry's voice quality changed perceptually as the target syllable length increased, and because there was a relative improvement in phonation in the three-syllable word, in line with the mean data (Table 5.11).

Love: [la]

There is noise visible on the spectrogram between the main harmonics (Figure 5.9), that reflects a perceived harsh vocal quality (section 1) and a breathy quality at section 3. The vowel during section 2 is perceptually modal, although there is still additional noise in the waveform and on the

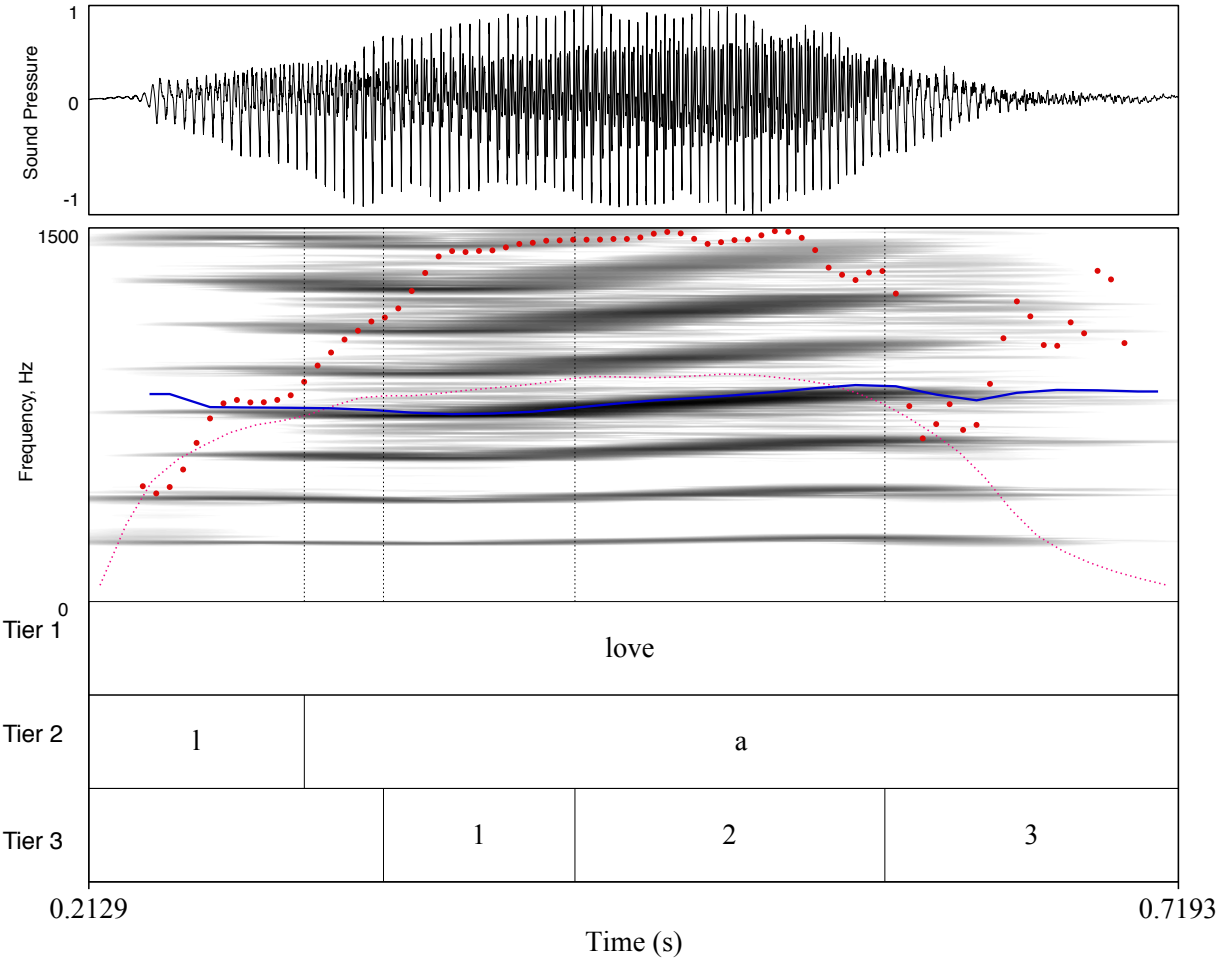


Figure 5.9: Waveform and annotated spectrogram showing Kerry's imitation of 'love'. The image shows the waveform and narrowband spectrogram (1500 Hz; window length 80 ms, 45 dB) of the word *love*, the pitch contour (50-450 Hz) and the intensity contour (45-100 dB), in pink. The first formant frequency is shown (red dots). Kerry's phonetic production is marked on Tier 2, and points of interest are numbered in Tier 3. These are discussed in the text.

spectrogram. The waveform increases and decreases gradually in intensity, and gradually increases in pitch between onset and offset.

Loving. Narrow transcription: [hʌ**vĩ**]

After initial dysfluency (see Figure 5.8), the onset of *loving* was perceptually harsh and marked by an increase in intensity near the onset (Figure 5.10: section 1). The second syllable was diplophonic at point 2 and the final syllable was produced as a nasalised mid-central vowel (Figure 5.10: Tier 3, section 3): the offset of the vowel was breathy (section 4).

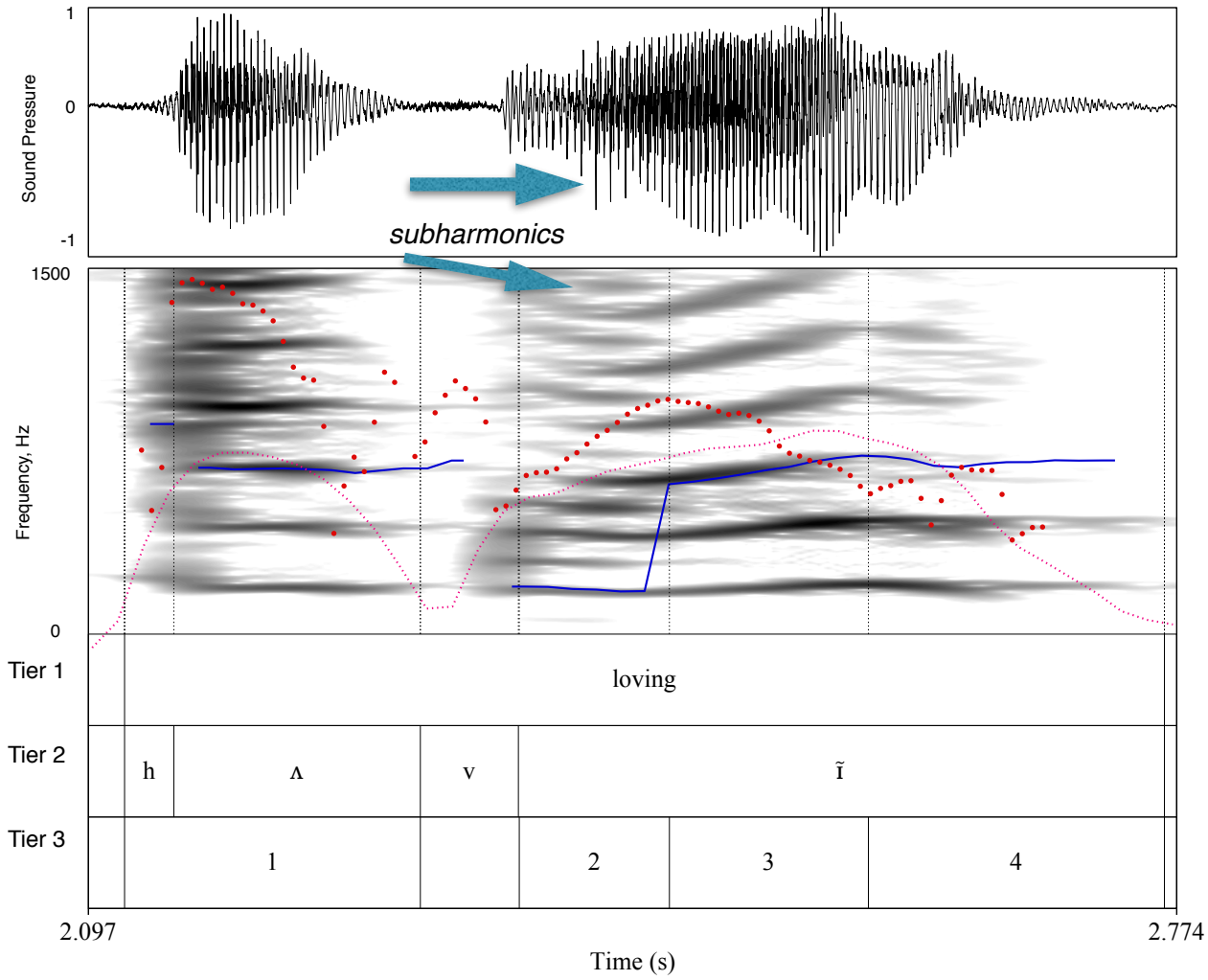
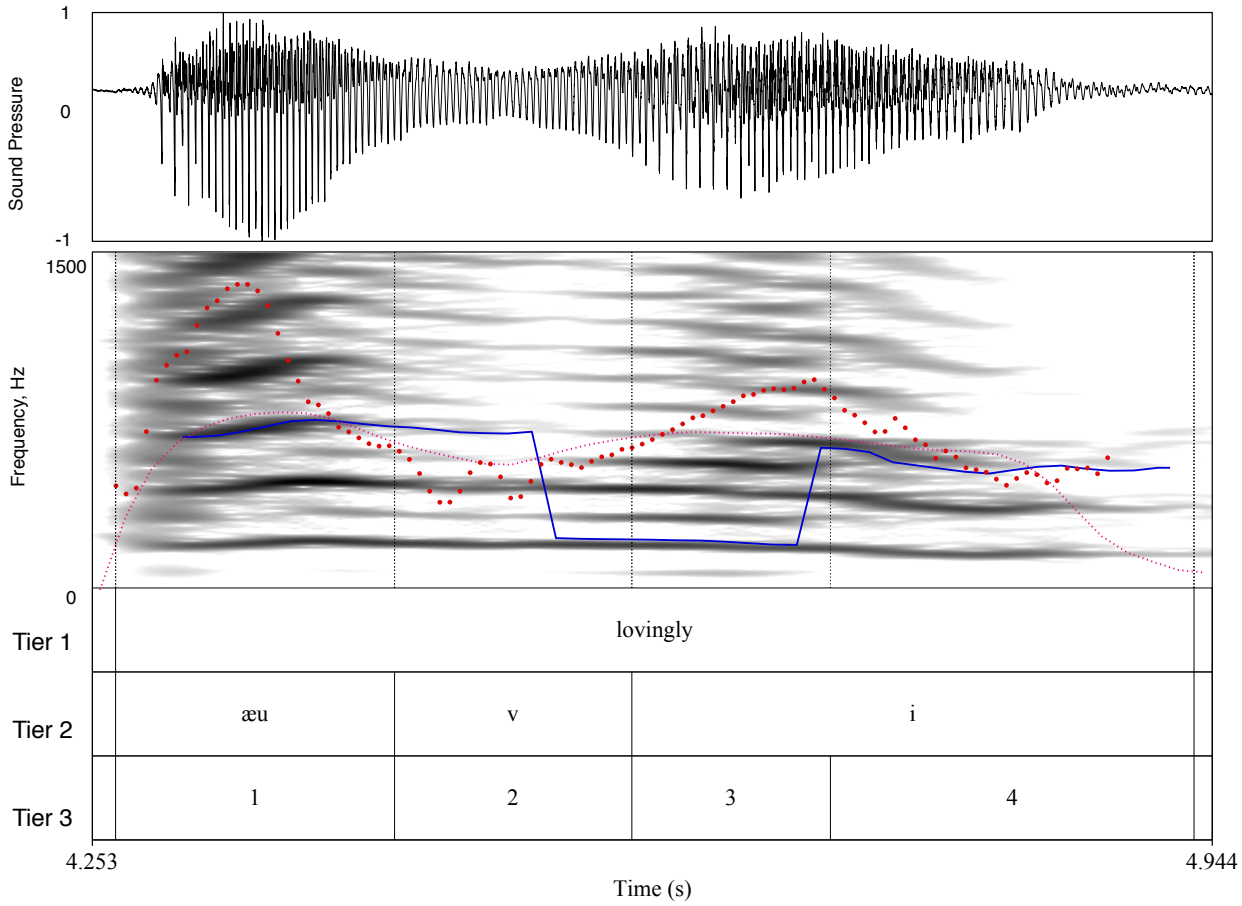


Figure 5.10: Waveform and annotated spectrogram showing Kerry's imitation of 'loving'. The image shows the waveform and narrowband spectrogram (1500 Hz; window length 80 ms; 45 dB) of the word 'loving', the pitch contour (50-450 Hz) is shown in blue, and the intensity contour (45-100 dB), in pink. Kerry's phonetic production is marked on Tier 2. The pitch contour shows a break during section 2 that coincides with perceived diplophonia, and additional sub-harmonics on the spectrogram. There are additional peaks on the waveform in this section, too, as indicated by an arrow. Other points of interest are numbered in Tier 3. These are discussed in the text.

Lovingly. Narrow transcription [**æu**vɪ̃]

The three-syllable word was produced as two syllables, with 1.59 seconds of initial blocking. The onset was harsh (Figure 5.11a: section 1) and the first syllable remained harsh in quality. The detail (Figure 5.11b) shows a subharmonic on the spectrogram that begins in section 2 and increases in

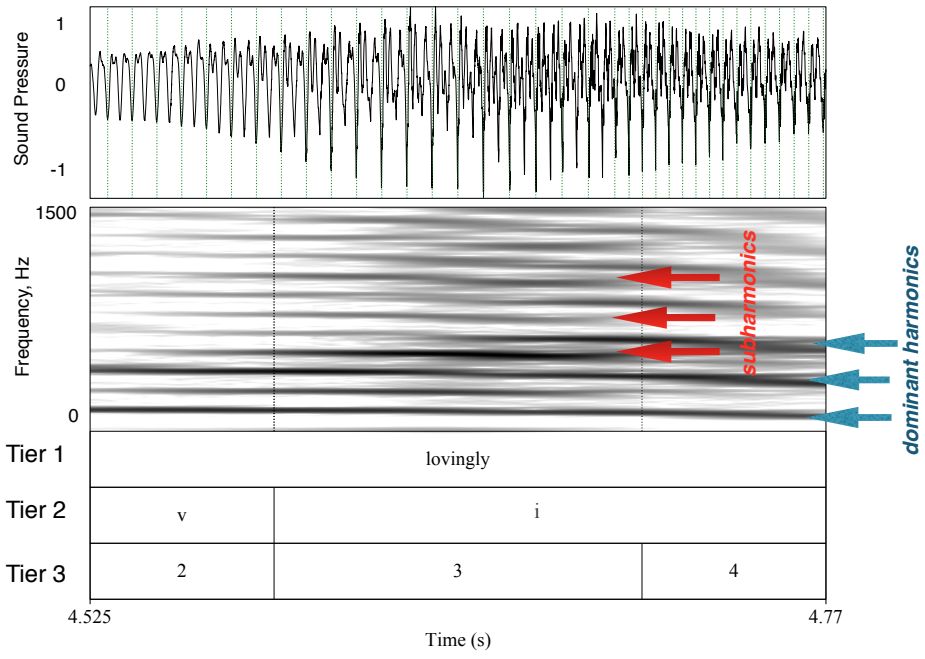


a)

Figure 5.11: Waveform and annotated spectrogram showing Kerry's imitation of 'lovingly'. The image shows the waveform and narrowband spectrogram (0-1500 Hz; window length 80 ms; 45 dB) of the word 'lovingly', after 1.59 s of initial dysfluency. Figure a) (above). Figure b) (below) shows a detail of sections 2, 3 and 4.

In Figure a), the pitch contour (50-450 Hz) is shown in blue, the intensity contour (45-100 dB), in pink and the first formant is shown in red dots. Kerry's phonetic production is marked on Tier 2: the onset of her second syllable was perceptually creaky and diplophonic (Tier 3: sections 2 and 3). Other points of interest are numbered in Tier 3 and are discussed in the text.

b) Detail of the waveform and spectrogram (as above) of section 3, and parts of sections 2 and 4. The pulse lines are also shown on the waveform. Subharmonics are visible on the waveform, between pulse lines, in sections 2 and 3; and on the spectrogram. The lower three subharmonics are indicated by red arrows, and the corresponding dominant harmonics are indicated by blue arrows.



intensity and complexity during section 3: the onset of this subharmonic in section 2 coincides with perceived diplophonia. The waveform at the onset of section 4 (Figure 5.11b) is similar in complexity but is perceptually harsh, without diplophonia. The final vowel was breathy and nasalised towards offset (Figure 5.11: section 4).

5.3.4 Nonword imitation

Data for each nonword Kerry reproduced are shown in Table 5.12. Kerry reproduced a maximum of four syllables for all words, and reproduced just the final two syllables of two five-syllable words (*defermication, reutterpation*). Despite omitted syllables, the pattern of syllabic stress of strong-weak syllables was similar to the demonstration in words of two target syllables (*hampent*, *pennel*) and in words of four target syllables and above. However, both words of three target syllables were reproduced with emphasis on the final syllable, rather than on the initial syllable. Despite omitted syllables, Kerry reproduced the pitch contour of all but one word (*sepretennial*) but her pitch range was smaller than the demonstration. Multiple interharmonics were present in all words, and three

Table 5.12: Prosodic features of the demonstration of nonwords (CNREP) and of Kerry's imitation. The table shows the perceived stress, and the segmental, rhythmic and melodic features of the demonstrated word and of Kerry's imitation. Kerry's production of each target word is shown with a phonetic transcription, with IPA symbols. Kerry's pitch contour is given and this is marked to show whether her production was the same (S) or different (D) to the stimulus. The percentage of phonation in vowels that was produced with interharmonics is also shown. The footer rows show the mean values for the demonstration and imitations: pitch range (ST), speaking rate (syllables/second) and totals, expressed as a percentage of the demonstration.

Demonstration				**Accuracy of production**					**Proportion of vowels produced with inter-harmonics %**
Target word and stress		**Mean range, ST**	**Duration, ms**	**Phonetic transcription with IPA transcription**	**No. of syllables**	**Pitch range, ST**	**Pitch Contour**	**Duration, ms**	
1	**Ha**mpent	3.2	647	**æ̃**p**ʊ̃**s	2/2	2.9	S	993	100
2	**Pe**nnel	5.4	735	**ɛ**wi	2/2	2.6	S	649	77
3	**Do**ppelate	7.5	895	ʌ̲ʊtɪ**jəʊ**	3/3	4.1	S	865	66
4	**Glis**tering	8.5	968	ɪ̲t^{h}ə**wi**	3/3	4.8	S	1171	80
5	**Blon**terstaping	6.9	1452	**ɛ̃s**jɪl	2/4	4.6	S	964	73
6	Cont**ram**ponist	5.8	1222	tʃ…**tʃæ̃**mpɪn	2/4	4.4	S	961	100
7	Per**plis**teronk	8.4	1127	ʔ…**ɪst**əklɔ̲**ŋ̲**q̲	3/4	4.5	S	897	100
8	Defer**mica**tion	5.6	1415	**kɐ̃ɪ**ʃən	2/5	3.5	S	959	72
9	Reutter**pat**ion	5.4	1152	**peɪ**ʃən	2/5	3.4	S	859	80
10	**Sep**retennial	4.3	1087	nt**ɛt**ʊə̲n̲	4/5	4.4	D	1051	83
Mean value		**6.1**	**3.5 syll/sec**		**25/37**	**3.9**	**9/10**	**2.5 syll/sec**	**83%**
Percentage of demonstration				**28.6% consonants correct**	**67%**	**64%**	**90%**	**73%**	

words (*hampent, contramponist, perplisteronk*) were produced with no phonation, as indicated by the presence in all vowels of subharmonics (Table 5.12). Two words were produced with dysfluency (*contramponist, perplisteronk*).

There is no clear link between an increase in the target syllable length and a decrease in the percentage of phonation produced with interharmonics (Table 5.13), or between voice measures and increasing syllable length. Mean CPP was lowest in words of four syllables, but HNR and the percentage of DUV and DVB in vowels was also highest in four-syllable words.

Table 5.13: Mean voice perturbation measures of Kerry's vowels in CNREP, arranged according to the number of syllables in the target word.The table shows mean data and SD for Mean F0 (Hz), DUV (%), DVB (%), HNR (dB), CPP (dB), and the mean proportion of vowels and voiced phonemes that were produced with interharmonics, expressed as a percentage of the total duration of the word. The data is arranged by the target number of syllables.

No. of target syllables	No. of words, n=	No. of vowels, n=	Mean F0, Hz (SD)	DUV, % (SD)	DVB, % (SD)	HNR, dB (SD)	CPP, dB (SD)	Mean proportion of vowels with interharmonics, %
2	**2**	**4**	169.15 (72.19)	23.87 (25.87)	2.36 (4.73)	9.78 (3.64)	3.34 (1.18)	89.00
3	**2**	**5**	188.04 (75.06)	5.04 (5.05)	0.00	9.28 (4.00)	3.52 (1.11)	73.00
4	**3**	**7**	216.19 (43.23)	29.43 (27.88)	22.27 (27.88)	11.14 (3.53)	3.16 (0.91)	91.00
5	**3**	**6**	189.29 (90.03)	9.84 (18.95)	0.00	10.73 (3.57)	4.32 (0.93)	78.00

Example: **Do**ppelate [ʌ̟ʊtɪ**jəʊ**]

The nonword *doppelate* was selected as an example because Kerry partially replicated the melodic contour and produced the correct number of syllables but her stress and phonetic production were markedly different to the demonstration.

After a pause of 1.417 s (Figure 5.12: Tier 2, between sections 1-4), Kerry produced three syllables. Kerry's production of the word was shorter in duration than the target (Figure 5.12: Tier 2): however, Kerry did not produce the initial and final consonant sounds. The relative duration of Kerry's first two syllables (Figure 5.12: Tier 2, sections 4 and 5) retain the rhythm of the target (Tier 2, section 1). Kerry produced the word with an overall decline in pitch contour (Figure 5.12: sections 4-6) that imitates the demonstration (section 1), but the changes between syllables differ to the demonstration: the researcher's production reduced in pitch between the first and second syllable, but the change in the pitch of Kerry's second syllable was less marked (section 5: blue contour). Whereas the researcher's production reduced in intensity with each syllable (Figure 5.11: section 1), Kerry increased the intensity on her third syllable (section 6; pink contour). The overall effect is a difference in perceived stress. Voice quality was perceptually modal, but only the /ɪ/ vowel and the offset of the /əʊ/ were produced without interharmonics.

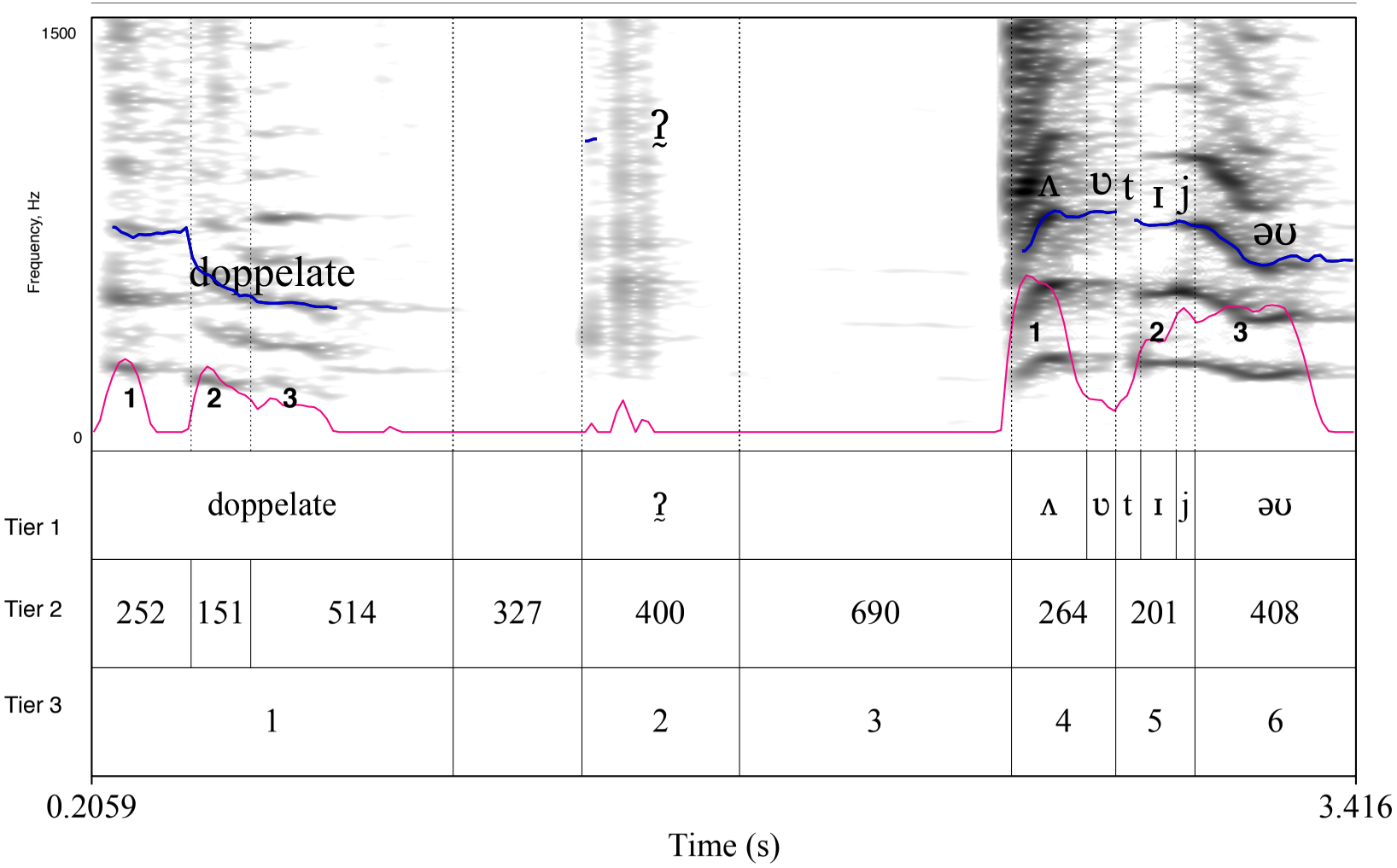


Figure 5.12: Annotated spectrogram showing Kerry's imitation of 'doppelate'. For both the stimulus and Kerry's imitation, the image shows the waveform, narrow-band spectrogram (0-1500 Hz;window length 150 ms; 35 dB), phonetic production, and syllable duration (milliseconds) and division between syllables. The phonetic production, pitch contour (blue: 50-450 Hz) and intensity contour (pink: 35-100 dB) are shown on the spectrogram. The syllables are also numbered to mark changes in intensity. Tier 1 shows the demonstrated word and Kerry's phonetic production. Kerry's production began at section 2 (Tier 3) with a glottal noise, followed by a pause (section 3). The duration between the end of the demonstration (Tier 3, section 1) and the onset of Kerry's imitation (Tier 3, section 4) is 1.417 seconds.

5.3.5 Imitation of phrases

Kerry omitted key words when imitating phrases and she omitted syllables from some words (Table 5.14). Two phrases were affected by dysfluency (numbers 2 and 7) and phrase 4 was produced with initial glottal stops. Kerry's phrases were lower in pitch range than the demonstration but phrases 5, 6 and 7 were within one semitone of the range of the demonstration. Missing syllables and words affected the perceived rhythmic grouping of syllables, but Kerry reproduced the melodic contour of all phrases except for phrase 2; and with the exception of phrase 6, Kerry reproduced the primary stress. In phrase 3, despite the missing word 'you', Kerry reproduced the prosodic contour and perceived rhythmic grouping of syllables: this phrase was produced with an unintelligible initial three syllables and the highest percentage of modal phonation.

Table 5.14: Prosodic features of the demonstration and of Kerry's imitation of phrases (Newton, 1999). The table shows the target phrase, which is marked with syllabic stress and the perceived grouping of words, and with the pitch range (ST) of the demonstration. Primary stress is indicated in bold type: secondary stress is underlined. The similarity of the rhythm, pitch range (ST) and pitch contour is indicated by S (same), D (different), and (P) partial replication. The percentage of atypical phonation as indicated by the presence of interharmonics on the spectrogram is also shown. ExtIPA symbols and VoQS indicate Kerry's voice quality, which was creaky in parts (V̰).

Demonstration: words , syllable stress, and rhythmic grouping	**Pitch change, ST**	**Kerry's production: words, IPA transcription and syllabic stress, and rhythmic grouping**	**Rhythm**	**Pitch change, ST**	**Pitch contour**	**Proportion of voiced segments produced with interharmonics, %**
_ /__ __ __ _ / ___ 1.The /red car went a/ **way**	7.3	__/ ____ __ __ / _ _/ _____ {V̰ ŋ /kha ɪʔ ʌʔ / we ə/ V̰} w**ei** the car it (uh) went aw**ay**	D	2.8	P	83.0
__ / ___ _ __ /___ _ ___ 2.They /robbed the bank /ˈ**ye**sterday	8.9	__ __ __ /__ __ {V̰ ə bæχ ɪs…s..s.s.. /**es**dei V̰} ˈthe bank yes..ss.. ˈ**yes**… s(ter)dayˌ	D	6.6	D	63.0
_/ __ _ _/ __ 3.You/ must clean your/ **teeth**	10	__ _ _/ __ **fe**ɪ̃fən jə **/tif** (unknown) your **teeth**	S	2.9	P	49.0
__ _ _ _/ __ 4.Mary's shoes are / **clean**	9.3	: ʔ.ʔ. / klɪ̃	n/a	6.5	P	65.0
_ __ _ / __ _ __ /_ _ __ _ 5.I gave the/ elephant/ a ba**na**na	7	__ _ _/ _ _ /__ _ {V̰ ei ʔɪ ʔi efə…V̰}/**an**əs I give the/ e(le)phant / **an**as	D	6.2	S	65.0
__ _ / _ _ __ 6.**Clare** ate/ all her lunch	6.7	__ _ _ /__ ʔŋ klɛɹ{V̰ ɛ̰ əʰə/ **lʌntʃ** V̰} Clare ate her/ **lunch**	D	6.6	P	75.0
_/ __ _ _ _ /__ __ 7.My uncle is a **far**mer	5.9	__ __ __ __ / __ __ {V̰ **ʌχɤ** ɪ əf…ff/ **fa**mə V̰} uncle is a f…/**fa**rmer	D	5.1	P	80.0
Mean	**7.8**	**4.4 syll/second (Demo: 6.6 syllables/second)**		**5.2**		**69.0%**
Percentage of demonstration		**69.0%**	**14.0%**	**67.0%**	**71.0% partial**	

Example: **Clare** ate all her lunch: [ʔŋ klɛɹ {V̰ ɛ əʰə **lʌntʃ** V̰}]
Phrase 6 (Figure 5.13) was chosen as an example as this was close to the demonstration in pitch range but was different in perceived stress and rhythmic grouping (Table 5.14).

Figure 5.12 shows the phonetic detail, pitch and duration of the stimulus and Kerry's imitation of the phrase. Kerry's voice was creaky from section 5, which coincides with multiple interharmonics that are visible on the spectrogram (Figure 5.13: Tier 3, sections 5 and 6). Despite pitch breaks the melodic contour shows a similar trend to the target phrase. Although the pitch contour of *Clare* is shallower on Kerry's production, there is a downward change in pitch between this and the second word, *ate*, just before the pitch break (section 5). Kerry produced most of the word *her* [əʰə] at a higher pitch than *ate*, but lower pitch than the first word *Clare*, emulating the pattern of the stimulus. In contrast to the stimulus phrase, Kerry emphasised the word *lunch*: the increase in intensity and the initial increase in pitch indicate this (Figure 5.13: section 6). The imitated phrase is similar to the target in pitch range (Table 5.15). The first two words are similar when comparing the relative duration (Figure 5.13: Tier 2) of the words 'Clare ate' as produced in the demonstration (section 1, Tier 3) to that of Kerry (sections 4 and 5, Tier 3). Kerry's final two words were longer in relative duration than in the target phrase (Tier 2) and Kerry's phrase was longer overall (Table 5.15).

5.3.6 Song imitation: Magic Book

Kerry required a prompt to recall the words of the song in all renditions, but she recalled more words independently in later renditions (Figure 5.14). Her production was fluent but there was some repetition in her third production (line 3), and of *abracadabra* in her first and final productions.

Kerry reproduced elements of the melodic shape accurately on her first attempt (Figure 5.15). Her second phrase *magic book* (K1) was close to the target interval and absolute pitch. However, she spoke the second phrase, in imitation (also see Figure 5.14). She recalled more words of the third phrase in her second rendition (K2) and the melodic contour was close to the target (Figure 5.15: in blue). The accuracy of her intervals was the least accurate in her third production, in which she tapped the beat whilst singing. In her final, solo, performance (K4), she independently reproduced all words except for *what's* (Figure 5.14) but she did not produce the appropriate melodic interval leading up to the word *the* in the final phrase. In all renditions, Kerry's starting note remained close to the target and she reproduced at least one interval accurately in each performance (Figure 5.15).

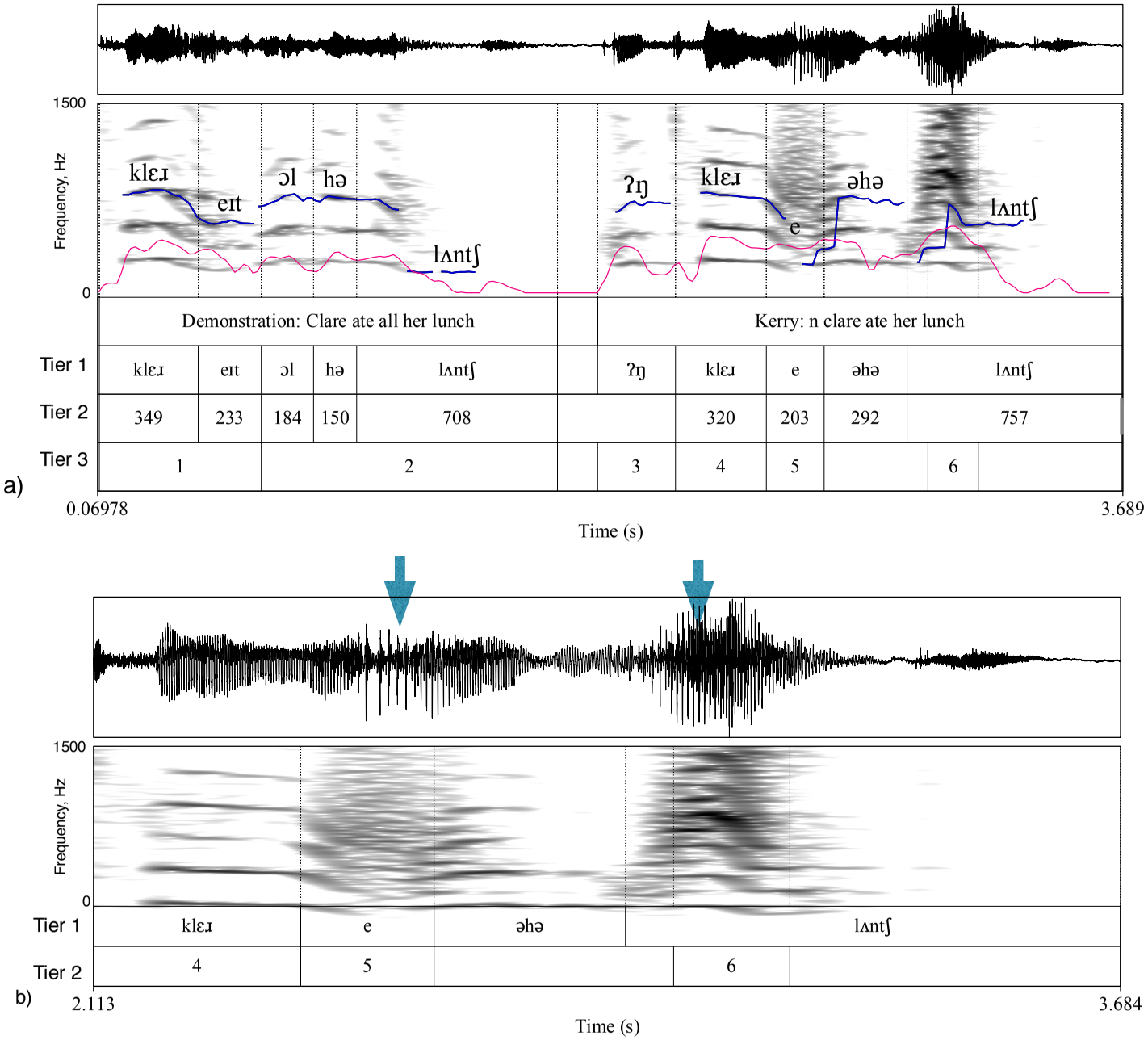


Figure 5.13: Waveforms and annotated spectrograms showing Kerry's imitation of 'Clare ate all her lunch'.

a) demonstration and imitation: the image shows the waveform, narrow-band spectrogram (0-1500 Hz; window length 150 ms; 35 dB), phonetic production (Tier 1), and syllable duration (Tier 2: milliseconds). The phonetic production, pitch contour (blue) and intonation contour (pink) are shown on the spectrogram.

b) Kerry's imitation: detail. The image shows Kerry's imitation only (settings as above). Her phonetic production is marked in Tier 1 and points of interest in Tier 2: these are discussed in the text. Kerry's voice quality contains multiple subharmonics and these are visible on the waveform of b) (marked with an arrow).

Target	Kerry 1	Kerry 3 (with taps)	Kerry 4 (solo)
1. **ma**gic **boo**k	1. {V̈ **bæʒɤ bœ**	1. **æ̃dʒɪk bʊk**	1. **bæ**ʒɪk b̲ʊ̲k
2. **ma**gic **Book**	2. **bæʒɪk bʊ** V̈}	2. **maʒəx bʊχ**	2. **bæ**ʒɪχ b̲əʊ̲k
3. **what**'s In**side** the	3. (*what's inside*) ʊ̃̈s ʔɪs̰daḭ	3. wʌ dɪ̰ns ə **ɪn**sait	3. (*what*) wʌ dɪ̰sdaɪ̰ ðɪ
4. **ma**gic **Book**	4. (t*he magic book*) **bɤ̃x**	4. (*the magic book)*	4. **bæ**ʒɪk bʊk
5. **ab**raca**da**bra	5. ɐpw**ʌts**…**dɛf̃**ɒ̃ħ	5. ʌbwæ ʊ **æ̃b**wə	5. sss.. **bæ**bwʌ

Figure 5.14: Phonetic transcription and syllabic stress of Kerry's first, third and final productions of 'Magic Book'. The figure shows the phonetic detail of Kerry's productions over a four week period. Prompts for words and starting pitches that were given by the researcher are shown in italics. Stressed syllables are highlighted in bold text. ExtIPA symbols and VoQs are used to indicate atypical phonation.

Figure 5.15: Graph showing the development in Kerry's recall of the words and melody of 'Magic Book', over a five-week period. K1 was her first imitation, in week 1, and her final, solo, production was K4 (week 5). In week 4 (K3), Kerry tapped the pulse as she sang. The target contour is shown in blue.



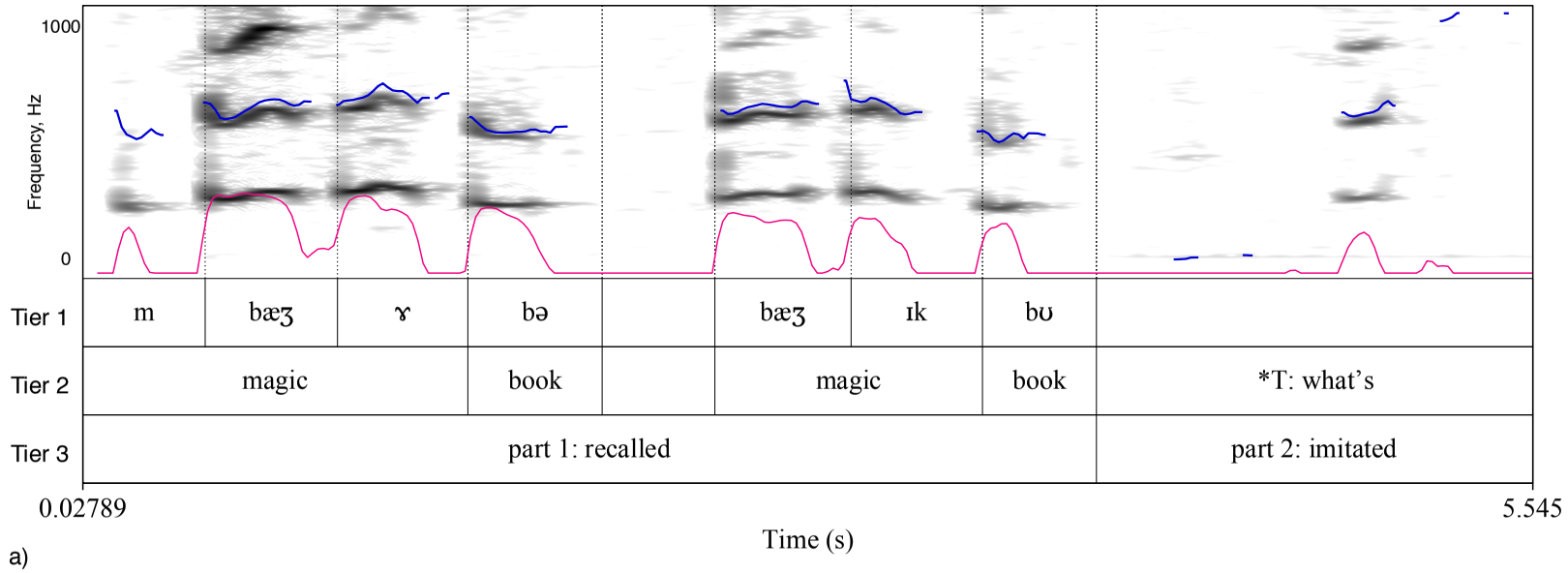
On her first attempt at singing the song (K1), Kerry produced the first two phrases of the song from memory (Figure 5.16a: part 1) and the second two phrases after prompts (Figure 5.16b: part 2). There is a difference in voice quality between the two conditions and the spectrogram reflects this. Noise on the spectrogram of the recalled phrases (Figure 5.16a), indicates a breathy voice quality; in the prompted phrases, interharmonics are visible in all Kerry's utterances. There are two pitch breaks (marked with a red asterisks), and her perceived phonation was harsh. This third phrase of the song was produced with atypical phonation on subsequent renditions (see Figure 5.14).

In her final rendition of the first phrase of the song (Figure 5.14: Kerry 4), Kerry produced more consonants and articulated the sounds more rapidly than in her first imitation (Table 5.15). However, she reproduced more of the melodic intervals to within 0.5 ST in her first rendition (Table 5:15: see bold type).

Figure 5.16: Annotated spectrograms showing Kerry's first production of 'Magic Book'.

a) Part 1: the first phrase that Kerry produced from memory, and the first prompt for phrase 2.
b) Part 2: the second phrase, that Kerry produced after prompts.

Both images show the waveform, narrow-band spectrogram (0-1000 Hz; window length 150 ms; 45 dB), phonetic production, pitch contour (blue: 40 - 450 Hz) and intensity contour (pink: 25-100 dB). Kerry's phonetic production is shown in Tier 2 and the target words are shown in Tier 3. Prompts are indicated in Tier 2, marked with a black asterisk. The pitch breaks in the imitated phrases (marked with red asterisks) coincide with perceptually atypical phonation.



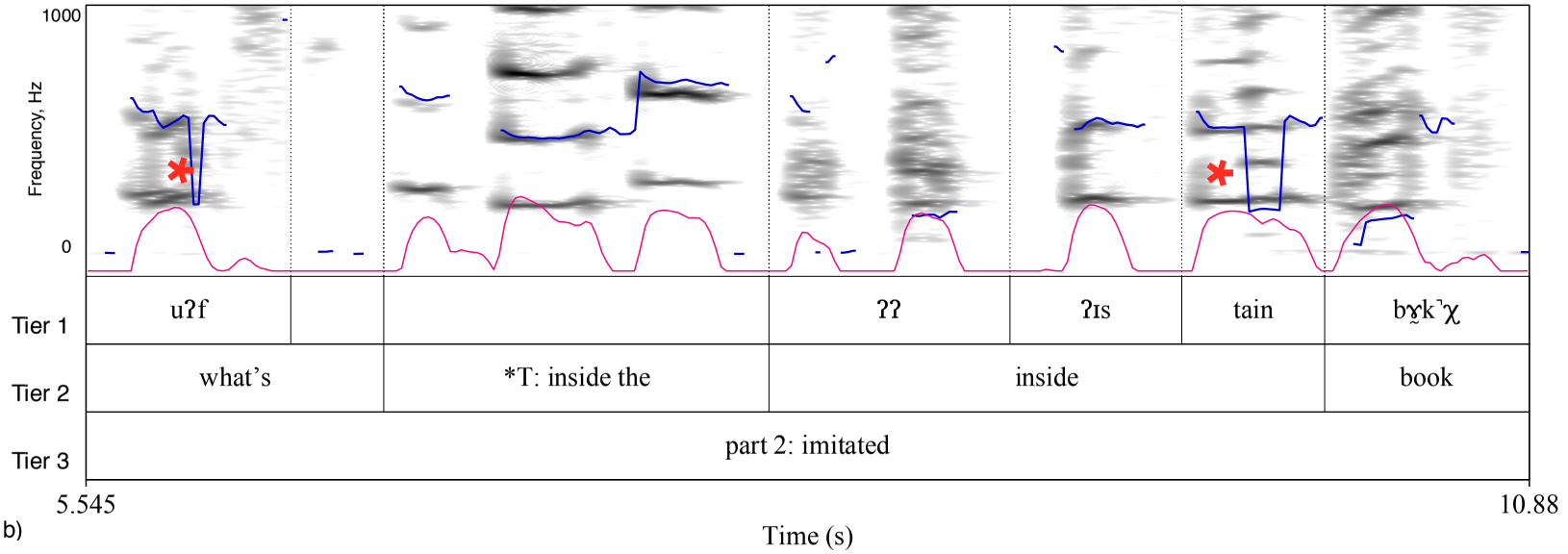


Table 5.15: Articulatory and musical features of the demonstration of the first phrase of 'Magic Book', and of Kerry's first and final productions. The table shows segmental, articulatory, melodic and rhythmic features of the demonstration and Kerry's initial and final productions. It shows the number of consonant sounds (C) Kerry produced in the phrase ('magic book, magic book'), and their relevance to the target sound in terms of manner and place. Articulatory features measure the rate of production of consonants and vowels and the relative duration of these. The mean change in semitones between syllables is given (melodic features). Direction is indicated (downwards (-), upwards (+)): those in bold text highlight pitch reproduction that was accurate within 0.5 ST. The table also shows the mean duration and relative duration of syllables in the words (rhythmic features).

Demonstration: key features			**First production**	**Final production**
Segmental production	**Number of C sounds**	**10**	6	10
	Number of relevant sounds	**10**	6	10
Articulatory features	**Total duration C, ms**	**856**	1660	1279 (77%) of speed
	Articulatory rate: sounds per second	**11.8**	3.6	7.8
	Total duration V, ms	**1577**	1426	1203 (84%)
	Average duration of C:V	**100:184**	100:112	100:106
Melodic features	***Magic*: Mean ST change**	**0**	+1	-1.3
	***Book*: Mean ST change**	**-3**	**-3.5***	-1.3*
	***Magic*: Mean ST change**	**0**	**0.2***	**+0.1***
	***Book*: Mean ST change**	**-3**	**-3.2***	**-3.2***
Rhythmic features	**Mean duration of *magic*, s**	**398**	504	336
	Mean duration of *book*, ms	**797**	944	706
	Relative duration of syllables *ma:gic*	**1:1**	**1:1**	**1:1**
	Relative duration of the words *magic:book*	**1:2**	**1:2**	**1:2**

5.4 Prosody and voice in production of speech and song from LTM

Kerry's spontaneous speech was mostly monosyllabic and her longest connected phrases were produced when singing and speaking *Happy Birthday.* This section will focus on the prosody and voice quality of a selection of these.

5.4.1 Spontaneous speech

Example: Really **good**: [ɪi̯˞ dɤəthˀ]

Figure 5.17 shows Kerry's comment on the singing performance of one of her peers. The comment is the longest spontaneous utterance she produced in teaching sessions, and is the only available sample of connected speech that was not imitated, or derived from song (see Section 5.4.2). The two syllables in *really* are defined rhythmically by changes in vowel [ɪ]-[i] (section1) and there is no gap in articulation between the first and second word. A gradual rise in pitch and intensity, coupled with a lengthier duration, placed emphasis on the onset of the word *good* (section 2). Kerry's pitch

spanned over nine semitones and dropped during the second part of the vowel of *good* (section 3). Her utterance was mostly modal in quality, but changes in the first formant frequency contour in sections 1 and 2 coincided with an atypical vocal quality: the waveform also reflects a change at these points. The vowel in section 3 was produced with interharmonics as the pitch contour descended, and with a perceptually harsh quality.

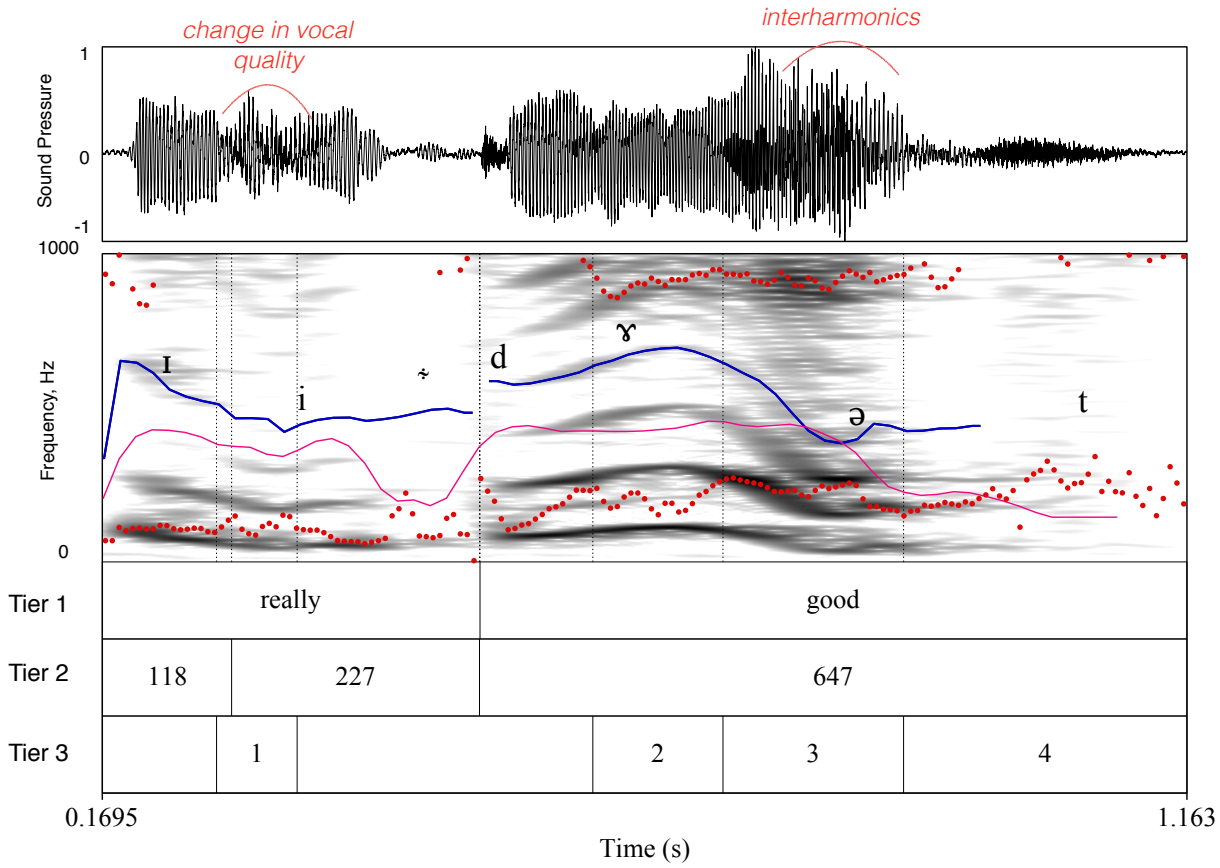
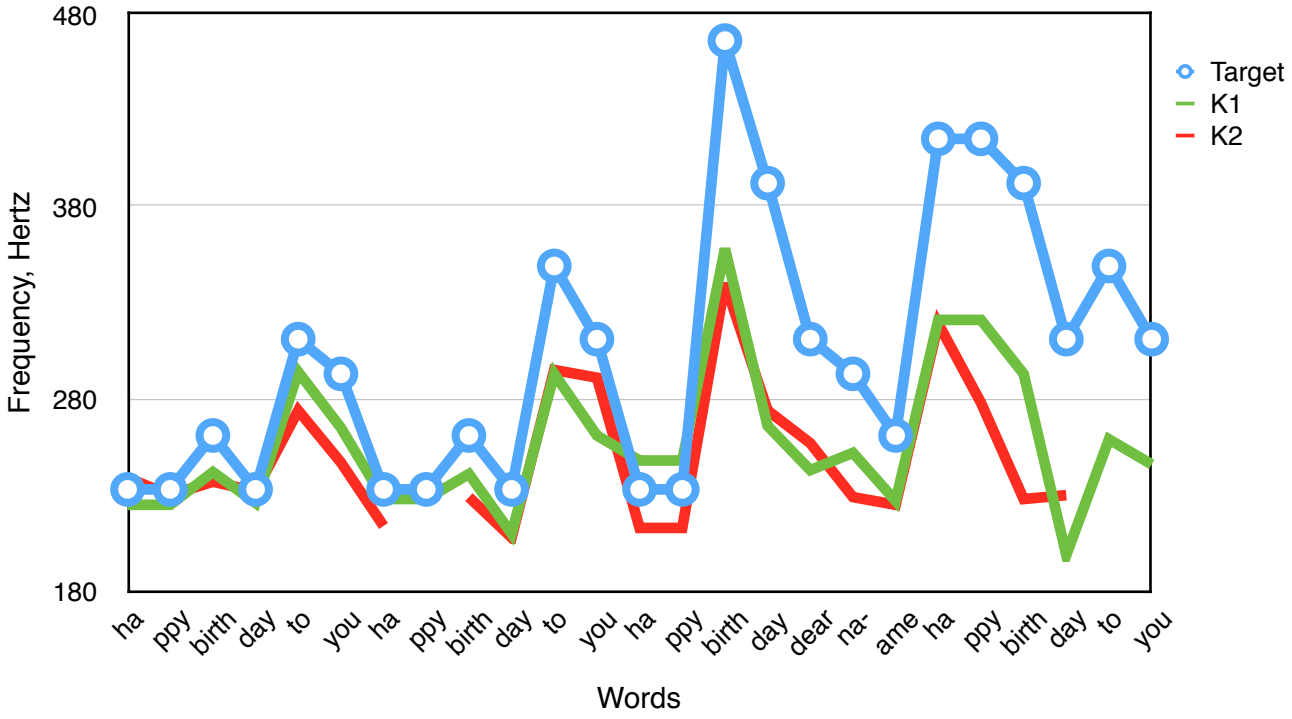


Figure 5.17: Waveform and annotated spectrogram showing Kerry's comment to a peer: 'really good'. The image shows the spectrogram (0-1500 Hz; window length 80 ms; 40 dB), pitch contour (in blue: 50-500 Hz), intensity contour (in pink: 35-100 dB) of Kerry's comment to a peer. The first formant frequency and part of the second formant frequency are also shown on the spectrogram (red dots). The duration (milliseconds) of each syllable is shown in Tier 2 and points of interest in Tier 3. These are discussed in the text.

5.4.2 Happy Birthday in sung and spoken conditions

Kerry sang the song on two occasions and she reproduced the melodic shape of the song each time (Figure 5.18). Her production of the words when speaking and when singing was fluent (Figure 5.19). Although Kerry's vocal production was melodic, her voice quality was intermittently creaky or harsh in quality when singing and speaking. In both conditions, a similar proportion of vowels in words were produced with interharmonics (Figure 5.19: expressed as a percentage of the vowel duration). Her sung production was similar to the target in terms of perceived stress but her spoken production placed emphasis on different syllables (Figure 5.19).

Figure 5.18: Graph showing the melodic shape of Kerry's productions of 'Happy Birthday' (sung). The image shows a plot of the mean fundamental frequency of each syllable Kerry produced on two separate occasions (K1 and K2). The target melodic contour is shown in blue, and is based upon Kerry's starting F0. The changes in mean F0 are plotted against time (see Methods, Chapter 3).



Target	Sung (version 2)	Spoken
1: happy **birth**day **to** you	1:hæbi **bə**deɟi **tu̯** ju̯ (41%)	2: hæ**bɪ** bɜθdi tɪ **ju̯** (17%)
2: happy **birth**day **to** you	2: hæpɪ **bəf**dei **tu̯** ju̯ (70%)	
3: happy **birth**day dear…	3: hæbi̯ **bə**:deɟi tiə…**tu̯** ju	3: **hæ**bə **bɜð**i (***kɐ̯****rry*) (30%)
4: happy **birth**day **to** you	(55%)	4: h..hæ**bɪ̯ bɜ̯f**deɟ tu **ju̯** (77%)
	4: **æ** bədi **tu** ju̯ (58%)	Interharmonics: 41.33%
	Interharmonics: 56%	

Figure 5.19: Phonetic transcription and syllabic stress of Kerry's sung and spoken productions of 'Happy Birthday'. The phonetic detail includes an indication of the perceived syllable stress and of voice characteristics. The atypical voicing as indicated by the presence of interharmonics is also shown, expressed as a percentage of the total vowel duration for each phrase, and as a total of each production. ExtIPA symbols and VoQs indicate voice quality.

Kerry's spoken phrases are shown in Figures 5.20, 5.21 and 5.22. Each phrase was produced with a different rhythm: the first phrase was produced with equal syllable timing and equal intensity in the first three syllables (Figure 5.20: Tier 2); the second was more variable in syllable duration and in intensity (Figure 5.21); and the third (phrase 4) was produced with pauses between each word. Phrase 4 followed a 2.25-second period of dysfluency and was produced with a greater proportion of subharmonics (Figure 5.22).

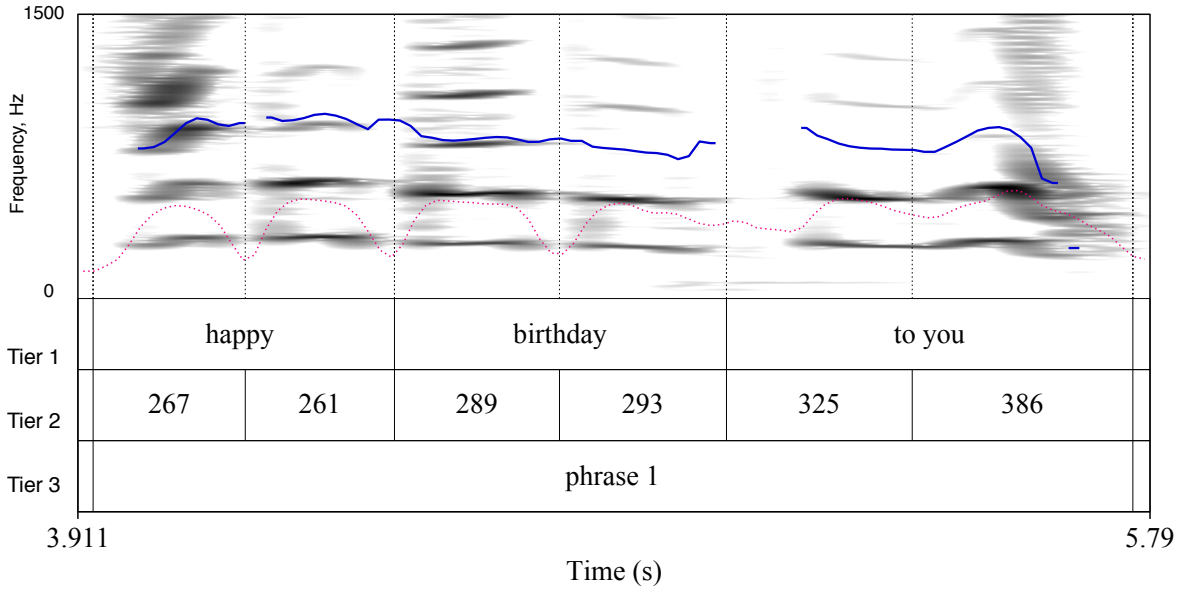


Figure 5.20: Annotated spectrogram showing Kerry's production of 'Happy Birthday', spoken. The image shows the waveform and spectrogram (0-1500 Hz; window length 80 ms; 40 dB), which is annotated with the pitch contour (in blue: 50-450 Hz), and the intensity contour (in pink: 35-100 dB). Tier 1 shows the words, Tier 2 shows the duration of syllables (in milliseconds) and Tier 3 shows the phrase number. Kerry produced the first phrase after a verbal prompt.

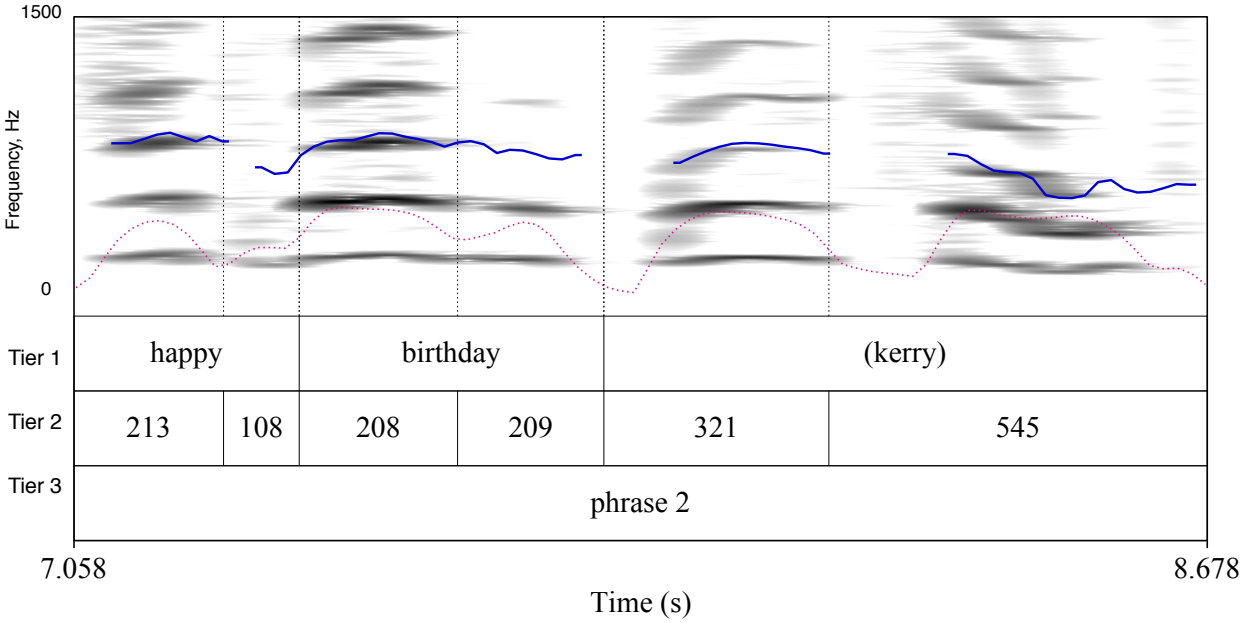


Figure 5.21: Annotated spectrogram showing Phrase 2 of Kerry's production of 'Happy Birthday', spoken. The image shows the waveform and spectrogram (0-1500 Hz; window length 80 ms. 40 dB), which is annotated with the pitch contour (in blue: 50-450 Hz), and the intensity contour (in pink: 35-100 dB). Tier 1 shows the words, Tier 2 shows the duration of syllables (in milliseconds) and Tier 3 shows the phrase number. Kerry supplied her own name at the end of the phrase and began to produce the word *happy* immediately after (not shown): this was followed with a 2.25 s period of dysfluency.

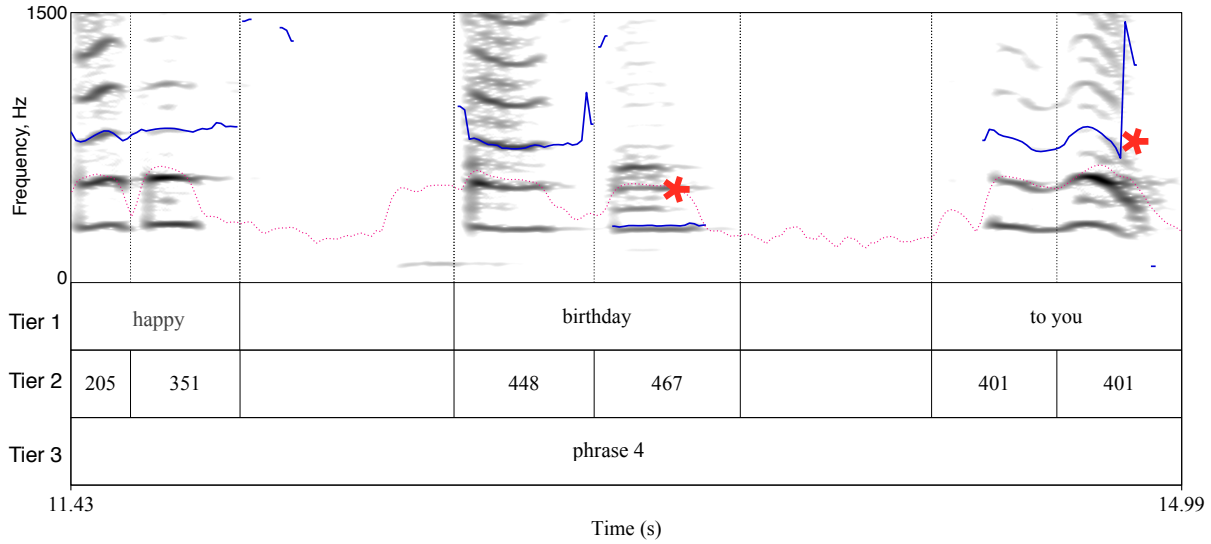


Figure 5.22: Annotated spectrogram showing Phrase 4 of Kerry's production of 'Happy Birthday', spoken. The image shows the waveform and spectrogram (0-1500 Hz; window length 80 ms. 40 dB), which is annotated with the pitch contour (in blue: 50-450 Hz), and the intensity contour (in pink: 35-100 dB). Tier 1 shows the words and further information, Tier 2 shows the duration of syllables (in milliseconds) and Tier 3 shows the phrase number. The final phrase was produced slowly and with pauses between words. The spectrogram shows subharmonics between the main harmonics, and breaks in the pitch contour in two words: these are indicated by a red asterisk.

The melodic and rhythmic data for the phrases that had the same target rhythm when sung (phrases 1, 2 and 4) are shown in Table 5.16. The data show a difference between Kerry's sung and spoken productions. Both sung phrases match the rhythm (RD) of the song but the rhythm of her spoken phrases are different to the song, and to each other (Table 5.16, Figure 5.20). The data show that Kerry reproduced the pitch changes of the song when singing and that she matched five intervals to within a semitone (Table 5.16: in bold). However, her melodic contour in the spoken phrases was different to the song and most melodic changes were close to one semitone.

Table 5.16: the relative pitch and relative duration of syllables in 'Happy Birthday' in Kerry's sung and spoken conditions, taken from the same recording session. The table shows the mean pitch change in semitones between syllables (ST), and the relative duration of syllables (RD) in comparison to the shortest syllable duration in the phrase. The (+) indicates a rising melodic interval, the (-) indicates a falling melodic interval: those in bold text highlight pitch reproduction that was accurate to within 0.5 ST. The mean speaking rate for each phrase is shown.

Target			**Singing, vs.2**				**Spoken**			
Phrase 1			**Phrase 1**		**Phrase 2**		**Phrase 1**		**Phrase 4**	
Syllable	**Pitch change, ST**	**RD**	**Pitch change, ST**	**RD 1=265ms**	**Pitch change, ST**	**RD 1=200ms**	**Pitch change, ST**	**RD 1=300ms**	**Pitch change, ST**	**RD, 1=200 ms**
ha	**n/a**	**1**		**1**		**1**	n/a	1	n/a	1
py	**0**	**1**	**0.0**	**1**	**0.0**	**1**	+1.1	1	+0.5	2
birth	**+2**	**2**	0.0	**2**	+1.0	**2**	-1.1	1	-1.7	2
day	**-2**	**2**	-0.5	**2**	**-1.8**	**2**	-1.4	1	0.0	2
to	**+5/+7**	**2**	+3.0	**2**	**+6.0**	**2**	+ 0.7	1	0.0	2
you	**-2**	**4**	**-1.8**	**4**	-1.2	**4**	—0.5	1	+1.0	2
Mean speaking rate, syllables per second			**2**		**1.9**		**3.3**		**1.7**	
Pitch range (max-min) ST			**6.76**				**3.24**			

5.4.3 Reading

Figure 5.23 shows Kerry's initial phrase in the reading test. Her speech sounds were mostly accurate in terms of place and manner (see also Table 5.3), although many sounds were aspirated or nasalised in production ([sʰwi̯m], [wə̃n]) and some sounds are replaced with glottal stops ([æʊʔsait]). Kerry's speech was produced slowly, at 0.87 syllables per second, and was affected by frequent pauses and by some hesitation, but without repetition of speech sounds.

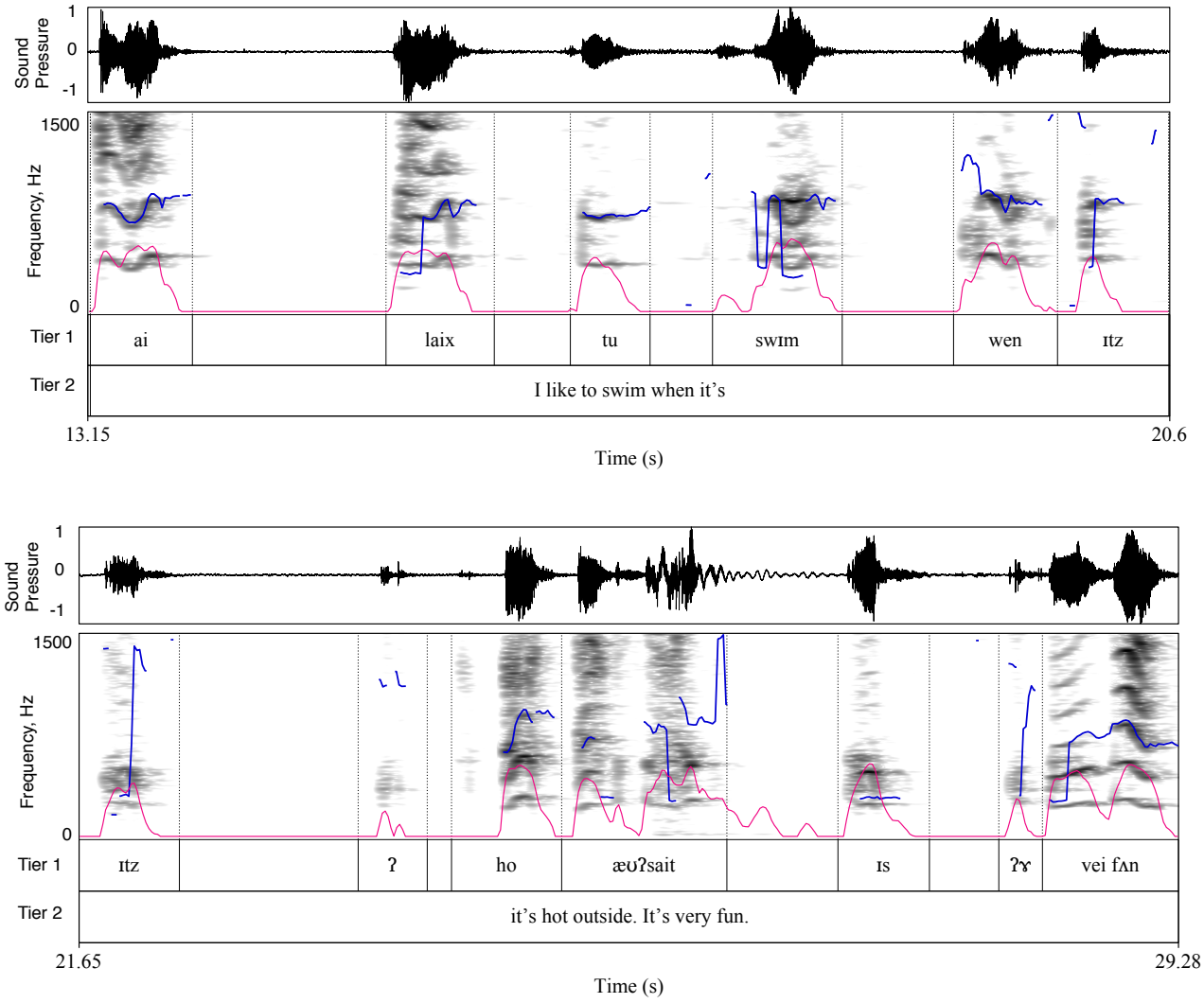
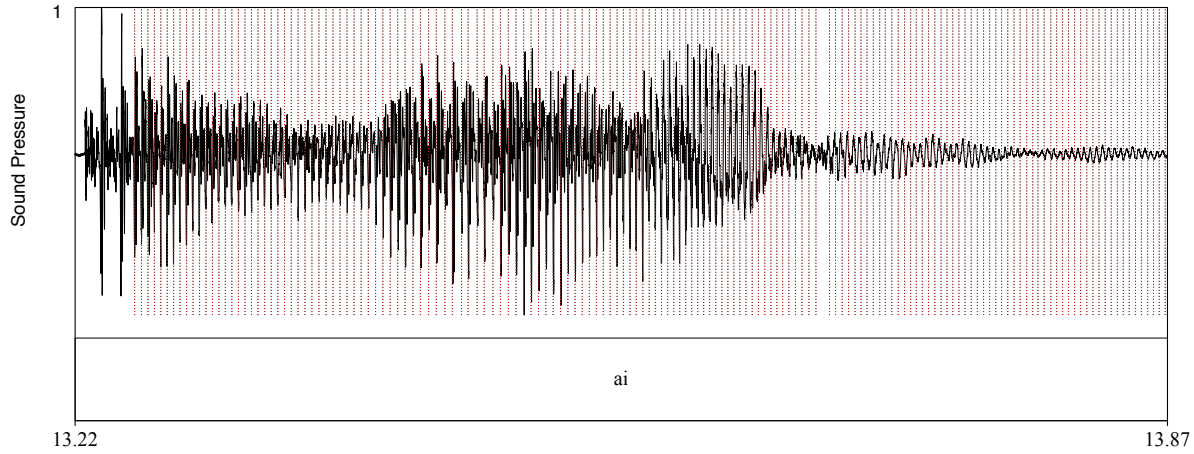


Figure 5.23: Annotated spectrogram showing Kerry's production of Phrase 1 of the reading task. The image shows the waveform and narrow band spectrogram (0-1500 Hz; window length 150 ms; 45 dB). The pitch contour (blue) and intensity contours (in pink: 35-100 dB) are shown: there are frequent breaks in the pitch contour that coincide with breaks in phonation. Tier 1 shows Kerry's production and Tier 2 shows the text.

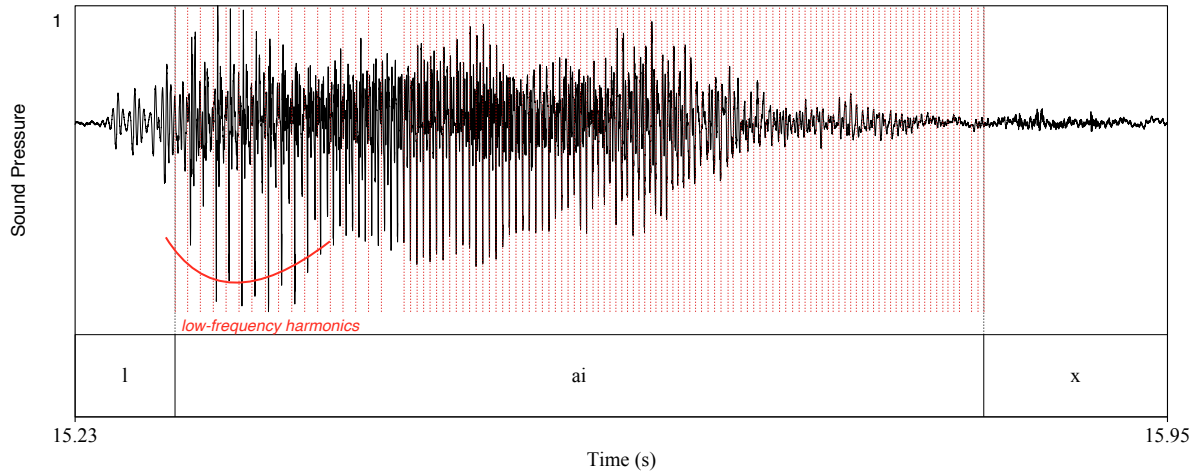
The words were produced within a pitch range of seven semitones. Pitch breaks affect the contour on the image (Figure 5.23) and as a result do not represent fully the perceived intonation: six of the words (e.g. *I*, *like*, *to*, *swim*, *hot*, *outside*) were produced with a rise-fall intonation. However, the words at the end of each sentence were grouped together; at the end of the first sentence, the

intonation moves upwards on the final word *outside*; and the pitch declines at the end of the second sentence, on the word *fun*.

Figure 5.24 shows waveforms of the words *I* and *like*, which are both high-frequency words in written English (frequency per million words: 6494 and 983, respectively: Leech, Rayson and Wilson, 2001). The word *I* was perceived as harsh in quality and the word *like* was perceived as creaky. Both words were produced with harsh initiation at onset. There are additional harmonic peaks between the pulse lines at the onset of the vowel in *like* (Figure 5.24b): these coincide with a 'creaky' quality and the pitch break (Figure 5.23).



a) 'I'



b) 'like'

Figure 5.24: Waveforms of the words 'I' and 'like', as produced by Kerry in the reading task. The pulse lines are shown on the waveform (in red). At the onset of the diphthong in 'like' there are high-intensity, low-frequency harmonics visible on the waveform that fall between pulse lines (indicated with a red contour).

Discussion

The discussion will begin with an examination of Kerry's rhythmic production in motor tasks, and in speech and song tasks (Section 5.5). This will be followed with a discussion of her melodic abilities and vocal qualities in speech and song. Section 5.6 will discuss the similarities and differences in her rhythmic and melodies skills in different domains, and will then consider the combined effects on Kerry's prosodic production in both domains of task demand and segmental context. The section will also consider the similarities and differences between her productions of *Happy Birthday* in spoken and sung conditions. Section 5.7 will draw together the findings and consider to what extent Kerry's speech and voice production may be supported or developed by musical activities.

5.5 Rhythm, voice and melody in speech and song

5.5.1 Temporal perception and production in music

In beat entrainment tasks, Kerry was able to synchronise within the first bar, which is within adult norms (Drake et al., 2010). Likewise, her performance in matching the beat at 132 bpm (Table 5.6) was within norms for adult non-musicians, who tap 20-50 milliseconds before the beat (Krause, Pollok and Schnitzler, 2010; Reifinger 2006; Repp, 2005). Her SD of ITI was variable (norms for drummers are about 2.5%: Repp and Su, 2013) but her mean AT of 5% at 132 bpm is close to norms for adult non-musicians (Reifinger, 2006). Her mean AT for one bar at 132 bpm was within 10 milliseconds (2.2% of ITI) of the target, which is less than the reported norms (Repp, 2005; Repp and Su, 2013). Given this, it is likely that Kerry's poorer ability to synchronise at the faster and slower tempi (Table 5.6) are either motoric or the result of poor sustained attention. What is worth noting is that her good timing skills in beat entrainment tasks are in contrast to recent research that showed that as a group, children who stutter tend to vary more in the duration of timing between taps (Olander, Smith and Zelaznik, 2010). Falk et al. (2014, 2015) reported that adolescents who stutter also tap earlier than controls (Falk, Müller and Dalla Bella, 2014, 2015). It is argued that a motor-timing deficit is common to the speech and to the motor domains of people who stutter (PWS) and of other speech-impaired groups (Corriveau and Goswami, 2009; Falk et al., 2015; Olander et al., 2010; Peter and Stoel-Gammon, 2009). Aside from potential methodological differences (see Chapter 3), Kerry's relatively superior music-motor abilities may result from her fondness for music and her training in drumming. Training in drumming can lead to enhanced motoric timing for beat keeping in relation to those without training (Krause et al., 2010).

Kerry did not mark the 'strong' beat of a bar by using extra motoric effort (Figure 5.1), a skill that develops in TD children by 7-8 years, in line with developments in motor control and attention (Drake et al., 2000; Drake and Gérard, 1989; Reifinger, 2006). In people with DS, limb motor

movements may be under-developed (Latash et al., 2007, 2008; Palisano et al., 2001; Rigoldi et al., 2011; Vicari, 2006), and Kerry's DAP score of 7;3 years (Table 5.1) indicates that she may not yet have developed the necessary motor control to mark meter. However, the ability to reproduce metrical accent is also dependent upon the ability to perceive and attend to changes in intensity (Gérard and Drake, 1990). Olakunbi, Bamiou, Stewart and Luxon (2010) reported a significant impairment in children and adolescents with diagnoses of APD to discriminate meter (Olakunbi et al., 2010). Differences in auditory processing have been identified in individuals with DS (Groen et al., 2008; Marcell et al., 1990; Marcell, 1995; Pettinato, 2009; Wuang and Su, 2011). Kerry's difficulties in discriminating consonant sounds (Table 5.1: test 13a) may reflect an auditory processing difficulty, or a hearing difficulty. Furthermore, HI can exacerbate auditory processing deficits in people with DS (Marcell, 1995). Any difficulties in hearing or auditory processing will disrupt the auditory-motor feedback system that is required for musical motor control (Sowínsky and Dalla Bella, 2013; Tierney and Kraus, 2013). Further testing would be necessary to establish the role of hearing or perception on Kerry's motoric ability.

Kerry was able to segment word rhythms during her second assessment (Figure 5.3), and to entrain to simple rhythms (Figure 5.2) but not to syncopated rhythms. The ability to entrain to a rhythm is correlated with motor development (Reifinger, 2006; Drake et al., 2010: Gérard and Drake, 1990) and TD children do not consistently produce asymmetrical rhythms until about seven years (Reifinger, 2006). For Kerry, as with metrical production, difficulties in these tasks could reflect a developmental delay. Similarly, data from rhythm imitation tasks suggest that she may have difficulties in controlling precisely timed rhythmic movements. In the TRaCoL task (Table 5.8), Kerry reproduced the number of claps in seven of the eleven rhythms that consisted four beats or less, with some minor alterations to the duration (Table 5.8), but she made errors in reproducing three of the rhythms that consisted of three claps. This error in reproducing rhythms that were within her memory span indicates either a failure to retain the relative duration, a failure to translate this into motor movements, or difficulties in producing precisely timed motor movements. Kerry's inability to reproduce rhythms of over four beats is consistent with her TAPS score of 3 (Table 5.1), suggesting that auditory working memory affected her performance at this level. The data are consistent with research in people with DS that has identified motoric deficits (e.g. Rigoldi et al., 2011; Latash et al., 2007, 2008; Vicari, 2006) and deficits in auditory memory (Baddeley and Jarrold, 2007; Carney et al., 2013; Costanzo et al., 2013; Lanfranchi et al., 2010) and attention (Lanfranchi et al., 2010; Will et al., 2014). However, as for metrical perception, her performance could also be affected by hearing or processing deficits.

5.5.2 Temporal perception and production in speech and song

The key factors that appear to affect Kerry's rhythmic production of speech are her dysfluency, and her inability to recall segmental or lexical detail (e.g, Table 5.12, Table 5.14). In naming and word-imitation tasks, dysfluency affected the onset of many single words (Table 5.3), which slowed her mean speaking rate. Dysfluency affected fewer nonwords than words (Table 5.3), but she was fluent when reading (Table 5.3), when singing (Figure 5.14, Figure 5.19), and when reproducing a short comment (Figure 5.16). Dysfluency is prevalent in people with DS (e.g. Kent and Vorperian, 2013), but the data indicate that different task demands affect her fluency, and therefore her rhythmic production: the potential impact of different processing styles on her fluency and voice production will be discussed further in this section. Additionally, the duration of some dysfluent segments in imitation tasks resulted in pauses of over one second between the demonstration and Kerry's production (e.g. the words *loving* and *zipper:* Figure 5.8), which may affect Kerry's retention of the word in phonological working memory.

Despite dysfluency and inaccurate reproduction of speech sounds, when Kerry reproduced the correct number of syllables in speech imitation tasks and songs, she tended to also reproduce the relative duration of syllables (e.g. Figure 5.8: Tier 1, Figure 5.12: Tier 2, Table 5.15: rhythmic features, Table 5.16: RD). Kerry also accurately reproduced the rhythm of some phrases even when lexical detail was poor (e.g. Table 5.14: phrases 3 and 7, Figure 5.13). For words and phrases, Kerry seems to have a secure concept of rhythmic structure at the macro-level, but comparison between Kerry's imitation and the demonstrations indicate that her accuracy is affected by segmental differences. For example, Kerry replaced the initial voiced alveolar fricative in the word *zipper* (Figure 5.8) with a voiceless glottal fricative: this change reduced the duration of Kerry's production, relative to the demonstration. In songs, too, her production was affected by segmental production: the rate of her production of words in *Magic Book* increased with practice but remained slower than the demonstration (Table 5.15). Additionally, Kerry spent proportionally more time producing consonants than vowels: for Kerry, it is very likely that articulatory demands affect her rhythmic production, and that in singing at least, practice can increase the rate of her production.

Whilst the data indicate that Kerry's rhythmic production in speech and song is primarily affected by production difficulties and memory for the content, there is evidence that her perception or her hearing may affect her production. Kerry did not reproduce the correct number of syllables in the four nonwords that began with a weakly-stressed syllable (Table 5.12: items 6-9) but she did reproduce four syllables of the strong-initial nonword, *sepretennia*l. As previously discussed (Chapter 4), such a deficit has been reported in DS subjects (Pettinato and Verhoeven, 2008). When learning *Magic Book*, Kerry also had difficulty in learning the words to the third phrase

(*what's inside the*: Figure 5.14, Figure 5.16). The syllable *in* and the word *the* both fell on unstressed beats in the song, which is comparable to the syncopated beats in a rhythm. As strong beats are perceived more easily than weak beats (Krumhansl, 2000), Kerry may have a particular difficulty in perceiving or isolating these sounds in music and in speech. However, Kerry did reproduce the word after practice, and it may be that she needs frequent exposure and practice in order to perceive and produce some sounds or some weakly-stressed sounds or syllables. In her final rendition of *Magic Book* (Figure 5.14), she sang the words with changes in intensity and relative pitch that conveyed a clear sense of metrical production. This, together with her production in *Happy Birthday* (Figure 5.19) and her ability to reproduce most stressed syllables in speech tasks (e.g. Table 5.12, Table 5.14), suggests that she does have an ability to perceive the subtle temporal cues that mark stress or meter in speech. However, the combined data indicate that in common with the DS population (e.g. Kent and Vorperian, 2013), she may have physiological difficulties in co-ordinating oro-motor movements, and in producing efficient phonation (e.g. Figure 5.4, Figures 5.9-5.11) which will hamper metrical production and rhythmic stress in speech.

5.5.3 Pitch perception and production in song and speech

Although Kerry was unable to complete the pitch perception tasks for speech or music (Table 5.1: tests 14 and 21), her production skills indicate that she can reproduce and therefore perceive changes in frequency to within half a semitone (e.g. Figure 5.7: *so-la-so*, Table 5.15, Table 5.16). This is equivalent to about a 6% change in fundamental frequency. However, as adults with normal hearing acuity can detect changes as small as 0.5% of the frequency (Kruhmansl, 2000), further testing would be needed in order to determine whether Kerry's perception of pitch is typical.

Despite being able to match single pitches within the lower range of her voice (Figure 5.6) and intervals within songs (Figure 5.15, Figure 5.18), Kerry was unable to match intervals (Figure 5.7). In the first assessment, she seemed to use the second or final pitch as a starting note for her own and in the second, she inverted a three-note melodic pattern (*so-la-so).* This may indicate a sequential memory difficulty, which has been reported in people with DS (Laws, 1998), and in non-musicians when asked to imitate melodic patterns of three notes and over (Williamson et al., 2010). The second assessment shows a poorer performance in terms of melodic shape, but the intervals that Kerry reproduced were more accurate than in the first assessment (Figure 5.7). Kerry's pitching was also closer to target pitches, and coincided on two occasions (*me-do* and the *me-do* interval of *so-me-do*). Potentially, Kerry developed her skills in matching absolute pitch during the teaching programme, but attentional, behavioural or vocal differences may have affected her abilities during either assessment. However, Kerry's behaviour in these tasks is typical of children when developing singing skills (Welch, 2009): in line with developing singers, she sings

most accurately within her spoken range (Rutowski, 1997: 'limited range singer') and is unstable in tonality (Welch, 1998: Phase 3 of 4).

Kerry reproduced the pitch contours of approximately 90% of words (Table 5.10, Table 5.12) and nearly two-thirds of phrase contours (Table 5.14). However, her pitch range was smaller than demonstrations. This is comparable to other studies that report reduced intonational range relative to controls for DS subjects in speech production tasks (Lee et al., 2009; Stojanovik, 2011; Zampini et al., 2015) and imitation tasks (Stojanovik, 2011). However, Kerry's accuracy in reproducing individual words and phrases was affected by atypical phonation (e.g. Figure 5.10, Figure 5.13) and by poor recall of the words in phrases and song. This notably affected her ability to repeat pitch contours in *Magic Book* (Figure 5.16). In comparison, her pitch range in *really good* (Figure 5.17) was nine semitones. The data suggest that Kerry's perception of pitch in song and speech allows for near-typical production, but that physiological and cognitive deficits impact upon her ability to produce pitch changes.

5.5.4 Voice quality, phonation and cognitive load

In common with descriptions of voice quality in DS (see Kent and Vorperian, 2013), Kerry's voice in speech is characteristically 'gruff', 'rough' or 'creaky'. Such labels may relate to different types of phonation (Aronson and Bless, 2009; Cavalli and Hirson, 1999; Keating et al., 2015): examination of the spectrograms for words and vowels has shown that these perceived qualities are linked to harsh production and intermittent diplophonia; and to the presence of interharmonics. The data from sustained vowels (Table 5.9) indicate that Kerry has difficulty in initiating voice, which in DS subjects may result from difficulty in applying the optimum degree of muscular tension to activate the vocal cords (Pryce, 1994). However, four sustained vowels were produced with unvoiced frames mid-vowel (Table 5.5) which demonstrates difficulty in sustaining phonation, once initiated. Although this difficulty can indicate an underlying organic problem (Aronson and Bless, 2009; Mathieson, 2001), in certain circumstances Kerry is able to produce stable pitches and controlled glides with minimal perturbations and with perceptually 'modal' phonation (e.g. vowels /u/1 and /a/ 3, Table 5.9). This suggests that Kerry has the physiological and neurological ability to produce efficient phonation. Given this, there must be one or more sources of interference with phonation. Data from Kerry's production of words in speech and song indicate that physiological and cognitive factors may affect the incidence of perturbation and perceived vocal quality.

Difficulties in initiating voice will affect the rapid onset and offset of phonation required for speech (Fourcin and Abberton, 2008). This effect is clear in Kerry's data. In imitated words (Table 5.11), Kerry's voice production changed in line with the number of syllables produced, whilst target articulatory movements remained similar. Perceptually, as the target number of syllables increased,

Kerry's vocal quality in the 'love' word-set changed from breathy (Figure 5.9) to creaky (Figure 5.10) to harsh (Figure 5.11). This could reflect a different vocal setting for longer words, when imitating. The breathy quality in the one-syllable word indicates a lax setting. Lax muscles allow too much airflow, reducing efficiency of phonation (Braunschweig et al., 2008; Laver, 1980; Mathieson, 2001; Titze, 1994), which would increase the 'noise' in the signal, leading to reduced HNR values. Both longer words were produced with increasing harshness which indicates an increasing degree of tension within the larynx (Aronson and Bless, 2009; Mathieson, 2001). The spectrograms also show an increase in the proportion of interharmonics and in the intensity of these. Excessive tension is likely to involve the false vocal folds (Bailly, Henrich and Pelorsonl, 2010), and can lead to instability (Beck, 2010; Tao and Jiang, 1998) and diplophonic production (Cavalli and Hirson, 1999), as was perceived in the three-syllable word (Figure 5.11b). For the task as a whole, a reduction in breathiness may account for the observed increase in HNR values in two- and three-syllable words, whilst other measures of perturbation increased (Table 5.11).

An increase in 'creakiness' and a reduction in breathiness is associated with increased cognitive load (Quatieri, Williamson, Smalt, Patel, Perricone, Mehta et al., 2015; Yap, 2012, 2015) and the data from the Syllables test (Table 5.11) show this effect. Silverman (2007) argues that processing load interferes with language ability in people with DS: as speech is not 'automatic' and requires effort, the process effectively drains resources for higher-order activities. Furthermore, for Kerry, the effects of cognitive load on her voice production may be exacerbated as a result of her dysfluency. PWS are more susceptible to increased demands on phonological memory and are less able to shift attention than people who do not stutter (Anderson and Wagovich, 2011; Bajaj, 2007), with consequences for achieving automaticity. PWS are also more susceptible to increased muscular tension within the larynx (Sebastian et al., 2013). Kerry's phonation was unstable even in contexts that placed minimal demands upon her memory or her articulation (e.g. sustained vowels: Table 5.9). This, too may stem from her dysfluency. Severe stuttering in children is associated with reduced F0, increased jitter and shimmer, and reduced control of intensity (Salíhovic, Junuzović-žunić, Ibrahimagić and Begonović, 2009). The authors suggest that the vocal instabilities are associated with weak laryngo-muscular control and reduced control of sub-glottal pressure (Salíhovíc et al., 2009).

Although Kerry's voice appeared vulnerable to increased syllable length in words, the degree of perturbation did not increase in line with target syllable length in nonwords (Table 5.12). However, in comparison to her imitation of words (Table 5.10), Kerry reproduced nonwords with fewer syllables (Table 5.12) and with fewer appropriate speech sounds (Table 5.3), but with comparable mean pitch contour and pitch range (compare Table 5.10 and Table 5.12). This is surprising, given that the CNREP task increases cognitive processing demands in other populations with speech

and language deficits (Gathercole and Baddeley, 2006). The difference in Kerry's voice production between tasks may indicate a change in attentional focus that could effectively reduce cognitive load. As the novelty of phonemes in CNREP would place increased demands upon Kerry's phonological working memory and upon her motor programming system (Anderson and Wagowich, 2011; Stackhouse et al., 2007), Kerry may have focussed less on segmental production and more on the prosodic features. In the Syllables task, the effect of cognitive load associated with planning for speech that was accurate in both segmental and prosodic aspects might be to increase Kerry's stress response. This can affect muscle tension at the vocal fold level (Aronson and Bless, 2009; Pinho et al.,1999; Roy, 2003; Salatoff, 2014). In comparison to imitation tasks, retrieval tasks from LTM would reduce the overall processing demands on speech and voice planning. The divide in voice quality that was noted between recalled and imitated sections of *Magic Book* (Figure 5.16) suggests that her production of even recently-learned phrases is superior to that of single words that requires more effort to recall.

When reading, Kerry's production was fluent, her segmental production was the most accurate of the tasks (Table 5.3) and her production was intelligible (Figure 5.23). However, the task resulted in a higher degree of perturbation measures (Table 5.5, Figure 5.23, Figure 5.24) than CNREP (Table 5.5), which was recorded within the same session and immediately preceded the reading task. Children with DS are slow to learn to read orthographically and still rely on the use of STM to decode written words (Ratz, 2013). The written form seems to have helped Kerry produce the speech sounds but at the expense of her ability to produce efficient phonation. Again, this effect on phonation may stem from the effort required in transcoding graphemes into speech sounds.

Kerry's phonatory characteristics are consistent with a diagnosis of dysphonia, which is believed to affect individuals with DS (Kent and Vorperian, 2013). This indicates a difficulty within the central nervous system that affects both speech muscles and laryngeal muscles and may also be central to Kerry's dysfluency. Ludlow and Loucks (2003) proposed that the same centralised motor control system may affect speech processing in PWS as those with dysphonia. Given that stuttering is known to increase with processing demand (Yap, 2012, 2015; Ludlow and Loucks, 2003), a centralised difficulty in control would explain the observed change in phonatory difficulties under different task demands. Likewise, stuttering and dysphonia are both observed to improve in conditions such as laughing, crying and singing (Ludlow and Loucks, 2003). A centralised motor difficulty would also explain Kerry's improved performance in phonation and reduction in stuttering in singing conditions; the relative exception being singing imitation, which was poorer in vocal quality measures, and which suggests that Kerry's phonatory stability is particularly sensitive to specific test conditions that require high degrees of attention and processing. This is consistent with the research that shows that 'non-meaningful' speech and the non-verbal communication of

emotion are produced with fewer stuttering or dysphonic symptoms and that symptoms rise as the processing demands increase (Ludlow and Loucks, 2003). In the reading task, Kerry's speech production was fluent but her phonation was high in perturbation: potentially, different aspects of the processing may affect how the central system controls speech muscles and vocal muscles, and that the vocal muscles are sensitive to the effects of cognitive load. However, data from CNREP indicate that the effects may be mediated by reduced attention to the accuracy of segmental production.

5.6 Links between musicality and speech

5.6.1 Difficulties common to speech and music domains

Kerry's accuracy in reproducing rhythms and beats appears to be affected by tempo, which suggest that her motoric skills are delayed or otherwise limited. Although her accuracy in synchronising to a beat does not indicate the degree of perceptual deficit found in other language-impaired populations (e.g. Corriveau and Goswami, 2009; Cumming et al., 2015; Huss, Verney, Fosker, Mead and Goswami, 2011; Peter and Stoel-Gammon, 2008), this does not rule out perceptual difficulties in addition to motoric limitations. Kerry does have difficulties in discriminating between consonant pairs (Table 5.1: test 13a) and this may reflect reduced processing or reduced hearing acuity. Kerry is able to produce speech sounds when reading, so she may have greater difficulty in perception of speech sounds, rather than an inability. Any perceptual difficulties will also be affected by noisy environments, which can delay processing time for speech sounds (Kraus, Strait and Parbery-Clark, 2012). If Kerry has a perceptual deficit, this may be more evident in speech imitation tasks than rhythmic motor tasks, due to the additional demands on phonological WM and the complexity of speech in comparison to that of a single rhythmic tap. Further research would be needed to clarify this, and to clarify to what extent her speech perception is affected by hearing difficulties (Table 5.1).

Perceptual deficits are implicated in Kerry's production of weak syllables and in some speech sounds in both speech and song tasks, but it is likely that her prosody is affected significantly by production difficulties. Prosody requires the control of phonation, duration, and relative intensity: a change to one parameter alters the overall effect. For Kerry, the articulatory, physiological and cognitive factors that affect her speech and phonation impact directly upon the intonation and rhythm of her speech. Well-rehearsed tasks or gestalt phrases seem to reduce the impact of some of these factors, allowing her to produce phrases that are more fluent, melodic and controlled in terms of rhythm and intensity (e.g. Figure 5.17). However, her production is affected in speech and in song by demands on her STM. Phonological demands in particular seem to impact upon her vocal and prosodic abilities. For example, her memory for prosody appeared to be preserved in the

absence of semantic or segmental detail (e.g. Table 5.14: phrase 3); whereas her voice quality was adversely affected in the reading task (Table 5.5: test 15), in which her segmental production was most accurate (Table 5.3: test 15). The degree of her dysfluency or the degree of omitted sounds/ words also increased with cognitive load (Table 5.3), which is in line with research into dysfluency (Bosshardt, 2006). The data suggest that cognitive load puts pressures on different aspects of her motor planning and/or programming system according to the task, and that her approach to the task may ameliorate or exacerbate this. For Kerry, her 'best' speech output may occur when speech output requires a lower level of conscious control. Examination of her production of spoken and sung *Happy Birthday* allows examination of her speech abilities in the absence of auditory memory demand and when phonological planning has been previously practised; it also allows examination of any differences in rhythm or melody. Together, these details will allow consideration of the potential for singing to support Kerry's speech.

5.6.2 Happy Birthday: evidence of transfer potential

Kerry's productions of *Happy Birthday* were mostly fluent (Figure 5.19) and produced with few voice perturbations. The incidence of interharmonics was low (Figure 5.19: 41.33%, 56%), relative to other speech tasks (Table 5.5). The data indicate that her ability to sing with fluency and with relatively improved phonation transferred to speech. However, there are differences between each version that indicate that despite Kerry's superior production in comparison to other tasks, the task demands associated with speech may still affect her production.

Kerry's intonation and rhythm in her spoken phrases were different from those of her sung version (Table 5.16). The first phrase of her spoken production was highly metrical and syllable-timed (Figure 5.20), resulting in a different time signature than her sung version (4/4 rather than 3/4). It has been argued that rhythmic production may focus attention (Hawkins, 2014), facilitate vocal output (Davidow et al., 2012; Thaut, 2008; Wan et al., 2010), and facilitate motor production in patients with aphasia (Schlaug et al., 2008; Stahl et al., 2013). Evidence also shows that an audible beat helps recall of text (Racette and Peretz, 2007), and reduces dysfluency (Alm, 2004; Wieland, McAuley, Dilley and Chang, 2015).

Kerry's second and third spoken phrases (phrase 3 and phrase 4, respectively) were qualitatively different to the first phrase in terms of rhythm, fluency and phonation (Table 5.16, Figure 5.20, Figure 5.21). Phrase 3 was naturalistic in rhythm and melody, but phrase 4 was initially dysfluent with lengthy pauses between each word, and contained a higher proportion of atypical phonation (Figure 5.19, Figure 5.22). In TD adults, recall of song lyrics are serial, based on 'lines' or grouped phrases, and memory for words is strongest for the initial lines (Racette and Peretz, 2007). The increase in pause duration and frequency in the latter phrases may indicate increased stress

(Buchanan et al., 2014) associated with an effort to recall the later text. Despite these difficulties in the fourth phrase, the task as a whole was produced with a lower incidence of dysfluency and interharmonics in comparison to other speech tasks (compare Figure 5.19 with Table 5.3 and Table 5.5). Singing words enables the formation of internal timing cues for speech which will reduce stuttering (Alm, 2004), with consequences for phonation (Loudlow and Loucks, 2003; Salihovíc et al., 2009). Although evidence suggests that PWS fail to generate an internal rhythmic template for their speech (Wieland et al., 2015), for Kerry, prior learning of the text in song may have provided such a template.

5.7 Summary and implications

Kerry's skills in rhythmic timing and in reproducing melody, aspects of which are within typical levels for adults (Levitin and Cook, 1996; Repp, 2005), indicate a sensitivity to temporal and pitch changes that is in advance of her verbal MA (Table 5.1). In terms of her singing development, Kerry exhibits characteristics of a limited range singer (stage 3 of 5: Rutowski, 1997) and is at Phase 3 of 4, in terms of singing development (Welch, 1998). Her sense of tonality in song places her at the equivalent of about 7 years (Gooding and Standley, 2011; Hargreaves, 1996); her inability to mark meter in rhythmic tasks places her skills within an equivalent age range of 5-9 years (Gooding and Standley, 2011; Hargreaves, 1996). The data from *Magic Book* demonstrate that Kerry has a capacity to learn and develop musicality, beyond what might be expected from her verbal MA and her abilities in speaking tasks. Potentially, perceptual deficits may affect her temporal production in rhythmic motor tasks and speech and song. However, Kerry's accuracy and sensitivity in temporal timing when performing rhythms, the fluency of her speech in singing tasks, and her ability to sing with metrical emphasis suggest that a perceptual deficit is not likely to be the sole cause of her speech production difficulties. The data suggest a core immaturity either in muscle movement or in motor programming, both of which are observed in people with DS (Almeida et al., 2000; Capio and Rotor, 2010; Latash et al., 2008; Rigoldi et al., 2011; Sacks and Buckley, 2003).

However, the difference between Kerry's segmental accuracy, prosody and phonation in different tests implicate that task effects also influence production, and may exacerbate any motoric difficulties. For example, in any one task if the processing demand is high, Kerry's motor-timing system may be sensitive to disruption: this may result in dysfluency, in pitch breaks, or both. An increase in attentional demands leads to difficulties in maintaining a steady /a/ vowel for TD adults (Tumber, Scheerer and Jones, 2014): the authors suggest that high attentional demands leave fewer resources available to monitor auditory feedback. They suggest that this potentially affects both vocal control and speech control. It is therefore possible that different modes of stimuli place stress on different parts of the speech system. For example, the visual stimulus of text in reading for speech sounds may assist oro-motor planning and programming but the effort involved in

decoding the letters for sounds impacts upon the phonatory process; whereas singing and speech tasks that place fewer demands on auditory working memory appear to place less pressure on the systems required for speech planning and production. Whilst there may well be motoric limits to the accuracy of Kerry's speech and voice, rehearsal seems to allow for her optimum production of speech.

The effects of improved speech timing and phonation do seem to have transferred to Kerry's spoken production of *Happy Birthday,* especially in the over-learned and temporally-structured first phrase. Kerry's production of the recently learned *Magic Book* from memory was also produced with fewer disturbances to rhythm or voice relative to her imitated section. This may show the effects of transfer from STM to LTM in progress: once words, rhythm and melody are stored in LTM, phonation ought to improve. This has clear implications for practice. Even children with deficits in phonological WM are able to form stable mental representations with practice (Gathercole, 2006). Thus, practice leads to increased automaticity of speech production, which will reduce cognitive load. Automaticity in speech may be specifically difficult to achieve for people with DS (Silverman, 2007) and for PWS (Anderson and Wagowich, 2011). However, if the act of learning words in the context of a song reduces planning demands and enables Kerry to more easily produce speech that is fluent and less dysphonic, then automaticity of words or phrases may be more easily achieved.

Singing may confer particular benefits for Kerry in terms of fluency. Studies indicate that people with DS may rely on their right hemisphere for language processing, unlike TD populations, but utilise the left (Elliott and Weeks, 1993; Chua et al., 1996; and see Chapter 2, Section 2.5.2). However, their processing of music is typical: melody is processed in the right hemisphere, and rhythm in the left hemisphere (Chen et al., 2015). In speech, greater use of the left hemisphere in DS is associated with an increase in speech errors (Bunn et al., 2007). In comparison to speaking, the act of singing activates functions in the right hemisphere that may lead to increased control of motor timing for oral articulation and laryngeal control (Jeffries, Fritz and Braun, 2003). Similarly, MIT is believed to assist recovery of speech in patients with left-hemisphere aphasia by utilising the right hemisphere for motor production (Schlaug et al., 2008). FMRI scans (Jeffries et al., 2003) have shown that the control of speech movements when singing text, rather than speaking the same words, occurs in the right hemisphere and may specifically assist timing for word production and greater control of the larynx and phonation. Singing words rather than speaking the same words also places greater demands on the cerebellar circuitry (Jeffries et al., 2003) which has been linked to timing control in speech movements (Peter and Stoel-Gammon, 2008). Therefore, learning to produce words whilst singing may assist speech production by laying down appropriate motor memory and by reducing cognitive load, whilst enhancing aspects of precision in motor

timing within the cerebellum. This would set up favourable conditions for fluency and phonation, especially if dominant control for speech were within the right hemisphere; and for potential transfer between domains (Patel, 2011, 2014).

The data indicate that Kerry would benefit from musical activities that promote efficient phonation and fluent speech. As for Andrew (Chapter 4), Kerry's voice production could be supported through exercises that promote healthy vocal habits, including relaxed onset of phonation, controlled exhalation and coordination of the breath with phonation (e.g. humming, wind instruments, kazoos). Singing activities could be used to support the learning of words and phrases. For example, the use of music and song to rehearse words or phrases may lead to a more accurate motor representation, storage and coding of timing, and reduce demand during recall. The inclusion of an auditory beat or rhythm may also support Kerry's production by reducing speech-timing demands. Of course, it is not possible to prepare for spontaneous speech in this way: typically, the cognitive load associated with linguistic planning can result in dysfluency (Bosshardt, 2006), and, for Kerry, in dysphonia. However, if activities included words and phrases that were commonly used by Kerry there may be transfer benefits to speech.

Chapter 6: Rachel

This chapter will present and discuss the results of Rachel's performance on a range of core speech and music tasks. Section 6.1 will present a summary of her speech, musical and behavioural abilities, based on her college records and her performance in core tasks (see Chapter 3, section 3.2). Results from motor-music tasks will be given in Section 6.2. Section 6.3 will present the voice data from sustained vowels, and her accuracy in reproducing pitch and prosody in pitch and interval-imitation tasks, and speech- and song-imitation tasks. Section 6.4 will focus on the prosodic features of her speech and song in production tasks that draw from long-term memory.

These results will be discussed in Sections 6.5-6.7. The discussion in Section 6.5 will examine the similarities and differences between Rachel's temporal production in motor-music, speech and music tasks. This section will also examine similarities and differences in her ability to reproduce or produce melody in speech and song. Section 6.6 will discuss the potential links between Rachel's rhythmic, melodic and vocal abilities in speech and song. Section 6.7 will discuss the key difficulties that support or hinder her learning and her ability to transfer learning from the music to the speech domain.

6.1 Summary of history, learning and performance in core assessments

Rachel is aged 23;10 years, and is described in her college records as having SLD, and mild but untested hearing loss. Her records state that she enjoys listening to music, singing and dancing and that she attended dance lessons as a child. She continued to participate in drama, music and singing lessons for three years at college. According to her SaLT report, Rachel's speech is 'clear for single words' and she has no oral motor difficulties. However, she is reported to have some dysfluency in connected speech, which is limited to single nouns and short phrases and which is of reduced intelligibility. Her SaLT commented that Rachel is a sociable and friendly young lady but that she becomes self-conscious if she thinks she is at fault: the SaLT report recommends patience in order to encourage Rachel to participate in tasks that she thinks might be too difficult. The college records state that Rachel understands Makaton, but does not habitually use it for communication.

6.1.1 Cognitive and perceptual abilities

Rachel's test results suggest a verbal MA equivalent of 3 years, but a nonverbal MA of 7;6 years (Table 6.1). Her TAPS score indicates an auditory digit span of three. Rachel had difficulty in discriminating between vowels (20% accurate) and some difficulty in discriminating between consonants (80% accurate) (Table 6.1: test 13). In the pre-test for word discrimination, she was

Table 6.1: Profile: a summary of Rachel's performance in core tasks. The table summarises the core assessments in nonverbal and verbal ability, speech perception and production and musical perception and production. Test numbers refer to the test descriptions in Chapter 2, Section 3.3.

	Test number	Test name	Results
Cognition	**1**	BPVS RAW score	25
	1	BPVS age equivalent	3;0 (2;8 -3;5)
	2	Draw-a-person	7;6 years
	3	TAPS - digit span	3
Perception (speech)	**7a**	Same-different discrimination non-words (Stackhouse and Wells App C.2)	feature: 3/5=60% sequence: 3/5=60%
	7b	Same-different discrimination real-words (Stackhouse and Wells App C.2)	feature: 4/5=80% sequence: 1/5=20%
	13a	Minimal pairs perception: discrimination of consonants	4/5 =80%
	13b	Minimal pairs perception: discrimination of vowels	2/5 = 20%
	14	Simple speech prosody test	NULL
Speech Production	**5**	DEAP: Percentage of Phonemes Correct (PPC)	69.5
	8	s/z	n/a
	9	DDK	n/a
Perception (music)	**17**	PMMA	n/a
	26	Pitch perception test: pitch matching	1/3 (33%)
	27	Rhythm perception test (same/different)	NULL
	27	Pitch and interval perception test (same/different)	NULL
	28	Interval and rhythm perception test: find the sound	NULL
Music Production	**17**	Gross motor score	35/40 = 88%
	4	Rhythm imitation (TraCOL)	7/14 = 50%
	22a	Pitch matching on 'la'	1/5= 20%
	22b	Pitch matching: sustained /i/, /u/, and /a/ vowels at three pitches	4/10 = 40%
	23a	Pitch direction: vocal glides on 'la'	2/4 = 50%
	23b	Interval matching	1/5 = 20%
Reading	**15**	Not attempted	n/a

able to state whether pairs of familiar words (her name, the researcher's name, *cat*, *dog*, *pig*) were the same or different. However, she had difficulty in distinguishing between real words based on the sequence of sounds (e.g. *rates* vs. *raced*) than on a difference in feature (e.g. *race* vs. *rate*) (Table 6.1: test 7b). Rachel did not complete the simple prosody test (14). She responded *don't know* to three questions, and to the remaining five questions she selected the picture that related to what she wanted, stating *Rachel wants…* (e.g. *chips*: see Appendix 2, p. 330). Rachel was not able to complete the training tasks for the simple pitch, interval or rhythm tests (26-28) and therefore did not attempt the PMMA.

6.1.2 Motor-music abilities

When observed against a checklist (Appendix 2, p.324), Rachel was able to imitate and create gross motor movements in time to the music and to tap in time to the pulse. When asked to walk in time to music, her timing was slow and lagged slightly behind the beat, but all other motor movements were produced at the tempo that was appropriate for the salient beat, and were closely aligned to it. In TRaCoL, Rachel scored 7/14 (Table 6.1). She correctly reproduced patterns of three claps and two patterns of four claps, and she partially reproduced patterns of four and five claps. Further details of her performance will be given in Section 6.2.

6.1.3 Oro-motor and phonological abilities

Rachel was unable to complete the DDK task or the s/z task. In the DEAP task she correctly produced 69.5% phonemes (Table 6.1). Most errors were gliding voicing errors or errors of omission (Table 6.2). For example, Rachel replaced most alveolar approximants, /ɹ/, and lateral approximants, /l/ with the glide /w/. One exception is her pronunciation of /ɹ/ in /zɛbɹə/ as [dʒɛbɹə]. Unusual errors included substitution and omissions. For example, Rachel substituted the initial voiceless alveolar stop /t/ of 'tiger' with a voiced alveolar lateral approximant [læ:gə]. She produced 'tomato' as [tenteu], omitting the 'strong' syllable. In the word 'gloves', she substituted the alveolar voiced fricative /v/ in /glʌvz/ with a voiced bilabial stop, [gwʌbs]. In the Syllables task (Table 6.3), errors in phoneme, consonant, vowel and omissions increased as the target word length increased to three syllables. The highest errors were made on the CNREP task and the phrase imitation task, which contained a high percentage of omitted sounds.

Table 6.2: A summary of Rachel's speech errors and error patterns in the DEAP task. The table shows the number of errors made in the task and the error patterns (n= 5 or more; except WSD* n=3), expressed as a percentage of the total number of errors made. The errors are coded according to the conventions outlined in the DEAP manual, which defines the following categories: D indicates delayed error patterns; U indicates unusual error patterns.

DEAP		
Percentage of error pattern/total errors		
Gliding	**D**	28.0%
Cluster reduction	**D**	15.3%
Voicing	**D**	9.6%
Weak syllable deletion*	**D**	5.7%
Other	**U**	7.6% examples : /m/ =[n]; /f/= [w]; t= [l]; m =[b]
Number of errors		**52**
Percentage of errors accounted for by error patterns		**66.2%**

Table 6.3: The percentage of speech errors produced by Rachel in core speech tasks. The table shows the percentage of errors in phonemes (PPE), consonants with habitual errors (glides) removed (PCE-R), vowel errors (PVE) and the percentage of sounds omitted (PEO). The number of words examined in each test and subtest (Syllables) is given.

Test no.	Test name	No. of words n=	PPE	PCE-R	PVE	PEO
5	**DEAP**	**42**	33.65	38.24	18.60	6.33
11	**Syllables**	**27**	50.71	62.79	31.48	18.57
	1-syllable words	**9**	33.33	33.33	33.33	10.00
	2-syllable words	**9**	57.78	62.96	50.00	18.60
	3-syllable words	**9**	69.23	78.95	55.56	26.15
6	**CNREP**	**9**	69.55	61.57	51.08	35.33
12	**Phrase imitation, (Newton, 1999)**	**5**	64.89	68.04	64.90	60.77

6.1.4 Voice range and measures of voice quality

Rachel's voice range when imitating pitch glides on /a/ was 20.2 semitones (two octaves) and her speaking range in conversational tasks was over 12 semitones (one octave) (Table 6.4).

Table 6.4: The mean pitch and pitch range of Rachel's vocal glides on /la/ and connected speech. For each task, the table shows the mean, minimum and maximum values of fundamental frequency (Hz), and the SD of F0 when speaking, in Hz and semitones (ST). The mean range in speaking pitch is also shown in ST. Mean F0 and SD of F0 is not given for glides, whose purpose was to examine vocal range.

Test no.	Test name	Mean F0, Hz	Min F0, Hz	Max F0, Hz	SD of F0, Hz, ST	Pitch range, ST
22a	**Pitch glides on 'la'**		152.40	472.40		19.58
10	**Picture description**	215.12	110.31	295.32	37.69; 3.04	17.04
n/a	**Spontaneous speech**	245.28	130.45	356.57	40.28; 2.85	17.41

Rachel produced the medial portion of sustained vowels with few perturbations and with higher levels of perturbation values during onset and offset (Table 6.5). The mean MPT for sustained vowels was 1.191 seconds. The vowels that Rachel produced in each syllable of words in DEAP, Syllables and CNREP were higher in perturbation measures than for sustained vowels (Table 6.5). Interharmonics were present during the central portion of two sustained vowels (Table 6.5). All measurable sustained vowels were produced with interharmonics during the initial 100 ms (for methods, see Chapter 3, Figure 3.11 and 3.12).

6.1.5 Learning

During the baseline assessments Rachel sang melodically but with no clear articulation of words or consonant sounds and with inaccurate melodic intervals. She was set the following learning goals, in addition to the group goals (Appendix 3, p. 340):

- Match pitch in five notes within her natural speaking range;

Table 6.5: Mean voice perturbation measures for Rachel's sustained vowels and vowels in words. For sustained vowels, mean measures are given and separate values are given for the initial 100 ms (onset), medial duration (Total duration - (onset+offset)), and the final 100 ms (offset). Measures are also given for the vowels in each syllable of words produced in DEAP, Syllables and CNREP tasks. The table shows the mean measures of voice breaks, unvoiced frames, maximum phonation time, harmonics-to-noise ratio, jitter (%), shimmer (%). The table also shows the percentage of vowels in Sustained vowels, Syllables and CNREP tasks that were produced with interharmonics.

Test no:	Test name	No. of valid vowels, n =	DVB, %	DUV, %	HNR, dB	Jitter, %	Shimmer, %	CPP, dB	No. and percentage of vowels produced with interharmonics
19	**Sustained vowels: total duration (T)**	**8**	0.00	2.20	23.50	0.48	3.21	5.19	8/8= 100%
	Onset (t=100 ms)	*8*	0.00	*14.89*	*14.07*	*1.52*	*6.88*	*4.01*	*8/8 =100%*
	medial duration (T - [onset +offset]),ms	*8*	0.00	0.00	*25.07*	*0.35*	*2.62*	*5.29*	*8/8 = 100%*
	offset (t=100 ms)	*8*	0.00	*23.28*	*17.91*	*2.30*	*10.28*	*4.33*	*0/8 = 0%*
5	**DEAP**	**52**	0.97	7.53	12.11	1.39	12.64	4.36	
11	**Syllables**	**43**	2.11	4.11	13.59	4.81	1.48	4.81	37/43= 84%
6	**CNREP**	**14**	0.00	2.54	10.26	1.77	14.25	4.00	11/14 = 79%

- Develop accuracy in singing *so-me* and *so-me-la* intervals in new songs;
- Sing words from memory in familiar songs with clear initial consonant sounds; and
- Develop sustained phonation/ability to hum through imitation of 'hum' or use of kazoos.

Rachel was able to use a kazoo to sustain phonation for at least three counts at comfortable pitches (Appendix 3, p. 343[1]) and she developed her skills in matching single pitches in songs with a range of up to five notes (Appendix 3, p. 348 [26]) and *so-me* intervals (Appendix 3, p. 345 [6]). She was able to learn key words to new songs when prompted (Appendix 3, p. 349 [5]; p. 354 [14]), but she did not develop the ability to articulate key words clearly. Section 6.3 will examine aspects of her learning in more detail, based on the recordings of sung pitches, intervals and songs.

6.2 Motor-music production

6.2.1 Beat entrainment

Rachel's accuracy in beat entrainment changed with the tempo (Table 6.6). Her timing between successive taps (ITI, ms) was least accurate at 104 bpm in terms of percentage of the target (ITI, %) and was also the most variable (SD of ITI, ms). At 132 bpm, Rachel played at double the target rate, but measurements were taken for every second beat. The timing between every second beat was 101% of the target and synchronised to within 2.4% of the target beat. Rachel's ITI (ms) was within 1% of the target at 160 bpm but her AT (ms) was on average 218 ms (58%) out of sync with the target beat. Her timing to the beat (AT, ms; AT, %) was most accurate at 132 bpm but her standard deviation of 155 ms demonstrates greater inconsistency than at other tempi.

Table 6.6: The accuracy of Rachel's timing in beat entrainment and extraction tasks, at 104 bpm, 132 bpm, and 160 bpm. The table shows mean ITI and AT in milliseconds and as percentage of the target interval, and the SD in milliseconds. *Rachel played at twice the target tempo for the target beat of 132 bpm: the data reflect measurements between every second beat.

Tempo, bpm	**104**	**132**	**160**
Interval time, ms	**576**	**455**	**375**
No. of taps produced	**16**	**41**	**20**
ITI, ms	512	452*	372
ITI, %	89	101*	101
SD of ITI, ms	79	26*	34
Mean range of ITI, %	4.3		
AT, ms	-51	-11*	+218
AT, %	9	2.4*	58
SD of AT, ms	79	155*	35
Mean AT, %	23		

6.2.2 Rhythm entrainment

Rachel did not match either rhythm: she produced a repeating pattern of double beats, but defaulted to the pulse after two bars. Visual modelling of the rhythm did not affect Rachel's performance.

6.2.3 Syllable segmentation

Rachel produced one clap for the word *cat* without a demonstration but she was inconsistent in her clapping of *monkey*, and *caterpillar*, despite being able to count the dots representing the number of claps. She did speak and clap the syllables to *caterpillar* (Table 6.7) but she was inconsistent in

Table 6.7: The number and relative duration of claps produced by Rachel in syllable segmentation tasks. The table shows Rachel's accuracy in producing the appropriate number of taps to represent the number of syllables in words, and her ability to produce these with appropriate relative duration between taps. Rachel produced 'caterpillar' twice, and she said the word as she clapped: the number of syllables and claps she produced on each occasion is given.

Words	**Support**	**Target relative duration of claps**	**Number of claps/taps produced**	**Mean relative duration of claps**	**Number of claps and spoken syllables**
			motor production only		with speech
Can you drum the words words words words words words Can you drum the words words words words words	**none: independent**	**1:1 and 2:2**	18/32	1:1	
	with demonstration and ongoing clapped support		17/19	1:1 and 2:2	
	with demonstration and notated support		19/19	1:1 and 2:2	
Monkey	**visual dots**	**1:1**	3, 5		
Cat	**none**	**n/a**			
Caterpillar (1)	**visual dots**	**1:1:1:1**	4	2:1:1:2	4 claps; 4 syllables
Caterpillar (2)	**visual dots**	**1:1:1:1**	5	2:1.5:2:3	2 claps; 4 syllables

the number of claps she produced. When independently drumming the words to *Can You Drum The Words?*, Rachel produced an isochronous beat (1:1), rather than a rhythm (Table 6.7). After a demonstration, and with ongoing clapping of the rhythm, she produced the relative duration of the words (1:1 and 2:2). However, her playing lagged behind the demonstration by almost two beats. When given visual notation as support, Rachel played all beats with the correct relative duration and the rhythm of the second phrase was closely aligned to target at beats 11-13 (Figure 6.1).

6.2.4 Rhythm imitation

Rachel accurately reproduced rhythms consisting of three claps (Table 6.8). She reproduced two of the four rhythms consisting of four claps and partially reproduced rhythms of two claps and of four claps. In imitating longer rhythms, she reproduced the final four notes of a five-clap rhythm, but was unable to reproduce rhythms of six or seven claps.

Table 6.8: Rachel's accuracy when imitating clapped rhythms in the TRaCoL task (Treharne, 1999). The table shows the accuracy of Rachel's responses in terms of reproducing rhythms in number and duration (N&D), and partial replication of the number (N) or relative duration (D). The notes produced in error are indicated in red text.

Target no. of claps	**Target rhythms, with notation**	**Rachel: responses and number of correct responses**	
		N+D	**Partial**
2	**1.** ♩ ♩ **2.** ♫	**1.** ♩ ♩	**2.** ♩ ♫ ♩
3	**3.** ♩ ♫ **4** ♫ ♩ **5.** ♩ ♩ ♩ **6.** ♪ ♩ ♩	**3.** ♩ ♫ **4** ♫ ♩ **5.** ♩ ♩ ♩ **5.** ♪ ♩ ♩	
4	**7.** ♫ ♫ **8.** ♫ ♩ ♩ **9.** ♩ ♩ ♫; **10.** ♩ ♫ ♩ **11.** ♩ ♩ ♩ ♩	**8.** ♫ ♩ ♩ **10.** ♩ ♫ ♩	**7.** ♫ ♩ **9.** ♩ ♫ ♩ **11.** ♩ ♫
5	**12.** ♩ ♫ ♩ ♩		**12.** ♫ ♩ ♩
6	**13.** ♫ ♩ ♩ ♫		**13.** ♩ ♫ ♩
7	**14.** ♩ ♫ ♫ ♩ ♩		**14.** ♩ ♫
TOTAL	**14**	**7**	**7**
Percent correct		**50%**	**50%**

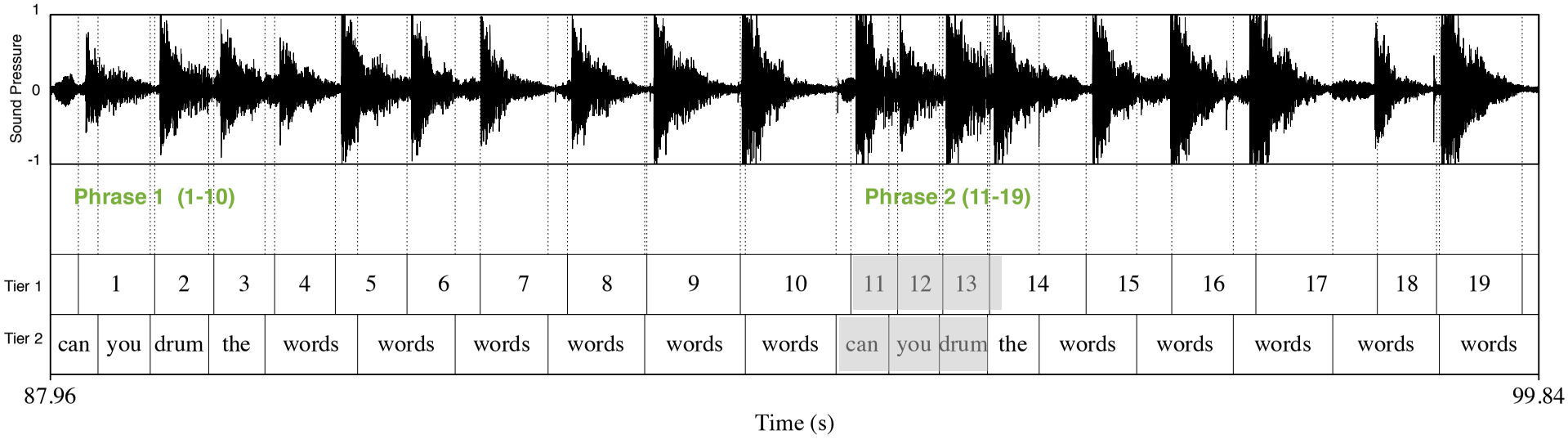


Figure 6.1: Annotated waveform showing Rachel's ability to drum word-based syllabic rhythms. The image shows the waveform of Rachel's drumming to phrases 1 and 2 of 'Can You Drum The Words?', combined with the stimulus, as sung by the researcher (Tier 2). Tier 1 shows the alignment to the stimulus in Tier 2 and the number of Rachel's beats. The boundaries to the beats in Tiers 1 and 2 indicate the duration beats. At beats 11-13, there is a clear difference in the relative duration of Rachel's beats in comparison to the beats either side and to her beats 1-3. The shading on these beats shows the similarity between Rachel's production (Tier 1) and the target pattern (Tier 2). The beats 11-13 are also closely aligned to the onset of the words.

6.3 Voice and prosody in imitation tasks

6.3.1 Sustained vowels

Data for all sustained vowels are given in Table 6.9. The target pitch was determined by the demonstration, but the mean SD of F0 indicates that Rachel sustained a relatively stable pitch. DUV was present in all but two vowels (/u/2 and /u/3) and was highest in vowels /a/2, and /a/4/: HNR was also lowest in these vowels, and both were produced with the highest overall proportion of interharmonics. Interharmonics were present for all measurable vowels and only /u/2 was produced with less than 100 ms of interharmonics at onset. Vowels /a/4 and /u/1 were produced with interharmonics that were not continuous with the interharmonics produced from onset. Examples of /i/1 (Figure 6.2) and /a/4 are (Figure 6.3) given below. The first example, /i/1 was produced with a low percentage of interharmonics, but with the lowest CPP and 1.5% DUV. The second vowel, /a/4, was produced with the highest proportion of interharmonics and lowest HNR, and with perceived diplophonia.

Table 6.9: The mean fundamental frequency and voice perturbation measures for Rachel's sustained vowels. Measures given are for mean F0 (Hz), SD of F0 (Hz), DVB (%), DUV (%), HNR (dB), jitter (%), shimmer (%) and CPP (dB). Two further vowels had been excluded as these were contaminated by another participant who sang at the same time. Rachel's production of two vowels (*) overlapped with the demonstration and the contaminated portion (shown in 'duration') is excluded from these measurements. The presence and duration (milliseconds) of single or complex (2+) interharmonics is shown, and their position in the vowel is indicated, as follows: onset (i; t=100 ms), medial (m: the period after onset, but excluding the final 100 ms), and final (t=100 ms). Where bifurcations are medial but are continuous with onset or offset, this is indicated with +. The interharmonic data in bold type indicates diplophonic production.

Vowel	Duration of overlap/ Duration of vowel, ms	Mean F0, Hz	SD of F0, Hz	DUV,%	HNR, dB	Jitter, %	Shimmer, dB	CPP, dB	1 Inter-H: Duration (ms) and position	2+ Inter-H: Duration (ms) and position
/i/1	**1387**	332.45	1.10	1.50	25.66	0.47	1.78	4.13	0.00	100 (i) +7 (m)
/i/2	**974**	267.02	1.29	2.14	24.57	0.57	2.26	5.65	0.00	100 (i) +6 (m)
/i/3*	907/ **1074**	246.26	0.97	0.92	24.43	0.37	4.45	5.34	n/k - onset contaminated	
/u/1	**1264**	299.03	2.50	1.32	26.71	0.59	2.66	5.40	182 (m)	100 (i) +49 (m)
/u/2	**1849**	269.27	0.83	0.00	28.15	0.25	2.03	6.10	96 (i)	0.00
/u/3*	691/**795**	252.23	0.72	0.00	30.34	0.24	1.88	6.72	n/k - onset contaminated	
/a/1	**1453**	295.57	0.92	1.14	21.29	0.40	4.76	4.62	0.00	100 (i)+56 (m)
/a/2	**1041**	275.45	12.13	7.20	18.40	0.63	4.28	4.26	0.00	100 (i) +104 (m)
/a/3	**1105**	257.21	1.79	3.02	20.89	0.45	2.68	5.01	0.00	100 (i)+53 (m)
/a/4	**975**	303.51	2.03	4.71	14.72	0.82	5.78	4.62	**622 (m)**	100 (i)+ 69(m)
Mean	**1191**		**2.34**	**2.20**	**23.52**	**0.48**	**3.26**	**5.19**		

Example 1: /i/1

Figure 6.2 shows the sustained vowel /i/1. The vowel was produced at a sustained pitch of 332.45 Hz which was her close to the demonstration (328 Hz, note E4). Onset was harsh (section 1: Figure 6.2) and several interharmonics are visible (marked with an arrow) between the main harmonic lines. The pitch contour (in blue) shows that pitch was unstable as the intensity (pink contour) rose. Rachel's production was otherwise perceptually modal and the pitch was perceived as stable during the main section (section 3). The pitch contour and formant contour (in red) rose towards offset (section 4) as the intensity declined.

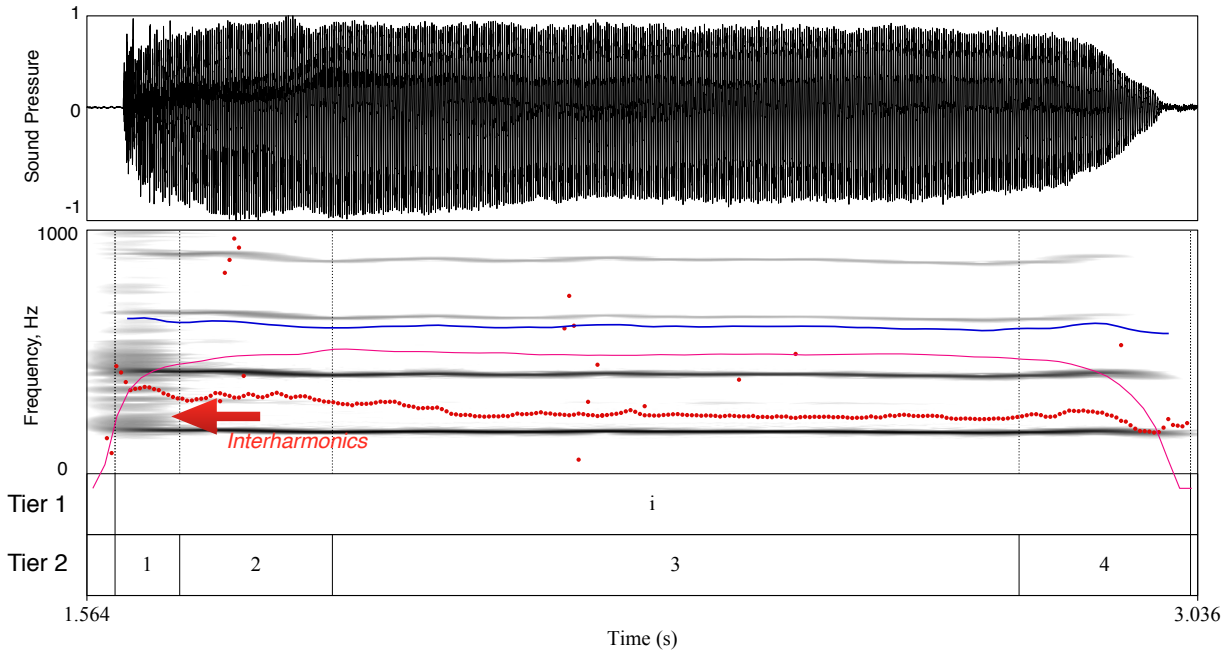
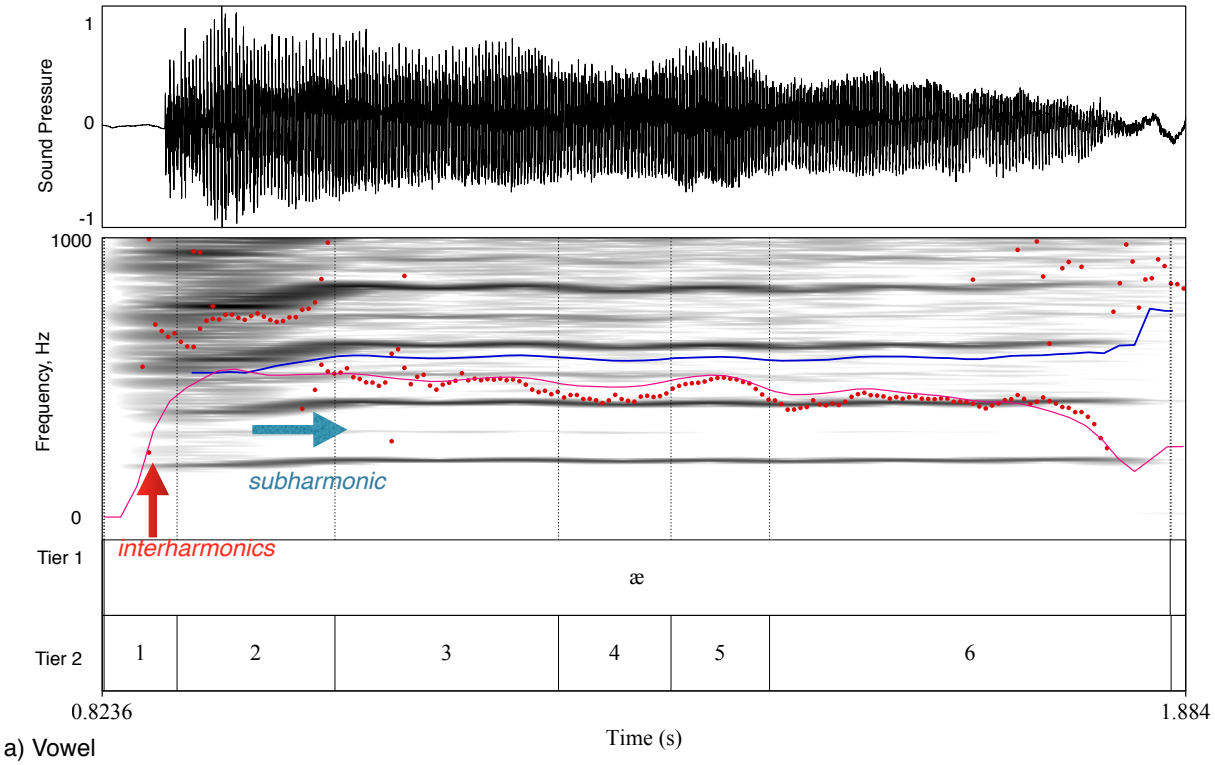


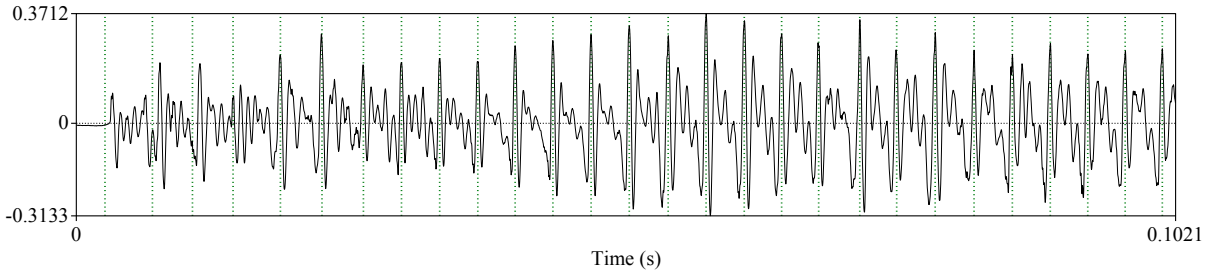
Figure 6.2: Waveform and annotated spectrogram showing Rachel's /i/ vowel at E4 (/i/1). The image shows the waveform, narrow-band spectrogram (0-1000 Hz, window length 150 ms, 45 dB), pitch contour (blue: 50-400 Hz), intensity contour (pink: 50-100 dB) and the first formant frequency (red dots) of Rachel's imitated /i/ vowel. Tier 2 shows the sections referred to in the text. Interharmonics are visible in section 1 of the spectrogram between the main harmonics: this is marked with an arrow.

Example 2: /a/4

Rachel's production of the /a/4 vowel is shown in Figure 6.3. The vowel was produced with harsh, glottal onset (section 1) and was mostly harsh in quality throughout. Multiple interharmonics are evident at onset (sections 1-2): the detail of the waveform (Figure 6.3b) shows the complexity of the signal between pulse lines at onset. One subharmonic is visible between the main harmonics in sections 2, 3 and 4 (marked with a blue arrow). The pitch (blue), intensity (pink) and formant (red) contours are unstable and their contours are very similar to each other, with peaks occurring in sections 3 and 5. Rachel's pitch rose as the intensity declined towards offset (section 6).



a) Vowel



b) Detail of waveform at onset (Section 1).

Figure 6.3: Waveform and annotated spectrogram showing Rachel's /a/ vowel at G4 (/a/4). a) The image shows the waveform, narrow-band spectrogram (0-1000 Hz, window length 150 ms, 45 dB), pitch contour (blue: 50-400 Hz), intensity contour (pink: 50-100 dB) and the first formant frequency (red dots) of Rachel's imitated /a/ vowel. Tier 2 shows the points of interest referred to in the text. Multiple and single interharmonics are indicated with red and green arrows, respectively. b) The image shows detail of the waveform of section 1: the pulse lines are marked in green.

6.3.2 Pitch and interval matching

In her initial assessment (Figure 6.4: R1) of matching single pitches Rachel sang consistently below target pitch (in blue) and did not repeat all pitches. In assessment two (R2) she produced all pitches more closely to target and matched notes D4 and E4. When matching intervals (Figure 6.5), she reproduced the correct number of pitches and the pitch direction. The *so-me* interval was accurately reproduced in the second assessment and all performances in the second assessment were closer to the absolute pitch of the targets than in the initial assessment.

Figure 6.6 and Table 6.10 show Rachel's pitch changes when she was asked to sing a five-note ascending and descending scale from a visual representation. Her notes were equivalent to half-semitones and tones (Table 6.10). After a demonstration of the scale, she repeated the task and

Figure 6.4: Graph showing Rachel's accuracy when matching pitches during initial and final assessments The fundamental frequency of the target pitch (in Hz) is shown in blue and Rachel's mean F0 is shown for productions in her first assessment (R1) and final assessment (R2), which were six weeks apart. The target pitch is shown in blue and the same note is shown at an octave below, in orange.

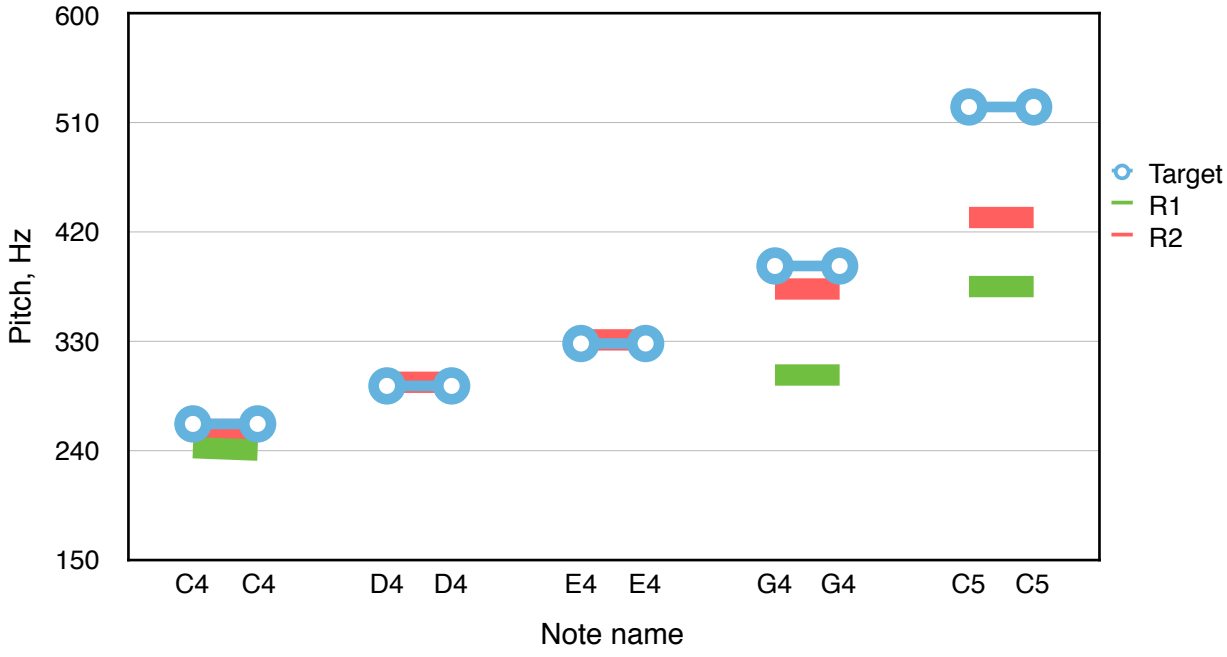
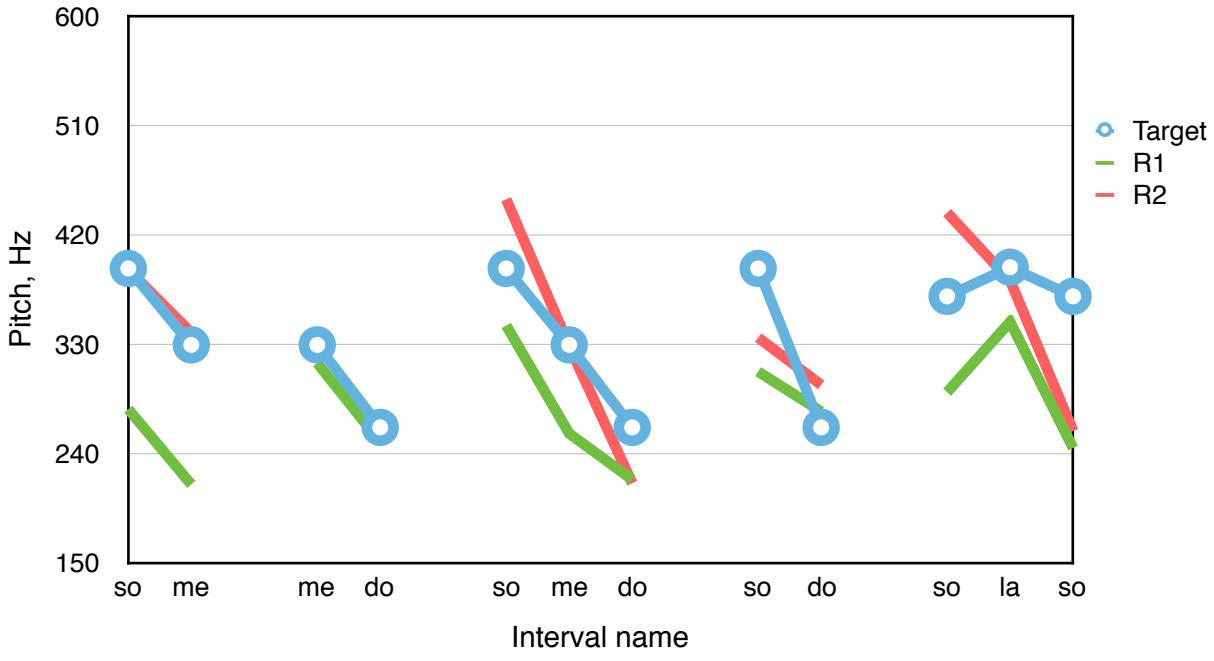


Figure 6.5: Graph showing Rachel's accuracy when matching melodic intervals during first and final assessments. The fundamental frequency of the target pitch (in Hz) is shown in blue and Rachel's mean F0 is shown for productions in her first assessment (R1) and final assessment (R2), which were six weeks apart. The target pitch is shown in blue and the same note is shown at an octave below, in orange.



produced tones and semitones (Table 6.10). Figure 6.6 shows that Rachel produced two notes in section 3, then stopped. With a prompt for the first two notes in the ascending scale, she independently completed the upward scale (section 4), and the descending scale (section 5).

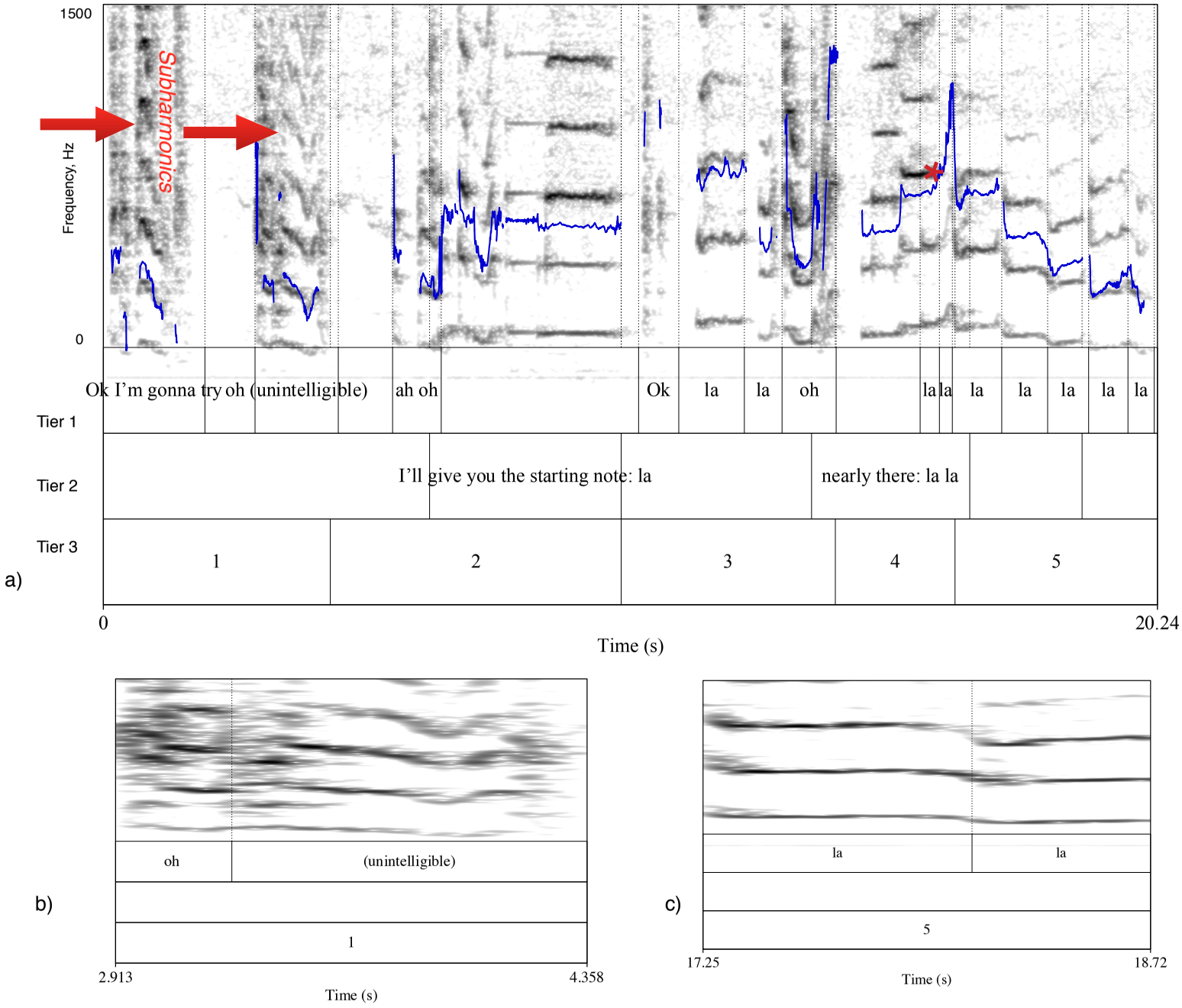


Figure 6.6: Annotated spectrogram showing changes in Rachel's pitch and voice quality when reading a 5-note scale from visual notation. The images show the spectrogram (0-1500 Hz, window length 150 ms, 50 dB), and pitch contour (in blue: 150-450 Hz) of Rachel's production. In Figure a), Tier 1 shows Rachel's words and interactions with the researcher, whose comments are shown in Tier 2. Tier 3 highlights points of interest. The pitch contour contains a number of breaks as a result of environmental noise and overlapping speech. The researcher demonstrates the ascending scale in section 4 and Rachel completes the ascending pattern where the red asterisk is placed (section 5): she independently produces the descending scale (section 5). The pitch contour is broken during two of her notes but Table 6.10 presents the perceived pitch of these notes.

A high number of interharmonics were produced in Rachel's spoken sections that are not apparent in her singing of 'la': the presence of interharmonics is indicated by red arrows on Figure a). Detail of the unintelligible section is shown in Figure b); detail of section 5 ('la') is shown in Figure c) for comparison. Each detail is 1.45 seconds in duration for comparison.

Figure 6.6a also shows Rachel's interactions with the researcher (Tier 2). Rachel's voice when speaking was produced with interharmonics and was perceptually harsh in quality (Figure 6.6a, section 1; Figure 6.6b), whereas her sung notes were perceptually modal in quality and were produced without interharmonics (Figure 6.6a, section 4; Figure 6.6c).

Table 6.10: Rachel's tonal production when 'reading' a 5-note descending scale. The table shows the mean F0 (Hz) produced by Rachel when singing a five-note descending scale from a visual image without assistance, and from the same visual image after a sung demonstration of the scale. Rachel's change in mean F0 is shown in semitones (ST) and in whole tones (T) and semi-tones (S): the target ST and tone change of the descending scale is shown for comparison.

Target changes		**Rachel: Independent**			**Rachel: After demonstration**		
ST change	**Tone change**	**Frequency, Hz**	**ST change**	**Tone change**	**Frequency, Hz**	**ST change**	**Tone change**
n/a	**n/a**	257			285		
2	**T**	222	2.5	T	248	2.4	T
1	**S**	192	2.5	T	224	1.8	T
2	**T**	187	0.5	S/2	202	1.9	T
2	**T**	181	0.6	S/2	190	1.1	S

6.3.3 Imitation of words of increasing syllable length

Table 6.11 gives the key features of pitch changes and relative duration of vowels in words of increasing syllable length, for the demonstration and for Rachel's imitation. Rachel's nPVI-V, pitch range, SD of pitch range, and her intonation contours as indicated by the percentage of rise and fall, were similar to the demonstration. In the demonstration and the imitation, words of three syllables were produced with the greatest contrast in duration of consecutive vowels (nPVI-V), and with a predominantly falling intonation, wider pitch range (ST) and increased variability of pitch range (SD of ST).

Table 6.11: The prosodic profile of words in the Syllables test, as demonstrated to Rachel and as imitated by her. The table shows the mean values of nPVI (variability of vowel duration), pitch range (semitones), intra- and inter-syllable trajectory, rate of pitch change, and percentage of pitch rises and falls. The mean values are shown for the demonstration and for Rachel's imitated production.

No. of Syll	**No. of nuclei**	**nPVI: vowel**	**Pitch range, ST**	**SD of ST**	**Traj Intra-syllable**	**Traj Inter**	**Rises, %**	**Falls, %**
1	**9**	31	7.5	1.5	16.6	0.0	11.1	0.0
demo	**9**	39	8.7	1.8	19.8	0.0	25.0	0.0
2	**18**	31	7.5	2.3	21.6	35.7	5.6	0.0
demo	**16**	37	8.6	2.3	21.2	10.2	25.0	0.0
3	**22**	42	15.6	4.5	23.9	40.4	0.0	16.0
demo	**27**	44	16.8	6.0	18.4	46.9	0.0	4.8
Mean percentage of		**87%**	**89%**	**82%**				

Figure 6.7 shows Rachel's imitation and the demonstration of words of three syllables. Rachel's pitch contours are similar to the demonstration, but the changes in intensity contour differ in four words: these occur in the second syllable of *lovingly, hardening*, *thickening;* and in the final syllable of *pleasingly*. Furthermore, consecutive peaks of intensity are similar to each other in Rachel's production and some three-syllable words are produced with only two distinct syllables.

Measures of voice quality changed as the target number of syllables increased (Table 6.12). In words of two and three syllables, jitter, shimmer and HNR reduced; but DVB and DUV increased. Mean F0 and SD of F0 were highest in words of three syllables, in line with the demonstration (Figure 6.7). The number of syllables produced with interharmonics decreased with target syllable length (Table 6.13). The number of consonants Rachel produced, irrespective of their accuracy (Table 6.3), also decreased with target syllable length.

Table 6.12: Mean voice perturbation measures for words of increasing syllable length. The table shows mean data and SD for fundamental frequency (F0, Hz), DUV (%), DVB (%), HNR (dB), jitter (%) and shimmer (%), and CPP (dB) for each vowel in each syllable that Rachel produced, arranged by the target number of syllables.

No. of target syllables	No. of vowels produced	Mean F0, Hz (SD)	DUV,% (SD)	DVB, % (SD)	HNR, dB (SD)	Jitter, % (SD)	Shimmer, % (SD)	CPP, dB (SD)
1	**9**	197.35 (12.47)	3.62 (8.04)	0.00 (0.00)	11.44 (4.85)	4.69 (1.89)	1.93 (1.56)	4.69 (1.89)
2	**18**	208.33 (27.18)	4.22 (8.45)	1.69 (6.25)	12.71 (5.34)	4.77 (1.39)	1.59 (1.15)	4.58 (1.54)
3	**22**	257.07 (81.08)	4.31 (14.28)	4.71 (15.62)	17.20 (5.25)	5.01 (1.68)	0.86 (0.68)	5.13 (1.38)

Table 6.13: Voice production and speech production in words of increasing syllable length. The table shows the number of the syllables in words that were produced with interharmonics, in words of increasing target syllable length. The table also shows the number and percentage of the target consonant sounds Rachel produced, regardless of their accuracy. The number of pitch and intensity contours that Rachel imitated are shown, expressed as a percentage of the number of words.

No. of target syllables	No. words in the set =	No. and percentage of syllables produced with interharmonics	No. and percentage of consonant sounds produced	No. and percentage of intensity contours reproduced	No. and percentage of pitch contours reproduced
1	**9**	8/9 = 89%	16/20 = 84%		7/9 = 78%
2	**9**	13/18 = 72%	20/27 = 74%	5/9 = 55%	9/9 = 100%
3	**9**	11/22 = 50%	20/37 = 54%	6/9 = 67%	8/9 = 89%
Mean		**65%**	**67%**	**61%**	**89%**

Example: 'please, pleasing, pleasingly'

The 'please' set was selected as an example (Figure 6.8) because the set contained the only example of dysfluency from the task, and because Rachel appeared to have difficulty in perceiving the three-syllable word 'pleasingly', despite having produced these sounds in words of one and two syllables.

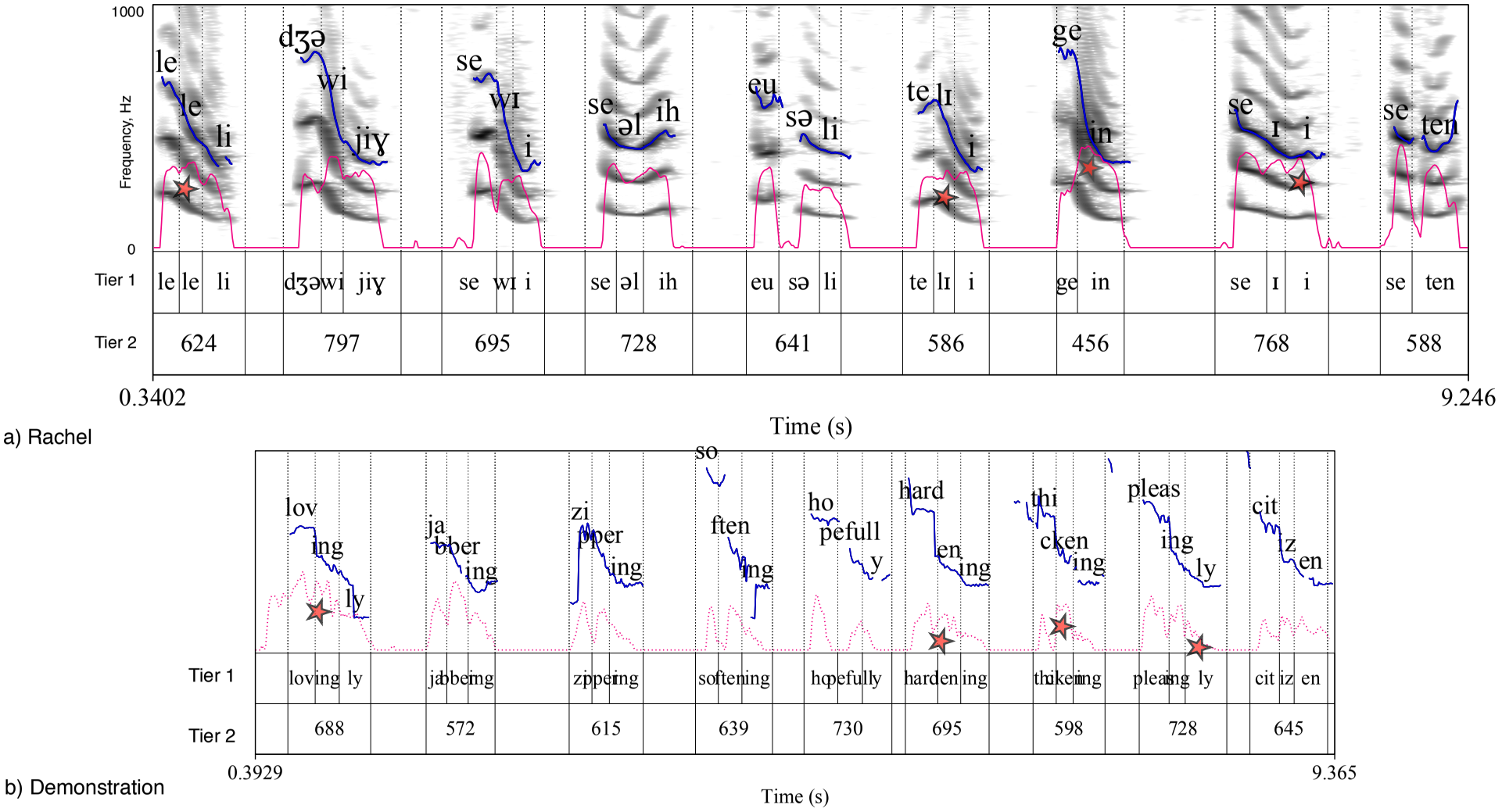


Figure 6.7: Annotated spectrograms showing the demonstration and Rachel's imitation of words of three syllables. The above image (a) shows the narrow-band spectrogram (0-1000 Hz, window length 150 ms, 50 dB) of Rachel's production of words of three syllables. The pitch contour (75-500 Hz) is shown in blue, alongside her phonetic production (in black), which is also shown in Tier 1. The intensity contour (40-100 dB) is shown in pink. The duration of words (milliseconds) is shown in Tier 3. The researcher's demonstration of these words is shown below (b), with target syllables shown and the duration of the word: as the recording was quieter, the overall intensity contour (in pink) is reduced and pitch contour (in blue) is affected in places by environmental noise. The red stars positioned alongside the intensity contours in four words indicate where the production of intensity between successive syllables differed between Rachel's imitation (a) and that of the demonstration (b).

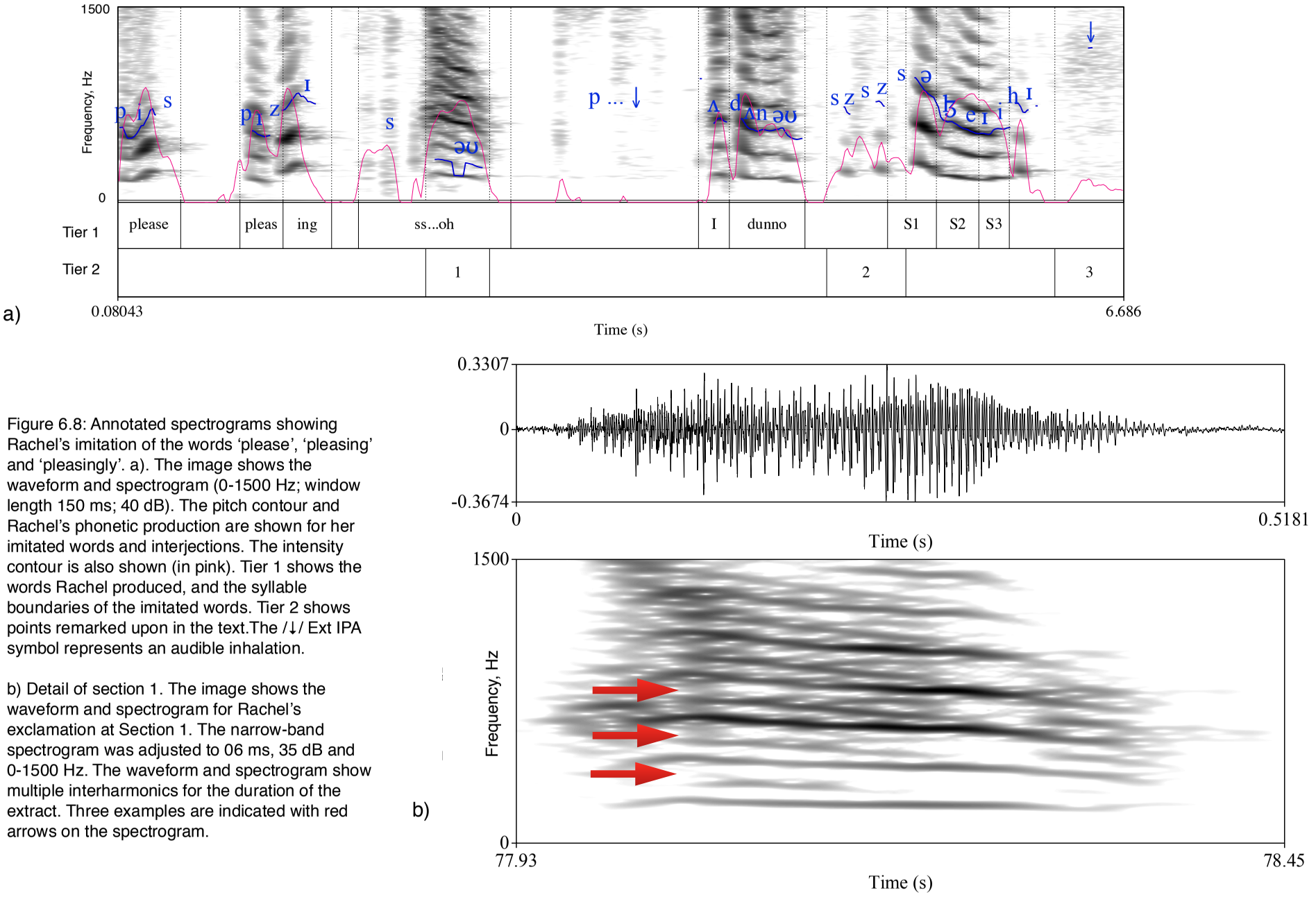


Figure 6.8: Annotated spectrograms showing Rachel's imitation of the words 'please', 'pleasing' and 'pleasingly'. a). The image shows the waveform and spectrogram (0-1500 Hz; window length 150 ms; 40 dB). The pitch contour and Rachel's phonetic production are shown for her imitated words and interjections. The intensity contour is also shown (in pink). Tier 1 shows the words Rachel produced, and the syllable boundaries of the imitated words. Tier 2 shows points remarked upon in the text.The /↓/ Ext IPA symbol represents an audible inhalation.

b) Detail of section 1. The image shows the waveform and spectrogram for Rachel's exclamation at Section 1. The narrow-band spectrogram was adjusted to 06 ms, 35 dB and 0-1500 Hz. The waveform and spectrogram show multiple interharmonics for the duration of the extract. Three examples are indicated with red arrows on the spectrogram.

Please: [pi̱s]; pleasing: [pɪ̱zi]

Both words were produced with a reduced initial cluster, /pl/, and with a perceptually 'harsh' voice quality. Two different qualities of interharmonics are visible on the spectrogram of *please* (Figure 6.9). A subharmonic is visible during section 2 (Figure 6.9), and multiple interharmonics are present in sections 3 and 4. These coincide with a change in perceived voice quality and a perceived rise in pitch and a change in the amplitude of the waveform (Figure 6.9).

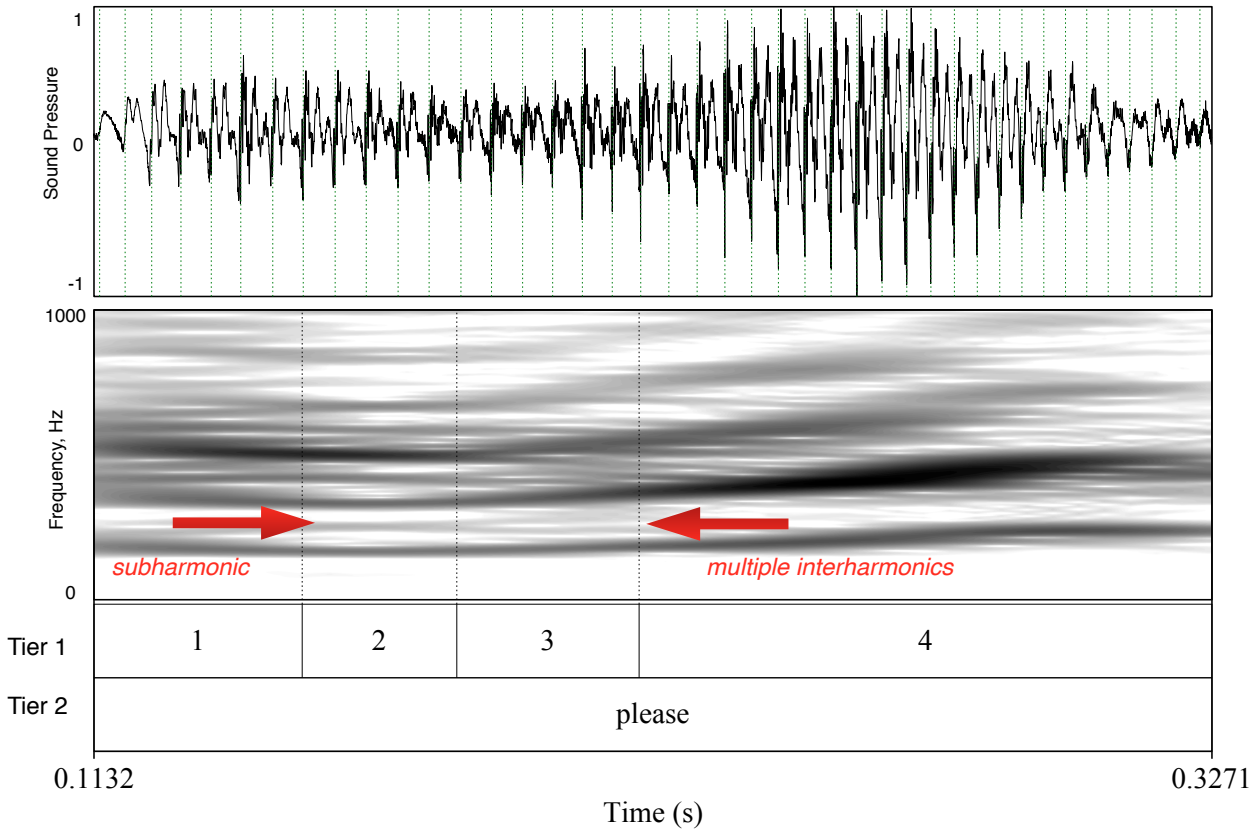


Figure 6.9: Waveform and annotated spectrogram showing a detail of Rachel's imitation of 'please'. The image shows the waveform of the /i/ vowel in 'please' and the narrow band spectrogram, adjusted to give increased resolution (0-1000 Hz; window length 70 ms, 45 dB). The pulse lines are shown on the waveform. One subharmonic is visible between the harmonics in section 2. In sections 1, 3 and 4 (Tier 1), multiple interharmonics are visible. These are indicated with red arrows.

Pleasingly: [zs..zs..səɮɪi:]

Rachel produced 'pleasingly' with a voiceless alveolar fricative, /s/ (Figure 6.8: point 1), before saying *oh* and *I dunno* (Tier 1, point 1). Her production of the word 'pleasingly' (labelled as S1, S2, S3: Tier 1) was initially dysfluent and she produced audible inhalations [↓] before and after her final production (Figure 6.8, between sections 1 and 2, and section 3). She reproduced the intonational contour of the word 'pleasingly' (Figure 6.8: in blue) and produced the three syllables with appropriate rhythmic duration. These were differentiated by two vowel qualities. Her exclamation of *oh* was ventricular in quality and the waveform and spectrogram show multiple interharmonics (Figure 6.8 b). A subharmonic is clear on the detail (Figure 6.8b) that reflects her diplophonic production.

6.3.4 Nonword imitation

Rachel's phonetic reproduction of nonwords was inaccurate and she produced fewer syllables than target (Table 6.14). She retained the perceived stress of the final syllables in some words (e.g. *pennel*, *doppelate* and *contramponist*). She reproduced the pitch contours of all but *contramponist* and *reutterpation* with a rising intonation, in contrast to the demonstration, and her pitch range was lower than the demonstration. Her phonation was atypical in seven words, with audible ventricular phonation in *glistering* and *reutterpation* (Table 6.14).

Table 6.14: Phonetic and prosodic features of the demonstration of nonwords (CNREP) and of Rachel's imitation. The table shows the perceived stress, the mean F0 (Hz) and the pitch range (in semitones) of the demonstrated word. Rachel's production of each target word is shown with a phonetic transcription with Ext IPA symbols marking perceptually creaky vowels. Mean values of F0 (Hz) and SD of F0 are given for the demonstration and Rachel's production. Rachel's pitch contour is given and this is marked to show whether her production was the same (S) or different (D) to the stimulus. The percentage of vowels in each word produced with interharmonics is shown. *'Blonterstaping' was produced with initial dysfluency of 323 ms and no periodic phonation. **'Contramponist' was repeated before Rachel imitated it.

Demonstration	Mean F0, Hz	Production (IPA transcription)	No. of syllables	Mean F0, Hz	Pitch Contour	Percentage of vowel produced with interharmonics,%
Pennel	242	pæ̰ʊ	2/2	186	rise: D	100
Doppelate	198	**tɪ̰**læd	2/3	222	flat: D	63
Glistering	206	**tʃe**əwɪt	3/3	200	rise: D	100
Blonterstaping*	226	s*..s**tɔ̰**bḭ	2/4	n/k -no periodic phonation	flat: D	100
Cont**ram**ponist**	219	**bɛ**ləlɪst	3/4	197	**fall**: S**	61
Per**plis**teronk	240	dʒ̰egə̰ɪ**wʌg**	4/4	200	rise: D	100
Defermi**cat**ion	245	**æ̰**keʃən	3/5	228	rise: D	100
Reutter**pat**ion	207	æ̰**ɸə̰**ʃə̰n	3/5	216	fall: S	100
Reutter**pa**tion 2	204	æ**pḛɪ**ʃə̃	3/5	200	rise: D	87
Mean	**221 Hz**		**25/35**	**206 Hz**		**90%**
Percentage of target			**71%**			

Rachel's imitation of nonwords was comparable to the demonstration for the degree of pitch changes within syllables vowels (Traj Intra) and within vowels (% rise and fall) (Table 6.15). Values of pitch range (ST), pitch changes between syllables (Traj Inter) and nPVI were lower than the demonstration. However, four samples were excluded (*glistering, reutterpation 1* and *2, blonterstaping*) from the analysis as a result of pitch breaks, which caused the program to mis-identify the number of nuclei (see Chapter 3, Section 3.6.3).

Table 6.15: Prosodic profile of the demonstration of nonwords (CNREP) and of Rachel's imitation. The table shows the values of the demonstration and Rachel's imitation for number of nuclei, nPVI-vowel, pitch range and standard deviation (semitones), pitch changes within (Traj Intra, %) and between (Traj Inter, %) syllables, and the mean percentage of rise and fall of pitch within vowels. Rachel's accuracy in reproducing the features is expressed as a percentage of the demonstration.

	No. of nuclei	nPVI: vowel	Pitch range, ST	SD of ST	Traj Intra-syllable, %	Traj Inter-syllable, %	Rises in vowel, %	Falls in vowel, %
Rachel	**17**	27	6.4	2.0	16.2	18..2	5.0	0.0
Demo	**19**	46	13.6	4.0	18.6	35.7	5.3	0.0
Percentage of demo	**89%**	**59%**	**47%**	**50%**	**87%**	**51%**	**94%**	**100%**

Voice perturbation measures (Table 6.16) show a high degree of voice breaks and shimmer in nonwords. Words with four or more target syllables were produced with higher values in DVB, jitter and shimmer, and with lower HNR and CPP values. Words of three target syllables were produced with lower measures of perturbation (HNR, jitter, shimmer and CPP). Rachel expressed difficulty with the task, irrespective of target syllable length (as measured by the frequency of *I don't know:* see Chapter 3, Section 3.7.2).

Table 6.16: Mean voice perturbation measures for Rachel's vowels in CNREP, arranged according to increasing target syllable length. The table shows the mean measures and SD of DUV (%), DVB (%), HNR (dB), jitter (%), shimmer (%) and CPP (dB). The table also shows the number of times that Rachel expressed difficulty ('I don't know').

Number of syllables in the word	Number of words, n =	DUV (SD), %	DVB (SD)%	HNR (SD), dB	Jitter (SD), %	Shimmer (SD), %	CPP (SD), dB	'I don't know', n=
2	**1**	0.00	6.82 (5.43)	8.82 (2.65)	2.96 (1.59)	16.48 (8.23)	4.51 (1.02)	1
3	**2**	0.00	2.03 (3.97)	10.91 (4.34)	1.37 (0.82)	12.69 (6.39)	4.80 (0.88)	2
4+	**6**	1.38 (4.58)	15.31 (16.33)	8.52 (4.30)	3.22 (2.16)	18.19 (8.69)	3.33 (1.31)	4

6.3.5 Phrase imitation

Rachel imitated five of the seven target phrases but she reproduced just 36% of the words and 44% of the syllables (Table 6.17). Rachel's intonational stress was different to the demonstration, but this was affected by the reduced lexical content. Her speech production was slower than the demonstration and all phrases were produced with atypical voice quality. The percentage of interharmonics in the initial two phrases was lower that in subsequent phrases.

The prosodic data (Table 6.18) confirm that Rachel's productions were different to the demonstrations: vowel duration was more variable (nPVI); Rachel's pitch range (ST) was reduced; pitch changes within and between syllables were lower (Traj Intra- and Traj Inter-syllable); and her pitch changes within vowels were different in direction.

Table 6.17: Phonetic and prosodic features of the demonstration and of Rachel's imitation of phrases (Newton, 1999). The table shows the target phrase and the stress pattern, mean pitch range (ST) and speaking rate (syllables per second) of the demonstration and of Rachel's imitation. The similarity of Rachel's reproduction of the pitch contour is indicated by S (same) and D (different). Primary stress is indicated in bold type: secondary stress is underlined. EXT IPA symbols indicate perceptually creaky production on vowels and nasals, and nasal production that is more pronounced than usual for Rachel. The percentage duration of the vowels that were produced with interharmonics that were visible at 35 dB is also given.

Demonstration: words and stress		**Pitch range, ST**	**Rate, syll/sec**	**Rachel's production: IPA transcription and stress**	**No. of syllables**	**Pitch range, ST**	**Pitch contour**	**Rate, syll/ sec**	**Percentage interharmonics in vowels,%**
1	You must clean your **teeth**	**10.5**	**3.6**	aiɲ̃ kli:ɪn **tif**	4	4.9	D	2.2	44
2	Mary's shoes are **clean**	**8.6**	**2.7**	glin	1	1.9	D	1.6	39
3	I gave the elephant a ba**na**na	**13.2**	**4.8**	**ɛ̃le**fən **æ̃næ̃**	5	1.2	D	2.3	100
4	**Clare** ate all her lunch	**7.8**	**3.1**	æ̃ **lʌntʃ**	2	3.6	D	2.4	100
5	My uncle is a **far**mer	**10.0**	**5.3**	fa**mə**	2	2.6	D	2.8	100
Mean		**10.0**	**3.9**			**2.8**		**2.3**	**76**
Percentage of demo					**44%**	**28%**	**0%**	**59%**	

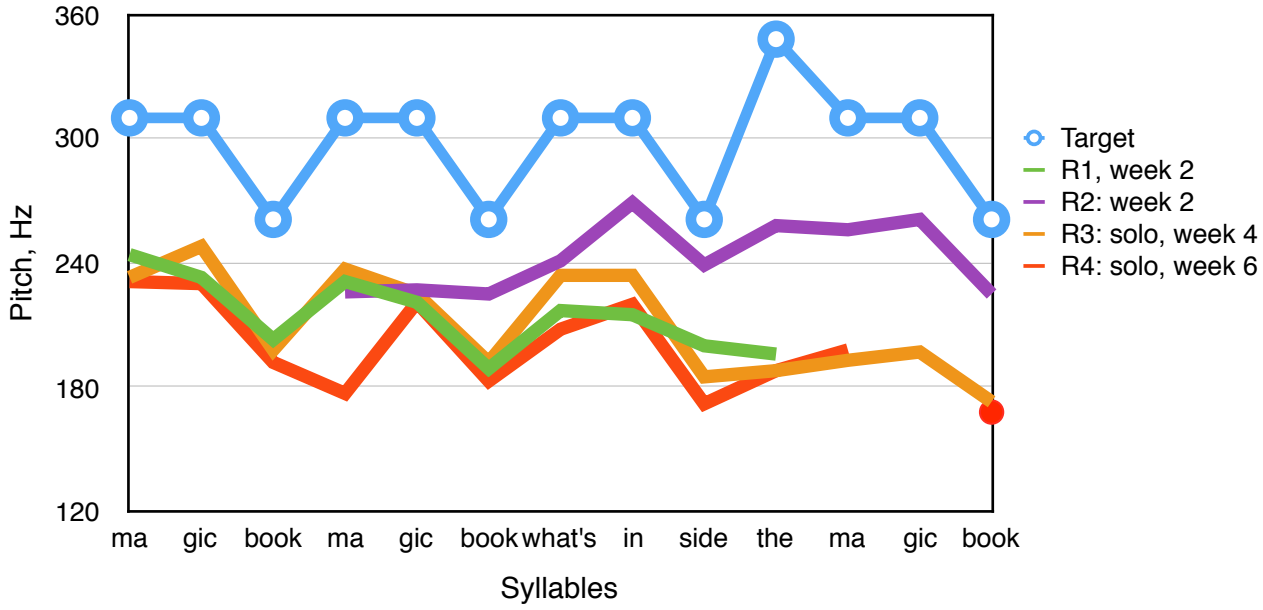
Table 6.18: Prosodic profile of phrase imitation (Newton, 1999). The table shows the values of the demonstration and Rachel's imitation for number of nuclei, nPVI-Vowel, Pitch range and SD (semitones), pitch changes within and between syllables, and the mean percentage rise and fall of pitch within vowels. The value of relevant measures is expressed as a percentage of the demonstration.

	No. of nuclei	**nPVI: vowel**	**Pitch range, ST**	**Std Deviation of ST**	**Traj Intra-syllable, %**	**Traj Inter-syllable, %**	**Rises in vowel, %**	**Falls in vowel, %**
Rachel	**13**	53	8.3	1.6	9.9	7.1	7.7	0.0
Demo	**28**	37	24.4	0.0	25.6	57.9	0.0	14.3
Percentage of demo	**46%**	**143%**	**334.0%**	**34.7%**	**38.7%**	**12.3%**		

6.3.6 Song imitation: Magic Book

The starting pitch of Rachel's productions of the song were consistent, and she produced the melodic contour of the song in most productions (Figure 6.10). She produced all the words in her third performance (R3, in orange). This production was also most accurate musically for the third phrase *what's inside*: the contour and intervals are very close to that of the target (in blue). Rachel recalled more of the words in each production, but she needed prompts at the start of phrases. Her final production (R4) was produced without prompts and with Rachel clapping the beat and rhythm, but she had practised the song with assistance prior to this. Although she recalled more words in the final production, her phonetic production was more accurate in her first rendition (Figure 6.11).

Figure 6.10: Graph showing the development in Rachel's recall of the words and melody of 'Magic Book', over a five-week period. The figure shows the melodic contour of Rachel's performance of the song, over a five-week period. R1 was her production in week 1, and her final production was R4 (week 6). The target contour is shown in blue.



Target	Rachel 1	Rachel 4
1. **ma**gic **boo**k	1. **meɪ**dʒi **bʊk**	1. **hæt**i **bʊk**
2. **ma**gic **Book**	2. **ei**ziu **bək**	2. **hæt**i **bʊk**
3. **what**'s In**side** the	3. əʊ (unintelligible)	3. wai is**ai** də
4. **ma**gic **Book**	4. **a**likə**læ**bwai	4. **mæd** bə
5. **ab**raca**da**bra	5. **kəɪæ**bwə	

Figure 6.11: Phonetic transcription and syllabic stress of Rachel's first and final productions of 'Magic Book'. The figure shows the phonetic detail of Rachel's first imitation of the song and her final, unprompted, rendition. Primary stress is indicated by bold type. Rachel's voice quality was perceived as almost typical throughout.

The features of four of Rachel's productions of the phrase *magic book* are shown in Table 6.19. Her segmental production was least accurate in her later productions but articulatory rate increased. The relative duration of C:V remained close to that modelled. Her production of nine (of sixteen) pitch direction intervals were accurate to within 0.5 ST (Table 6.19: bold type). The relative duration of the words were accurate for productions R2-R4.

Rachel's third production of *Magic Book* (Figure 6.12) was the most accurate in terms of segmental production (Table 6.19). However, her segmental production of the word *magic* changed within each production and she produced the word 'what' as *why*. Rachel tapped the beat of the first four phrases and the rhythm of *abracadabra*. Her taps were almost synchronised with the onset of words that fell on the strong beat (see those in bold, Figure 6.12), marking the pulse, but the timing

between her taps was not regular (Figure 6.13). She produced the correct number of taps to mark the rhythm of *abracadabra* but she produced an additional spoken syllable.

Table 6.19: Articulatory and musical features of the demonstration of the first phrase of 'Magic Book', and of four of Rachel's productions. The table shows segmental, articulatory, melodic and rhythmic features of the demonstration and Rachel's four productions by session (R1, R2, R3, R4). Segmental production details the number of consonant sounds (C) Rachel produced in the phrase ('magic book, magic book'), and their relevance to the target sound in terms of manner and place. Articulatory features measure the rate of production of consonants and vowels and the relative duration of these. The mean change in semitones between syllables is given (melodic features), and direction is indicated (downwards (-), upwards (+)). Data in bold text indicate pitch reproduction that was accurate within 0.5 ST. The table also shows the mean duration and relative duration of syllables in the words (rhythmic features). *Andrew (Chapter 4) accompanied Rachel on this task.

Demonstration: key features			**R1: prompted**	**R2: prompted***	**R3: solo, tapping beat**	**R4: solo, tapping beat**
Segmental production	**Number of C sounds**	**10**	7/10	7/10	9/10	8/10
	Number of relevant sounds	**10**	7	7	6	4
Articulatory features	**Total duration C. ms**	**856**	1008	838	1003	729
	Articulatory rate: sounds per second	**11.8**	7	8	9	11
	Total duration V, ms	**1577**	1695	2021	2156	1618
	Average duration of C:V	**100:184**	100:170	100:241	100:200	100:222
Melodic features	***Magic*: Mean ST change**	**0**	**0**	**0**	1.1	**0**
	***Book*: Mean ST change**	**-3**	**-2.6**	-1	-4	-3.3
	***Magic*: Mean ST change**	**0**	-1	**0**	0.9	+3.1
	***Book*: Mean ST change**	**-3**	**-2.6**	**-2.8**	**-2.7**	**-3.4**
Rhythmic features	**Mean duration of *magic*, s**	**398**	398	482	400	336
	Mean duration of *book*, ms	**797**	594	553	736	527
	Relative duration of syllables *ma:gic*	**1:1**	1:1	1:1	1:1	1;1
	Relative duration of words *magic:book*	**1:2**	2:3	1:1	1:2	1:2

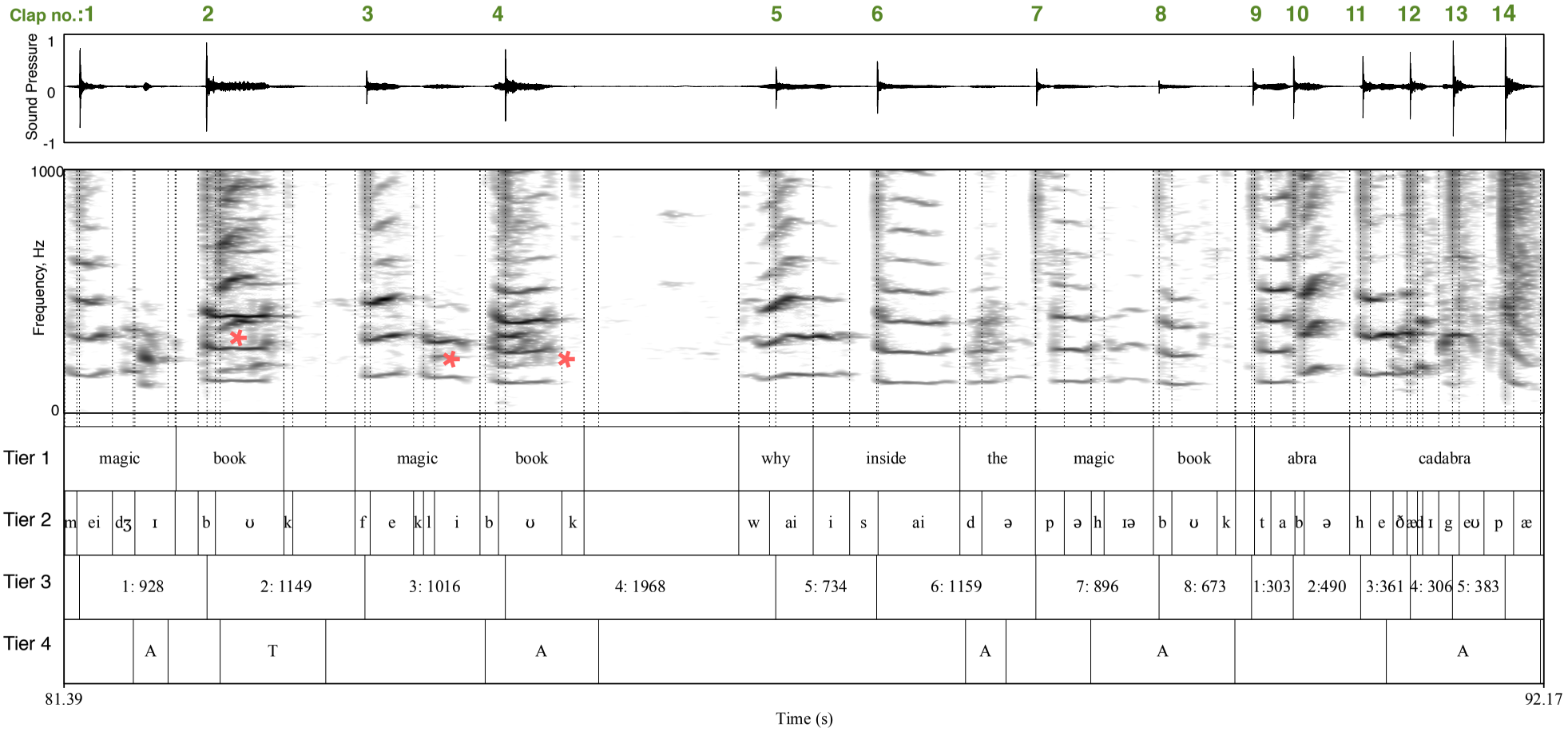


Figure 6.12: Annotated waveform and spectrogram showing Rachel's production of 'Magic Book' (R3). The image shows the waveform and spectrogram (0-1000 Hz; window length 150 ms; 40 dB) of Rachel's third production of 'Magic Book'. in which she simultaneously tapped the pulse and the rhythm of 'abracadabra'. The words she produced are shown on Tier 1 and the sounds on Tier 2. Tier 3 shows the number and duration of the pulses (1-8) (milliseconds), and of the rhythm of 'abracadabra' (1-6). Each tap that Rachel produced is also marked, in green, above the waveform. Tier 4 shows where the section is contaminated by Andrew singing and whispering (A) and by the researcher (T) asking him to remain quiet (T). The subharmonics that are visible on the spectrogram in phrase 1 (marked with red asterisks) are caused by environmental sounds. There is also contamination on the spectrogram from the claps that Rachel produced as she sang.

6.4 Prosody and voice in production of speech and song from LTM

This section will present details of Rachel's spontaneous comments and her spoken and sung productions of the song 'Happy Birthday'. The prosodic profiles of two spontaneous comments (1 and 2) and of the song *Happy Birthday* in spoken and sung conditions are given in Table 6.20. Her comments were similar to each other in pitch range (ST) and pitch movements (SD of ST) but Comment 2 was produced with a faster speaking rate, and with reduced nPVI values. Her spoken and sung productions of *Happy Birthday* were produced with higher nPVI values, a wider pitch range (ST) and increased pitch changes within (Traj Intra) and between syllables (Traj Inter).

Table 6.20: Prosodic profiles of Rachel's spontaneous comments and of her spoken and sung productions of 'Happy Birthday'. The table shows prosodic features of Rachel's spontaneous comments and her spoken and sung productions of Happy Birthday. The table shows the number of nuclei, nPVI-vowel, pitch range and standard deviation (semitones), pitch changes within (Traj Intra, %) and between (Traj Inter, %) syllables, and the mean percentage of rise and fall of pitch within vowels.

	Comment 1	Comment 2	Happy Birthday, spoken	Happy Birthday, sung
Number of nuclei	16	36	15	16
nPVI: vowel	47	36	73	68
Pitch range, ST	12.3	12.0	17.0	12.8
SD of ST	2.6	2.9	7.1	4.9
Traj Intra-syllable	16.7	14.5	24.4	12.6
Traj Inter-nucleus	14.3	15.5	37.2	39.8
Rises, %	0.0	0.0	0.0	0.0
Falls, %	18.8	8.3	13.3	0.0
Rate, syll/sec	3.3	5.1	6.4	3.8

6.4.1 Comments

Rachel's comment to Kerry (Chapter 5) was not completely intelligible (Figure 6.13). The pitch contour on Figure 6.14 (in blue) shows that there is a decline of overall pitch within each phrase (marked in Tier 3), and a global decline in pitch over the whole phrase. Rachel emphasised the following syllables: S1, *good* (phrase 1); (unintelligible), *times* **(**phrase 2); *you win* and the first

Rachel	Interpretation
R1: (**S1** S2)'s [kiæðə gɤət]	(name) is rather **good**
R2: [həwɪs tuɪæp aʔ **taɪ**]	her is too (*unintelligible*) at **times**
R3: [juɪ e **fæ**tə]	you win X **Fa**ctor

Figure 6.13: Phonetic transcription and syllabic stress of Rachel's comment to Kerry (Comment 1). Rachel produced three phrases (R1-R3), which are marked with primary stress (bold) and secondary stress (underlined). An interpretation is given, based on the context and the prosodic information (Figure 6.14). Rachel produced another participant's name and this has been substituted with S1 and S2, to convey the perceived syllabic stress. Rachel's voice quality was close to typical throughout.

syllable of *factor*, (phrase 3). The peaks of amplitude of the waveform coincide with relative peaks in pitch and with relatively longer vowel duration on the syllable. The combination of features results in the perceived stress.

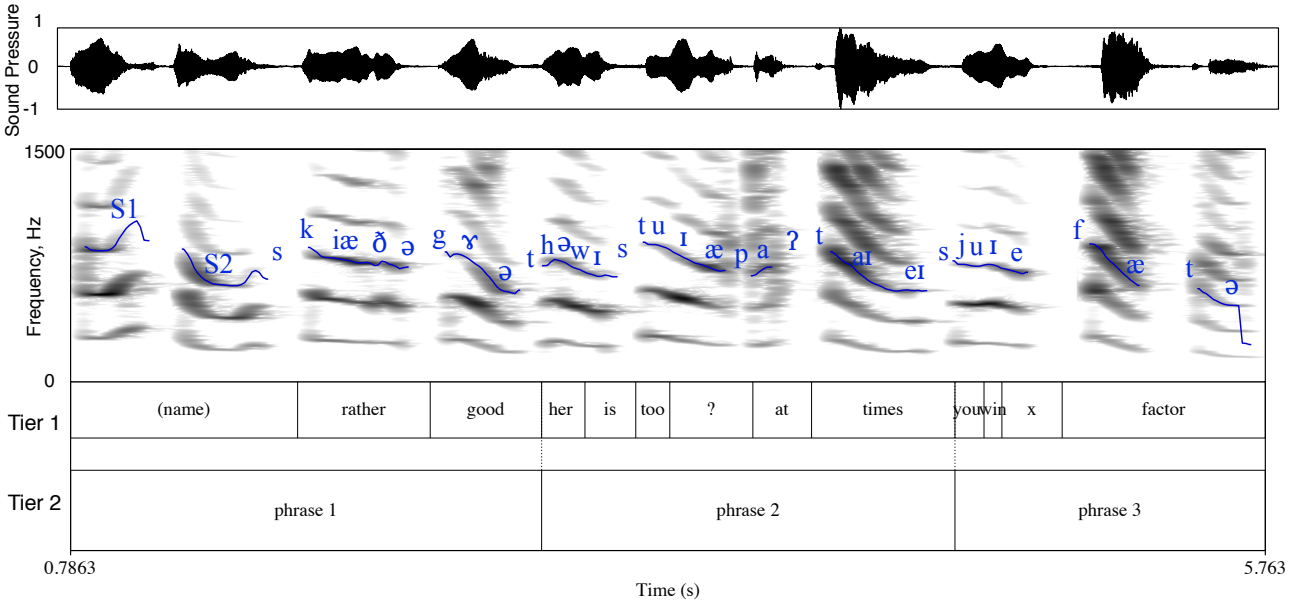


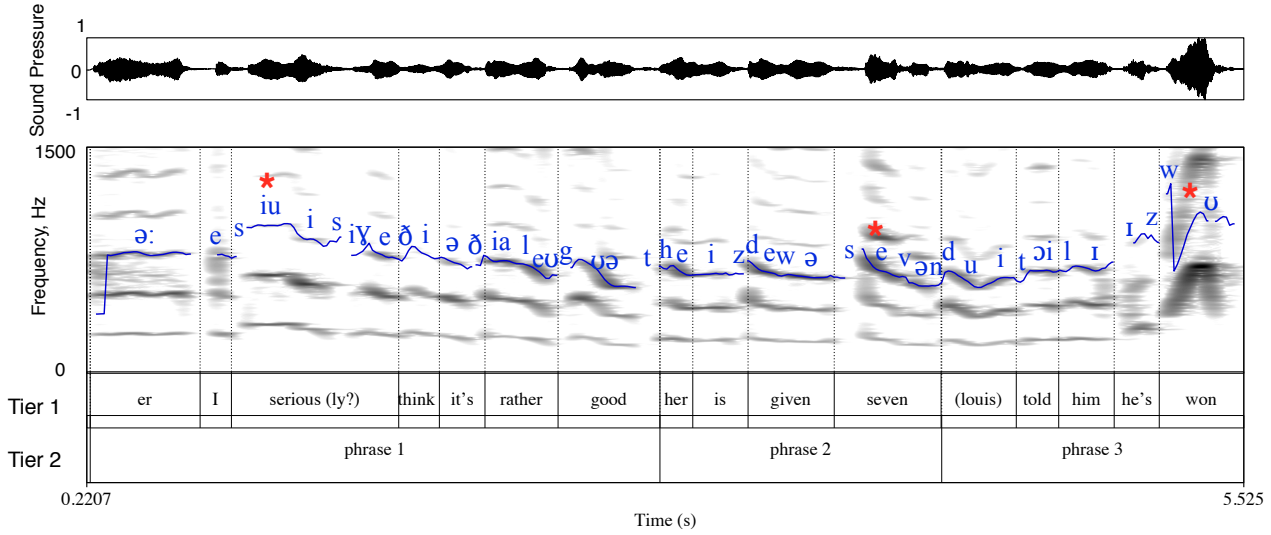
Figure 6.14: Waveform and annotated spectrogram showing Rachel's comment to Kerry (Comment 1). The syllables of Kerry's real name are given as S1 and S2 to protect her identity. The spectrogram (0-1500 Hz; window length 150 ms; 40 dB) is marked with the pitch contour (blue: 50-500 Hz) and phonetic production. The tiers below show the words (Tier 1), where known, and perceived phrase groups. The unintelligible section is annotated with ? (Tier 1). The phrase number are given in Tier 2.

Comment 2 (Figure 6.15, Figure 6.16) was produced rapidly, at a mean rate of 5.1 syllables per second (Table 6.20). Rachel's speech intelligibility was poor (Figure 6.15) but it was known that Rachel was a fan of the TV show, *X Factor.* Therefore, her meaning in phrases 1-3 was inferred from a combination of her speech sounds and the context. Rachel used a mixture of pronouns in her interactions with the researcher (Figure 6:15: R4, R5) which further hinders the interpretation of her speech. Rachel's production was melodic and her voice spanned an octave (Table 6.20). An

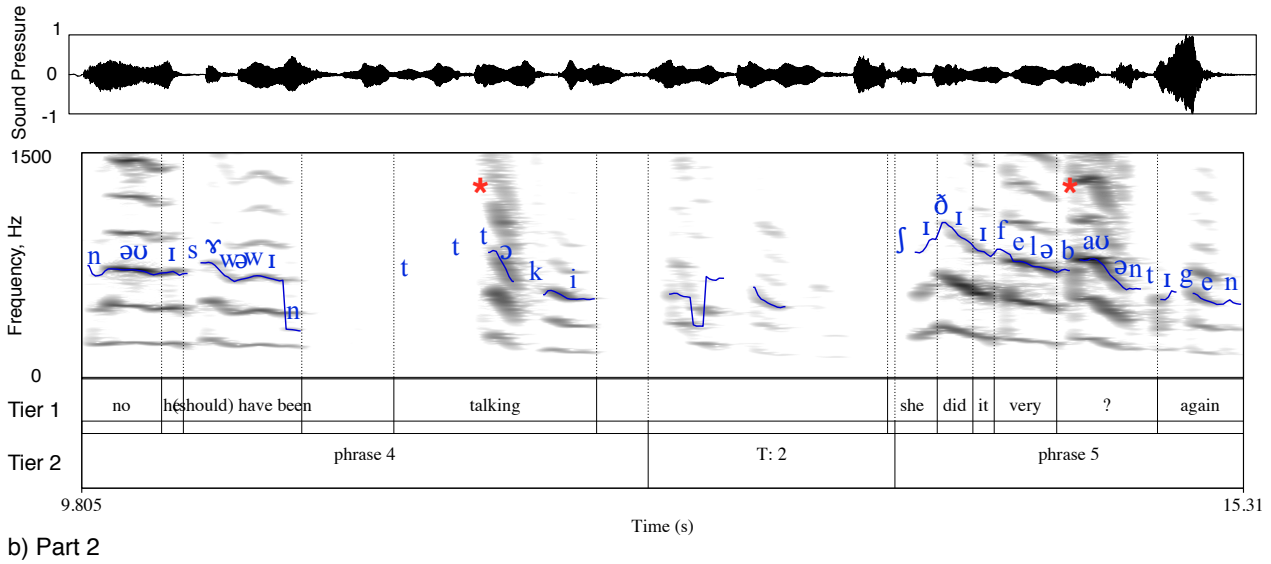
Rachel	Translation
R1: [ə̰: e **siui**si ɣe ði əð ialeʊ g̲ʊ̲ə̲t]	R1: er I **seriou**sly (?) think it's rather good.
R2: [he iz dewə **se**vən]	R2: Her is given **se**ven.
R3: [dui tɔɪl i i̲z **wʊ**]	R3: Louis told him he's **won**!
T: OK, so… did he sing it well?	
R4:[nəʊ ɪ s̲ɤ̲wə̲wɪn t..t..**tɔ**ki]	R4: **no**, he should (nt?) have been **talk**ing
T: who was talking?	
R5: [ʃi **ðɪ** ɪ **fə**lə **baʊ**ənt ɪgen]	R5: she **did** it **ve**ry (*unintelligible:* **brilli**ant?)

Figure 6.15: A phonetic transcription of Comment 2, showing Rachel's comments on Robert's singing (R1-5) and the researcher's comments (T). Not all words were intelligible and possible interpretations are given. The syllables that were produced with perceived stress are shown in bold. Incidences of creaky voice quality are indicated with ExtIPA symbols.

increase in her F0 coincided with a relative increase in intensity and vowel duration: these factors resulted in the perceived stress of five words (Figure 6.16a, marked with an asterisk).



a) Part 1



b) Part 2

Figure 6.16: Waveforms and annotated spectrograms showing Rachel's Comment 2, part 1 (a: top) and part 2 (b: below). The image spectrogram (0-1500 Hz; window length 150 ms; 40 dB) is annotated with the pitch contour (in blue: 50-500 Hz) and sounds that Rachel produced. The tiers below show the words that Rachel produced, with possible interpretations indicated in parentheses. The stressed syllables that coincide with relative increase in pitch and intensity are marked with an asterisk. The ? (phrase 5) marks a section of unintelligible speech. Tier 2 shows the phrase boundaries (also, see Figure 6.15).

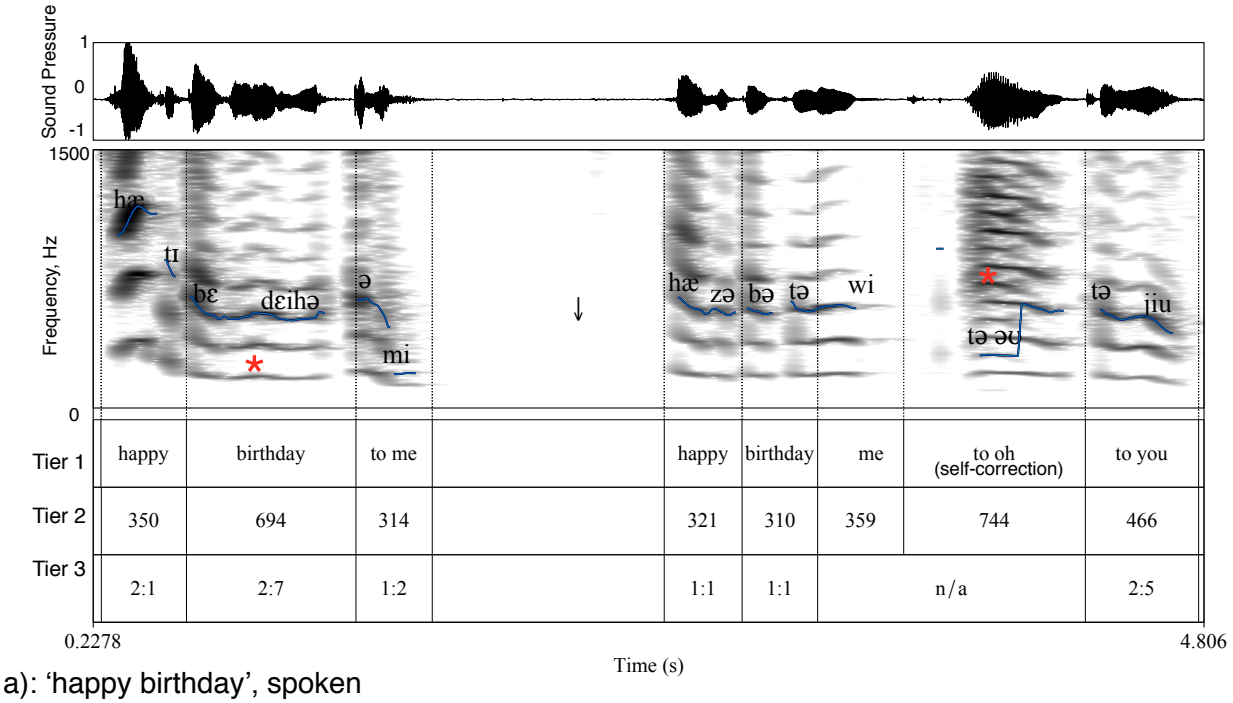
6.4.2 Happy Birthday

Rachel's production of the speech sounds in *Happy Birthday* varied within and between the productions (Figure 6.17). Her use of pronouns was inconsistent but she corrected herself in the spoken version (Figure 6.17: phrase 2). Although her voice quality was almost modal in quality, there are some features of atypical voice quality and production. For example, additional noise on the spectrogram indicates a breathy quality at the onset of *birthday* (Figure 6.18b). There is a visible change in the intensity and pitch contours during the second syllable of *birthday* (Figure

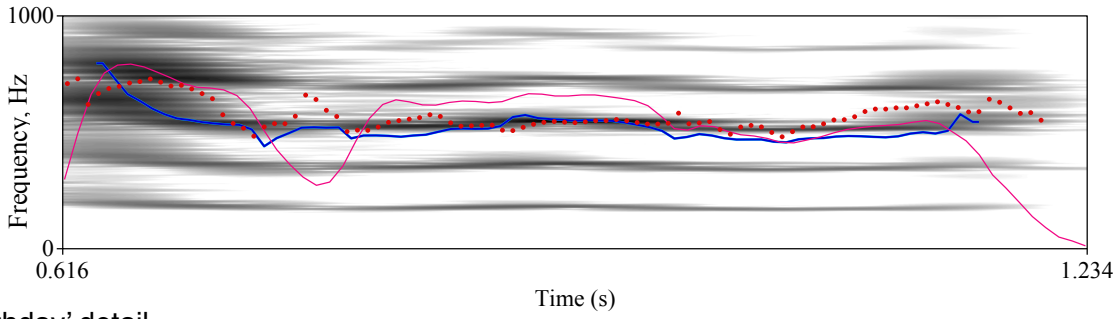
6.18b) that coincides with a ‘shimmer’ quality. Rachel’s interjection *oh* [{V!! **əʊ** V!!}] (Figure 6.17: phrase 2) was produced with a subharmonic (Figure 6.18c) and audible diplophonia. When singing, Rachel repeated the same melodic contour on three phrases (Figure 6.19): this differs from the anticipated melody. Her changes in pitch direction between the words *to* and *you* was consistent in all four phrases, and echoed the changes in pitch direction of the song. Rachel’s pitch changes when speaking were different to the song in other phrases in both direction and interval size (Table 6.21). The duration of the syllables between words differed to each other and within each production (Table 6.21, Figure 6.18, Figure 6.19: Tier 3) and differed to the target production of the song (Table 6.21).

Target	Sung	Spoken
1: happy **birth**day **to you**	1. hætə bei **ti** ju	1: **hætɪ bɛ̰**diəhə ə mi
2: happy **birth**day **to you**	2. hætə **wɛtɪ** ti ju	2: **hæ**zə bətə wi tə {V! **əʊ**
3: happy **birth**day dear…		V!} tə **jiu**
4: happy **birth**day **to you**	4. hætə **bei**təʊ **tu** bi	
	Interharmonics: 67%	*Interharmonics: 70%*

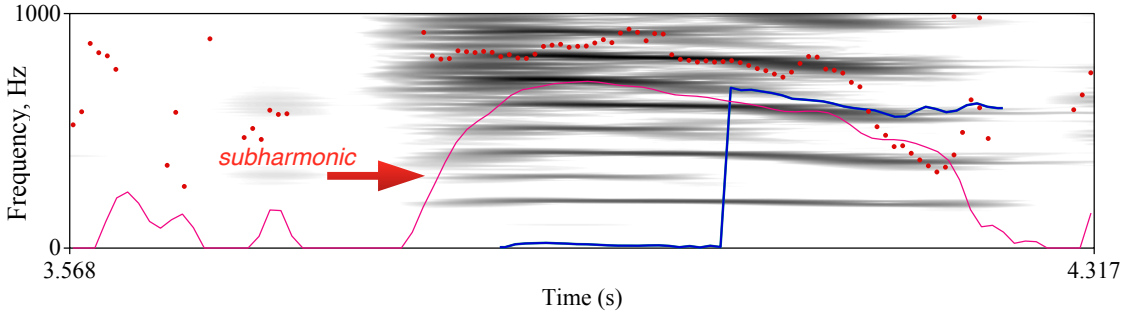
Figure 6.17: Phonetic transcription and syllabic stress of Rachel’s production of ‘Happy Birthday’, sung and spoken. The target production of each phrase, numbered 1-4 is shown. The phonetic transcription of Rachel’s sung and spoken productions are shown and are marked with intonational stress (primary stress in bold text, secondary stress in underlined text), and with Ext IPA symbols showing creaky voice (æ̰). VoQS indicated ventricular phonation (V!). The mean percentage of the duration of vowels that were produced with interharmonics is also shown.



a): 'happy birthday', spoken



b): 'birthday' detail



c): 'to oh' detail

Figure 6.18. Waveform and annotated spectrograms showing details of Rachel's spoken production of 'Happy Birthday'. The image a) shows the waveform and spectrogram (0-1500 Hz; window length 150 ms; 40 dB) of Rachel's entire spoken production (a: above), annotated with the pitch contour (50-500 Hz) and phonetic transcription of syllables (in blue). Below this, Tier 1 shows the words Rachel produced, Tier 2 shows the duration of speech segments (in milliseconds) and Tier 3 shows the approximate relative duration of syllables between subsequent syllables. The spectrogram shows multiple interharmonics during Rachel's first production of 'birthday' and at the point of her self-correction, marked by a red asterisk. Details are given for the word *birthday* (b) and her self-correction, *to oh* (c, below). The settings were changed for greater resolution, as follows: spectrogram (0-1000 Hz), F0 (100-250 Hz), intensity (50-90 dB). The images also show the first formant (standard settings). The spectrogram shows some additional noise in b) and there is a loss of intensity during the second syllable of *birthday* that coincides with a fluctuation in the pitch contour. Figure 6.18c shows a break in the pitch contour that coincides with the onset of a subharmonic and perceived diplophonia. The first subharmonic is indicated with an arrow.

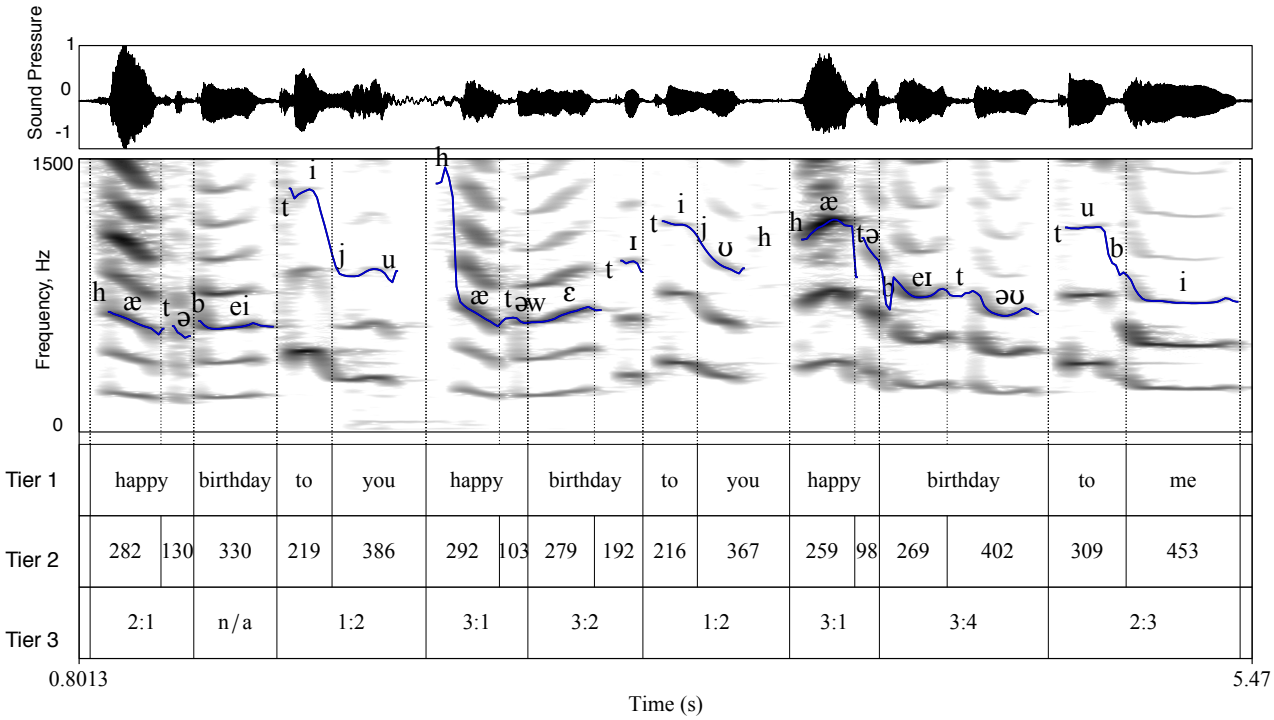


Figure 6.19: Waveform and annotated spectrograms showing Rachel's sung production of 'Happy Birthday'. The spectrogram (0-1500 Hz; window length 150 ms; 40 dB) is annotated with the pitch contour (50-500 Hz) and phonetic transcription of syllables (in blue). Tier 1 shows the words Rachel produced, Tier 2 shows the duration of syllables (in milliseconds) and Tier 3 shows the approximate relative duration of syllables between subsequent syllables.

Table 6.21: Relative timing of syllables in Rachel's productions of 'Happy Birthday' in sung and spoken conditions. The table shows the change in mean pitch (in semitones) between syllables (ST), and the relative duration of syllables (RD) in comparison to the shortest syllable duration in the phrase: these are marked in Figures 6.19 and Figure 6.20 (Tier 2 and 3). The (+) indicates a rising melodic interval the (-) indicates a falling melodic interval: those in bold text highlight pitch reproduction that was accurate to within 0.5 ST. Mean speaking rates are given for each phrase and for each task.

Target			**SInging**						**Spoken**			
Phrase 1/2/3			**Phrase 1**		**Phrase 2**		**Phrase 3**		**Phrase 1**		**Phrase 2**	
Syllable	**Target pitch change and direction**	**Target relative duration**	**Pitch change, ST**	**RD**	**Pitch change, ST**	**RD**	**Pitch change, ST**	**RD**	**Pitch change, ST**	**RD**	**Pitch change, ST**	**RD**
ha	**0**	**3**		2		3		3		2		1
py	**0**	**1**	+2	1	-1	1	-3	1	-2.1	1	-1.6	1
birth	**+2**	**2**	-13		0	3	-4	3	+1.2	2	+1.3	1
day	**-2**	**2**	0		+0.5	2	**-2.4**	4		7	+6	1
to	**+5/+6/+5**	**2**	+2.8	1		1	+9	2	+13.9	1	+4	2
you	**-1**	**4**	-5	2	-0.75	2	-7.8	3	-7	2	-4	5
Mean speaking rate (syllables/second) per phrase			**4.1**			**4.9**		**3.3**	**4.4**		**5.5, excluding correction**	
Mean speaking rate (syllables/second) per task			**4.1 syll/sec**						**4.9**			
Mean pitch range (min-max), ST			**12.80**						**17.00**			

Discussion

Section 6.5 will examine Rachel's rhythmic and melodic abilities in musical tasks, and in spoken and sung phrases and words. Rachel's performance in music-motor tasks will be discussed, followed by a comparison between her temporal perception and production abilities in speech and song. Her melodic abilities in music and speech tasks will be discussed separately. Common factors affecting prosody and speech production will be considered in Section 6.6: these include the influence on learning of memory, attention, cognitive load and Rachel's emotional response to tasks. Section 6.6 will also consider the evidence that Rachel may benefit from song-based music tuition in order to support her speech. The case study will conclude with a summary and consideration of whether and how Rachel's speech may benefit from music or singing.

6.5 Rhythm, voice and melody in speech and song

6.5.1 Temporal perception and production in music

Rachel's mean Anticipation Time scores (AT, ms) at 104 bpm and 132 bpm (Table 6.6) were close to norms (Reifinger 2006; Loehr and Palmer, 2009). The accuracy of her inter-tap intervals (ITI of 1%) at 132 bpm and 160 bpm (Table 6.6) were also within adult norms (<2% of ITI: Repp, 2005; Repp and Su, 2013). Her mean ITI at 104 bpm (89% of the target tempo) exceeded these norms. However, her performances at faster tempi were atypical. At 160 bpm Rachel played 8% from antiphase (50%), on average (Table 6.6). Playing in antiphase can be more challenging than playing in phase (Pfordresher and Dalla Bella, 2011; Repp, 2005; Repp and Su, 2013), resulting in reported norms for adults of 10% AT (Repp, 2005). However, playing between the target beat can support production by subdividing the beat, effectively halving memory demands (Repp, 2005; Repp and Su, 2013). Likewise, Rachel's doubling of the period at 132 bpm may reduce attentional or memory demands. This could indicate a difficulty in forming a stable temporal template for relatively slow tempi. Rachel's poorer performance at 104 bpm (89% ITI, 9% of AT) would be consistent with this. Her performance could reflect a personal preference for the faster beat hierarchy, as is common to young children and nonmusicians (Drake et al., 2000; Jungers, Palmer, and Speer, 2002; Krumhansl, 2000; Repp and Su, 2013; Thaut, 2008). As previously discussed (Chapters 4 and 5), poorer performance at a slower tempo could result from immature motor development, from attentional differences, or from reduced memory span (e.g. Drake et al., 2000; Flaugnacco et al., 2014; Grahn and Schuit, 2012; Phillips-Silver, 2009; Repp and Su, 2013; Saito, 2001; Snyder, 2000, 2009).

Rachel's inconsistent ability to imitate clapped rhythms that exceeded three notes (Table 6.8) is in line with her digit span of 3 (Table 6.1). When auditory memory demands were supported with gestural demonstration and visual notation, Rachel drummed the rhythm of *Can You Drum The*

Words? with the appropriate timing (Table 6.7, Figure 6.1). Visuo-spatial modelling supports adults with DS in performing bimanual drumming tasks (Ringenbach et al., 2006). The inclusion of the words in the *Can You Drum The Words?* task provided additional verbal guidance. This can support rhythmic production in TD children (Grahn, 2012; Reifinger, 2006; Su, 2014) and people with DS in some tasks (Robertson et al., 2002). It is possible that the visual stimulus clarified the task expectation, and enabled Rachel to draw on her knowledge of word rhythms, as she did for the rhythm of *abracadabra* (Figure 6.12: Tier 3). The 'live' demonstration of the rhythms may have conferred an additional advantage as rhythmic motor performance is enhanced in social situations (Deutsch, 2013; Phillips-Silver et al., 2010; Repp and Su, 2013).

In contrast, visual modelling of the rhythm entrainment task through gesture did not help Rachel to entrain to the rhythm (Section 6.2.2). The ability to entrain to rhythms does not develop until 8-10 years (Drake et al., 2000; Hargreaves, 1986), and this task may have been beyond Rachel's current developmental level for motor-music tasks. Furthermore, regular rhythms are easier to process that irregular rhythms (Schellenberg, 2016), and may be simplified in memory (Deutsch, 1986b in Krumhansl, 2000). In this task, Rachel simplified the pattern by producing a more regular pulse, which is how children between 5-7 years respond to complex rhythmic patterns (Drake and Gérard, 1989). This suggests that Rachel's ability in this task was limited by her overall delay in cognitive and motor development. Given the apparent benefit of visuo-spatial modelling in the *Can You Drum The Words?* task (Figure 6.1), it is probable that her difficulties in rhythmic motor tasks stem from her memory deficit, in line with previous research Flaugnacco et al., 2014; Grahn and Schuit, 2012; Repp and Su, 2013; Saito, 2001; Snyder, 2001, 2009).

6.5.2 Temporal perception and production in speech and song

The data from the speech perception tasks suggest that Rachel has difficulty in discriminating some speech sounds (Table 6.1: test 13). Poor perception of some sounds may be affected by her suspected hearing loss (Section 6.1) but results from the word-discrimination task indicate that she also has a difficulty in discriminating sequences of sounds (Table 6.1: test 7b). Rachel's production of speech sounds was poor in all tasks (Table 6.3) which may be explained by perceptual difficulties, production difficulties, or by a combination thereof. However, she made a higher number of speech errors in imitation tasks than in naming pictures, especially in words of increasing syllable length and in non-words (Table 6.3, Figure 6.7). This indicates that a specific deficit in STM affects her ability to reproduce speech sounds.

Data from speech imitation tasks indicate that Rachel may have initial difficulty in imitating temporal details. Rachel omitted syllables in words of two syllables and above (Table 6.11) and she omitted the initial syllables in four of the five nonwords of four syllables and above (Table 6.14).

This is consistent with a study by Laws (1998), who found that the majority of DS children (mean age of 11;6 years) in the study were unable to reproduce nonwords of four syllables in CNREP. Laws (1998) argues that CNREP is a reliable indicator of phonological memory deficit in the group. Rachel's rhythmic production of nonwords differed to the demonstration (Table 6.15: nPVI) but she reproduced the relative syllable duration of most words of up to three syllables (Table 6.11: nPVI, Figure 6.7).

Rachel's rhythmic imitation of phrases was affected by poor recall of the words (Table 6.17). Her imitation of phrases in *Magic Book* was also affected initially by poor word recall (Figure 6.10, Figure 6.11) but she was able to recall and reproduce the words when prompted with the initial sound or syllable (Table 6.19: R1, R2). With practice, she improved in articulatory speed and in the accuracy of the relative duration of syllables (Table 6.19). Rhythm in songs is more 'overt' than in speech (Huss et al., 2011), and Rachel's perception of the word rhythms may have been supported by melodic cues that can aid perception, recall and production (Rinta and Welch, 2008; Schlaug et al., 2008; Stahl et. al., 2013). Furthermore, a rhythmic structure can also assist production of speech (Schlaug et al., 2008; Stahl et. al., 2013; Thaut, 2008) which may be due to heightened attention (Hawkins, 2014). Rachel's third production of *Magic Book* (Figure 6.10) and her sung production of *Happy Birthday* (Figure 6.19) were both produced with a discernible beat: however, her phonetic production remained inconsistent in both (Figure 6.11, Figure 6.17). The improvement in Rachel's articulation rate in *Magic Book* may have supported her rhythmic production, but this is distinct from her ability to recall and produce the speech sounds.

In all speech tasks, just three words were produced with initial dysfluency. In word imitations these were: *pleasingly* (Figure 6.10) and *blonterstaping* (Table 6.14), both of which Rachel initiated on a voiceless alveolar fricative, /s/. In her connected speech, she stuttered on the initial voiceless alveolar plosive, /t/, of *talking* (Figure 6.15, Figure 6.16b). However, Rachel's spontaneous speech was at times rapid (Table 6.20: Comment 2 — 5.21 syllables/second) and difficult to interpret. She demonstrates a hastening of speech (festination) in Comment 1 (Figure 6.14), which is a characteristic found in populations with dysarthria, and speech and motor disorders (Thaut, Peterson, McIntosh, and Hoemberg, 2014). The speech sounds and the number of syllables she produced in songs was inconsistent: e.g. her production of the word *birthday* varied between one syllable and three syllables (Figure 6.17). The lack of pauses between phrases in her spontaneous speech (Figure 6.16) and instances of festination are characteristic of 'cluttering'. Cluttering is a rhythmic disorder that has been identified in DS populations, often alongside stuttering (Kent and Vorperian, 2013; van Borsel and Vandermeulen, 2008). Research into its prevalence in DS is limited, but van Borsel and Vandemeulen (2008) reported symptoms in 78.9% of their 76

individuals with DS. The data indicate that a motor disorder affects the rhythmic timing of her speech.

Links between rhythmic motor production and speech production

Rachel's data across CNREP, TAPS (Table 6.1) and TRaCoL (Table 6.8) suggest that her memory for speech sounds, digits and rhythm is limited to three items. This indicates a deficit in memory for temporal information in musical and speech domains. However, she also has difficulties in motor production in segmentation tasks, and in combining rhythmic motor movements with speech output. In segmenting word rhythms, the relative duration and number of Rachel's claps were inconsistent even when her knowledge of the required number of taps was secure (Table 6.8: *monkey*, *caterpillar*). In clapping the rhythm of *caterpillar 2* (Table 6.7), she produced five claps; when saying and clapping the word, she produced two claps but the correct number of syllables. Conversely, when clapping and saying the word *abracadabra*, she produced the correct number of taps but produced an additional syllable (Figure 6.12: taps 10-14). A study by Hall and Gathercole (2011) showed that the concurrent production of speech and tapping of an unrelated manual motor rhythm can result in poorer performance in TD adults, as a result of additional cognitive load. They suggest that motor rhythm tasks and speech rhythm tasks may draw on the same resources required for motor planning or temporal coding, and that production in one domain can interfere with production in the other. Although the speech segmentation tasks in this study required Rachel to produce the same rhythm in both domains, it is likely that the dual task demands affected her ability to produce the motor rhythms. However, Rachel was able to sings the words and clap the pulse in her third and fourth productions of *Magic Book* (Table 6.19: R3 and R4, Figure 6.12). This improvement in coupling speech with motor rhythm is likely to stem from reduced processing demands as her knowledge of the words increased, and as a result of a more simple motor movement.

6.5.3 Phonation and voice quality

The presence of interharmonics during the initial 100 ms of eight sustained vowels and during the medial section of two sustained vowels (Table 6.9) indicate that Rachel has difficulty in initiating and sustaining modal phonation. The mean incidence of interharmonics was similar for words as for sustained vowels, but there was an increase in the number of subharmonics in some tasks (Table 6.5). For example, Figure 6.6 shows a difference between Rachel's speech (Figure 6.6a) and her the sung notes on 'la' (Figure 6.6b) in terms of the number of interharmonics produced; and there is a visible increase in the interharmonics Rachel produced when she expressed uncertainty or corrected a mistake (Figure 6.8b, Figure 6.18c). The increase in interharmonics reflects a change between a perceptually modal quality and a more creaky quality.

The data show no clear pattern of increases in perturbation or interharmonics within tasks. In the Syllables task, the mean duration of atypical phonation decreased as the target number of syllables increased, as measured by the percentage of interharmonics (Table 6.13) and by HNR, CPP and shimmer (Table 6.12). The percentage duration of vowels produced with interharmonics was higher in CNREP (Table 6.14) and in phrase imitation (Table 6.17), which does indicate a difficulty associated with task demands. However, the CNREP data show no clear change with production length, and they show a different pattern to that observed in the Syllables task. This will be discussed further in Section 6.6, in relation to memory demand, articulation and prosody.

6.5.4 Pitch production and perception in song and speech

Although Rachel did not complete the pitch perception tasks (Table 6.1: tests 26, 27, 28), her ability to reproduce single pitches, pitch direction, melodic shape and the relative interval of two-note melodies (Figure 6.4, Figure 6.5, Table 6.19) demonstrates accurate perception and mental representation of relative pitch. Her reproduction of three-note melodies (Figure 6.5) was inconsistent but this may result from her reduced digit span (Table 6.1). Accurate singing relies on auditory memory (Dalla Bella et al., 2011; Dalla Bella et al., 2012; Krumhansl, 2000; Snyder, 2001, 2009) and an impaired phonological WM, in particular, is associated with reduced accuracy in tonal rehearsal (Williamson et al., 2010). However, even TD adults with no musical training have difficulty in processing melodic sequences of over three tones (Williamson et al., 2010). Therefore, Rachel's ability to repeat sequences of four intervals in her first imitation of the first phrase of *Magic Book* (Table 6.19) indicates that her memory for melodic interval may be close to typical abilities. This ability may result from her knowledge of musical key, which underpins the learning of novel melodies (Deutsch, 2013; Dowling, 1978, 1982). Dowling's research indicated that if the musical key is learned then melodic intervals are mapped onto it, which effectively reduces the processing required by limiting the choices of note and interval (Deutsch, 2013; Krumhansl, 2000). Rachel's ability to sight-read a descending scale (Figure 6.6) from notation suggests that she does have tonal knowledge: after a demonstration, she did not reproduce the intervals in the given order, but her production of five notes and intervals of a semitone and three tones (Table 6.10) suggests learning of the musical key, rather than recall of the notes themselves.

Rachel's imitation of intervals were more accurate on a simple C-V combination ('la') than when singing intervals in songs. Producing intervals is easier in isolation than in songs (Welch, 2006) and easier still if segmental and lyrical content are reduced (Berkowska and Dalla Bella, 2009a; Welch, Sergeant and White, 1996, 1997, 1998; Welch, 2016). These pitch matching tasks are likely to be the least challenging in terms of cognitive demands (Dalla Bella et al., 2012) and will therefore reflect Rachel's optimum ability. However, her production of melodic intervals in *Magic Book* was similar to the target from the first session (Table 6.19). With the exception of her third

production (R3), in subsequent renditions she continued to reproduce three of the four intervals in the first phrase to within 0.5 ST. Her relatively poor reproduction in R3, in which she sang and tapped, may result from the additional processing demands associated with concurrent articulation and tapping tasks (Hall and Gathercole, 2011). However, in comparison to the accuracy of her melodic intervals in *Magic Book,* her melodic production in *Happy Birthday* was poor. Rachel reproduced the traditional pitch contour of the final two words in phrases 1 and 2, and the traditional melodic shape of phase 4 in her phrase 3 (Table 6.21) but her intervals were larger than the traditional melody. It is possible that during her initial learning of *Happy Birthday*, she may have processed the course-grained features of the song, but not the intervals. This characteristic is typical of 'musically-naïve listeners' (Thompson, 2013).

Rachel's ability to imitate the broad-grained features of melody extends to imitation of words. Her imitations of words of increasing syllable length were close to the demonstration in pitch range and movements (Table 6.11) and in contour (Figure 6.7). Her imitation of nonwords and phrases were less accurate and most were produced with a rising inflection, rather than the downward contour that was modelled (Table 6.14, Table 6.17). This is unlikely to reflect a failure to recall the intonational contour, given Rachel's secure melodic production in songs and her close imitation of pitch contour in three-syllable words (Figure 6.7). Instead, Rachel's ascending pitch contour suggests uncertainty or questioning: this uncertainty, in turn, may stem from her poor retention of the sounds or words.

6.6 Links between musical and speech domains

6.6.1 Difficulties common to speech and music domains

Planning and cognitive load

Speech planning at the prosodic level is influenced by the intended linguistic structure (Krivokapic, 2012). Rachel demonstrates an ability to plan short and complex phrases with intonational planning (Figure 6.14, Figure 6.16): even though the rhythmic and segmental structure of her phrase was atypical, the melodic changes indicate the phrase boundaries (Figure 6.14: phrase 1-3). The prosody of phrases may be planned independently of their lexical content which is mapped on later (Krivokapic, 2012; Redford, 2015). Mapping the words might be an area of weakness for Rachel. For example, her dysfluent production in the fourth phrase of Comment 2 (Figure 6.15) indicates a planning difficulty at the segmental level (Bosshardt 2006; van Borsel and Vandermeulen, 2008): this affects her rhythm, but her intonation of *talking* preserves the declination of the final phrase and allows Rachel to convey her meaning. Mis-timings between the phonological plan and motor movements can both result from and indicate cognitive load, particularly in tasks that require phonological coding (Bosshardt, 2006).

Cognitive load depends on working memory capacity and long-term memory (Sweller, 2011). The response to cognitive stress is idiosyncratic and speech output can be affected by task demands and by self-imposed pressure (Buchanan et al., 2014; Bosshardt, 2006). Although the effects of cognitive load are idiosyncratic, it is known to affect phonatory instability (Quatieri et al., 2015; Yap, 2010), speech rate (Buchanan et al., 2014) and articulation rate (Johnstone and Scherer, 1999). Rachel's response to phrase imitation and nonsense words (CNREP) demonstrate these effects (Table 6.14, Table 6.15, Table 6.16). In longer words in CNREP, the incidence of anxiety, as measured by the frequency of *I don't know* increased. The longer words were also produced with higher perturbation measures (Table 6.16) and with a ventricular quality. Voice quality is an indicator of stress (Aronson and Bless, 2009; Johnstone and Scherer, 1999; Patel et al., 2011; Roy, 2003; Yap, 2012) and Rachel's voice quality is likely to reflect her perception of the task (stress-response) but also reflects the cognitive load, that leads to stress. This may be directly linked to her deficit in STM, and the demands placed on her phonological WM in CNREP, but it may also reflect her perception of the task's difficulty.

Rachel's imitation of nonwords (Table 6.14, Table 6.15, Table 6.16) and phrases (Table 6.17; Table 6.18) were less accurate than demonstrations in terms of prosody and were atypical in voice quality. In contrast, her imitations of words of up to three syllables were similar to the demonstration in rhythm and melody (Table 6.8) and were produced with lower measures of voice perturbation (Table 6.11, Table 6.12, Table 6.13). In comparison to CNREP, the segmental content and rhythmic structure of words in the Syllables task were simpler. Additionally, the task involved repetition of the initial syllables, which may have reduced processing demands and articulatory planning demands. Therefore, it could be argued that there were fewer processing demands in the Syllables task, which resulted in better prosodic production. However, further comparison suggests that Rachel may have focussed her attention differently to each task. In the Syllables task, HNR and CPP values increased as target syllable length increased (Table 6.12); in CNREP, these values decreased (Table 6.14). The voice and prosodic data indicate that Rachel responded to this task with less stress. One possible explanation is that the repeated sounds and task structure primed Rachel for the task, reducing her stress response. However, Rondal (1985) suggests that people with DS narrow their focus to specific aspects of speech processing as task demands increase (in Comblain and Rondal, 1996). Therefore, a second explanation is that when processing longer words, Rachel switched her attention from the whole word to processing just the prosodic element. This would explain the decline in segmental accuracy (Table 6.3) in words of three syllables, despite the opportunity to practise the initial syllables. As a result, her cognitive load and associated stress may have been reduced, resulting in improved voice quality. This is comparable to the improvements noted in singing performance when processing demands are reduced

(Berkowska and Dalla Bella, 2009a; Dalla Bella et al., 2012; Racette and Peretz, 2007; Welch, 2016; Wise and Sloboda, 2008).

Phonological processing and cognitive load

With practice, Rachel increased the rate of her production of the words *Magic Book*, but this was at the expense of segmental precision (Table 6.19). Rachel's speech production (Table 6.1) was below average for DS subjects in DEAP (Kent and Vorperian, 2013). Although the presence of stuttering and cluttering indicate motor-deficits in speech tasks, her poorer reproduction of speech sounds in the CNREP task indicates a specific weakness in phonological WM (Gathercole, 2006; Laws, 1998). This will affect the quality of the representation of unfamiliar words in long-term memory, which will also affect retrieval and the quality of production (Gathercole, 2006; Jarrold, Baddeley, Hewes, Leeke, and Phillips, 2004). Poor quality representation of words may explain Rachel's inconsistent productions of repeating words in *Happy Birthday* (Figure 6.17). It is therefore likely that her production of 'nonsense' sounds in her final rendition of *Magic Book* (Figure 6.11) reflects poor storage of the words and speech sounds. In contrast, her production shows that she had stored the melodic and rhythmic properties of the words. This is comparable to TD learners when they learn a new song. Racette and Peretz (2007) found that when their subjects omitted words they substituted either nonsense syllables that preserved the rhythmic structure or produced alternative words that were semantically close. However, Rachel's substitution of the word *why* for *what's* in phrase 3 of *Magic Book*, and her production of *me* in place of y*ou* in HB indicate that Rachel did not process the words semantically. Recall is improved when DS subjects process semantically (Purser and Jarrold, 2005; Smith and Jarrold, 2014) so Rachel's failure to do so would have further hampered her production.

The data indicate that Rachel's production of melody is a key strength in her singing, in her imitation of short words, and in her spontaneous speech production (a predominantly right-hemisphere task). Her poorer imitation of melodic contour in CNREP (Table 6.14, Table 6.15) may indicate the additional demands of processing speech segments, as discussed above. Rachel's superior production of prosodic features is consistent with research in DS that indicates a preference for global processing (Porter and Coltheart, 2006), and with Rondal's suggestion that individuals may narrow their attentional focus (in Comblain and Rondal, 1996). However, a particular feature of processing words in music may have influenced Rachel's production in songs, and has potential implications for her speech. In song, consonants may be processed separately to vowels, which are strongly connected to the melody (Kolinsky, Lidji, Peretz, Besson and Morais, 2009; Lidji, Jolicœur, Kolinsky, Moreau, Connolly and Peretz, 2010). If consonants are processed and stored separately from vowels in songs, an increase in segmental processing will increase cognitive load, leading to reduced overall performance. This effect of dual processing for words

may account for Rachel's poorer production of consonants in her final production of *Magic Book,* but it may also account for her more accurate production of vowels (Figure 6.11).

6.6.2 Happy Birthday: evidence of transfer between song and speech

Rachel's sung performance of *Happy Birthday* (Figure 6.19) was qualitatively different to her spoken production (Figure 6.18). Her sung production was more melodic, and was produced with elongated vowels and a slower rate of production (Table 6.21). These features make it more song-like than speech-like (Elmer, 2011; Wiesnewski, Mantell and Pfordresher, 2013). Rachel paused between phrases in her spoken production and became confused in her use of pronouns (Figure 6.17). In contrast, her recall of the lyrics was accurate in the song and uninterrupted. The argument that learning words as part of a song facilitates word recall is contested in research (e.g. Racette and Peretz, 2007; Gordon et al., 2010). Where research does indicate that singing benefits recall, the effects are attributed to a reduction to memory load by rhythmic 'chunking' (Schlaug et al., 2008), reduced presentation rate (Kilgour, Jakobson and Cuddy, 2000), heightened perception of rhythm (François et al., 2013; Hawkins, 2014) and better encoding during learning (Thaut et al., 2008). Thaut et al. (2008) found that adults with MS and memory impairments were better able to recall a list of words when the stimulus and the response were sung. However, Racette and Peretz (2007) found that word recall in healthy adults deteriorated if a text was presented as a song. They argue that the musical demands may interfere with word recall, but they also discuss the additional processing required in order to extract the text of a song for speech (Racette and Peretz, 2007). Rachel did isolate the words from the melody but the process of transfer may have placed additional cognitive demands which interfered with her recall of the words. The data support the notion that for Rachel, the words and melody of *Happy Birthday* were more easily recalled as one unit.

6.7 Summary and implications

As outlined in Section 6.1, Rachel's verbal ability is lower than her nonverbal ability. The results from music tasks show that her musical abilities lie within these measures: her rhythmic abilities are comparable to those of a TD child of between 5 and 9 years (Gooding and Standley, 2011; Hargreaves, 1996). Rachel uses a wide vocal range and exhibits characteristics of an initial range singer (stage 4 of 5: Rutowski, 1997). In speech and song she is able to reproduce the 'coarse-grained' features, but she can reproduce intervals in song with directed teaching, and can produce most speech sounds in some situations. Rachel's lack of accuracy in reproducing intervals, especially in the context of singing a song, places her at phase 2 of 4, in terms of singing development (Welch, 1998). However, the data indicate that it is primarily cognitive factors that affect her learning and production, and that cognitive load in turn affects her behavioural and emotional state, and therefore her performance.

Rachel's strengths in both domains are melodic, which may reflect an attentional bias for global processing in people with DS (Porter and Colheart, 2006). The data do suggest a perceptual deficit for fast temporal information in the speech domain. In particular, Rachel may have difficulty with processing or producing speech segments in situations that require multiple processing: e.g. speech sounds *and* rhythm *and* melody *and* meaning. As the processing demands of a task increase, Rachel may focus her attention on prosodic processing in order to reduce the cognitive load. Whilst research indicates that singing is more demanding than speech as a result of increased processing (Christina and Reiterer, 2013; Racette and Peretz, 2007; Gordon et al., 2010; Zatorre and Baum, 2012), tasks that focus upon separate elements of the task may support Rachel's initial learning. For example, Rachel could develop her knowledge of the melodic shape (melody and rhythm) on a simple vowel sound, which would reduce cognitive load (Berkowska and Dalla Bella, 2009a; Dalla Bella et al., 2012) and improve vocal accuracy (Dalla Bella et al., 2012; Welch, 2016). Rachel could then listen to the words presented in song: as speech production is slower in song than speech, this can support processing of speech segments (Schlaug et al., 2008; Schön, Boyer, Moreno, Besson Peretz and Kolinsky, 2008; Wan et al., 2010; Welch, 1998). Deconstructing the task would reduce the cognitive demands associated with first learning the words. Once the global elements are in place, Rachel may more easily attend to phonetic details.

Thus, singing may support Rachel's processing of speech but the data from *Happy Birthday* indicate that subsequent transfer to speech may be more challenging. Rachel's ability to imitate and learn words in speech and song is affected by her poor auditory STM. With repeated practice her recall of the text of *Magic Book* improved*,* in line with TD children (Davidson et al., 1981; Davidson, 1985) but the variability in her production of speech sounds in words suggests that her phonological representation of words is poor. This may be the result of reduced perception or hearing, but it may also be linked to poor retention of the words in memory. She had difficulty in recalling the words to *Happy Birthday*, and the accuracy of her speech in *Magic Book* deteriorated with practice (Table 6.19). People with DS can develop stable representations with sufficient exposure (Gathercole, 2006). However, Rachel would therefore need encouragement to practice speech segments, words and motor tasks slowly and accurately during the learning process. Accurate and 'error-free' rehearsal can reduce the effort associated with whole-word production in people with AOS (Whiteside, Inglis, Dyson, Roper, Harbottle, Ryder, Cowell and Varley, 2012). A focus upon the accurate rehearsal of words when learning may support automatic production of words (Whiteside et al., 2012). For Rachel, a similar strategy could reduce the cognitive load for speech production. Futhermore, Rachel's memory of words in songs may be supported by teaching her the meaning of the words. Her poor recall of the words in *Magic Book* and *Happy Birthday* indicate that she did not process the words at a meaningful level. According to Smith and

Jarrold (2014), people with DS code speech input phonologically and their interpretation of words is affected by other lexical representations. Alongside phonological tuition, the authors conclude that specific tuition in the meaning of words may support semantic coding and to support recall (Smith and Jarrold, 2014). Therefore, Rachel's production may be assisted if she received specific tuition in the meaning of words and phrases in songs.

In the long term, repetition and the teaching of specific aspects of the task in isolation will help to reduce cognitive processing demands and support retention. However, Rachel's own perception of the difficulty of the task, or of her own ability to succeed first time seemed to affect her willingness to participate and her voice quality. If Rachel were to benefit from learning words or sounds in a musical context, a range of strategies will be needed to support her confidence, as well as to reduce task demands. According to Patel's OPERA hypothesis (2011; 2014), transfer of learning will only take place if the musical activity places more demands on attentional and cognitive resources than are required for speech. Singing is inherently more demanding than speech (Christiner and Reiterer, 2013; Racette and Peretz, 2007) but careful management of the teaching would be necessary in order to ensure that Rachel was not stressed and that she remained focussed, without reducing the 'precision' of learning required.

Chapter 7: Robert

The first section of the chapter will describe Robert's background, based on his college records and his performance in core assessments (Chapter 3, Section 3.2). Section 7.2 will present the results of Robert's performance in music-based motor timing tasks, including beat and rhythm entrainment, and in syllable segmentation and rhythm imitation tasks. Section 7.3 will present data from voice and speech tasks. This section will also present the development of his learning in the songs *Magic Book* and *Monkey Man*. Section 7.4 will compare Robert's vocal qualities and prosodic abilities in continuous speech and when singing and speaking previously learned text (*Happy Birthday*).

The results of these tasks will be discussed in Sections 7.6-7.8. However, Section 7.5 will first consider the extent to which his hearing loss affects his perception in speech and music tasks. Section 7.6 will then examine Robert's temporal production in motor-music, speech and music tasks and his ability to reproduce pitches, melodic intervals and melodies. Section 7.7 will discuss Robert's learning and development in songs. It will consider the difference between his own, previous learning of the song *Happy Birthday*, and the potential to support his learning and to support transfer between spoken and sung domains. Section 7.8 will discuss the implications of using music and singing activities as a means of supporting Robert's ability to produce speech.

7.1 Summary of history, learning and performance in core assessments

Robert is 23;10 years old, and described as having SLD and moderate hearing loss. Within the adult DS population, hearing loss is believed to affect a third and to worsen with age (Määttä et al., 2011). The college record states that Robert's hearing appears to have deteriorated since his enrolment, three years prior to the study. At the time of his enrolment, his hearing was 'thought to be adequate', but a recent audiogram by the college SaLT shows that he has moderate bilateral hearing loss. The audiogram of his right ear shows a loss of between 50-60 dB between 250-1000 Hz, and at 4000 Hz; there is a slightly reduced loss of 46 dB at 2000 Hz. The loss in his left ear is 40 dB at 250 Hz but is between 60 dB and 65 dB for the range 500-4000 Hz. Robert does not wear a hearing aid: the SaLT report recommends the use of pictures or Makaton to support his speech comprehension and expression.

According to his SaLT, Robert has poor speech intelligibility and limited oral movement for speech. Specifically, his records state that he has a limited lingual movement, a reduced ability to move his tongue upwards, and incomplete closure of the lips and palate. During his time at college, speech and language therapy sessions focussed on: his use of lips to produce sounds; the production of

final sounds in words; his use of 'forward breath' for the sounds /f/, /w/, /p/ and /b/; and on social communication skills. His language understanding is described by the SaLT as limited to key words, and his output is also limited to key words, supported with Makaton signs. He is described by the SaLT as socially motivated, but lacking in confidence; and his voice is quiet, breathy and pubophonic. He has over three years' experience in taking part in singing and drama activities whilst at college. The extent of SaLT therapy or education prior to his enrolment at college is not known.

7.1.1 Cognitive and perceptual abilities

Robert's nonverbal ability of 4;3 years as measured by DAP is higher than his verbal ability (Table 7.1). His TAPS score of two (Table 7.1) is lower still. Robert was unable to complete the pre-screening tasks for same-different discrimination tasks in speech (Table 7.1: tests 7 and 14) or in music (Table 7.1: tests 17, 21 and 26). As a result, he did not attempt the PMMA task. In the minimal pairs based on different consonants (Table 7.1: test 13a), he failed to discriminate between /bat/ and /rat/, and between /mat/ and /bat/. In vowel-based pairs (Table 7.1: test 13b), he failed to discriminate between /cart/ and /cot/.

Robert was unable to read and he did not name all pictures correctly in the DEAP test (Table 7.1: test 5): for example, he labelled 'thumb' as *foot* and 'girl' as *boy*. However, in naming 'fish', his use of the correct Makaton sign showed his knowledge but he produced the sound as: [sɪsɪ]. When the researcher responded with *yes, fish*, Robert imitated the consonant sounds accurately, producing [fiʃ]. It is not clear whether Robert's responses in picture-naming reflected an inability to identify the pictures or a difficulty in producing the sounds.

7.1.2 Motor-music abilities

Observations of Robert's gross-motor movements to music (see Appendix 2, p. 324) indicate that he had difficulties in coordinating movements to the beat. He was given visual support throughout the task to model the desired movements and the timing of these to the dominant hierarchical beat, but his movements were typically slower than the demonstration and lagged behind the beat. Robert seemed to have difficulty in alternating leg movements and moved just his left leg in response to the beat. When asked to walk in time to the music, he moved slowly and watched his own feet as he did so. He was also unable to alternate his arms when asked to lift one at a time from his side: instead he moved his upper torso in the indicated direction. In imitating claps (Table 7.1: test 4), Robert produced all rhythms with just two claps.

Table 7.1: Profile: a summary of Robert's performance in core tasks. The table gives the results of the core assessments in nonverbal and verbal ability, speech perception and production and musical perception and production. Test numbers refer to the test descriptions in Chapter 2, Section 3.3.

	Test number	Test name	Results
Cognition	**1**	BPVS RAW score	25
	1	BPVS Age equivalent	3;0 (2;8 - 3;5)
	2	Draw–a-person	4;3
	3	TAPS - Digit Span	2
Perception (speech)	**7a**	Same-different discrimination non-words (Stackhouse and Wells App C.2)	NULL
	7b	Same-different discrimination real-words (Stackhouse and Wells App C.2)	NULL
	13a	Minimal pairs perception: discrimination of consonants	3/5 = 60%
	13b	Minimal pairs perception: discrimination of vowels	4/5 = 80%
	14	Simple speech prosody test	NULL
Speech Production	**5**	DEAP: Percentage of Phonemes Correct (PPC)	41%
	8	S/Z	n/a
	9	DDK	n/a
Perception (music)	**17**	PMMA	n/a
	26	Pitch perception test	NULL
	27	Interval and rhythm perception test	NULL
Music Production	**18**	Gross Motor score	13/40 = 33%
	4	Rhythm Imitation (TraCOL)	7.1 % (n=1)
	22a	Pitch matching on /la/	0/5 =0 %
	22b	Pitch matching: Sustained /i/, /u/, and /a/ vowels at three pitches	0/11 = 0%
	23a	Pitch direction: vocal glides on /la/	2/4= 50%
	23b	Interval matching	0/5 = 0%
Reading	**15**	Swimming passage	n/a

7.1.3 Oro-motor and phonological abilities

Robert was unable to produce the sound sequences for DDK or the s/z tasks (Table 7.1: tests 8 and 9). This is not unexpected, given the reported difficulties he has in lip closure, or his reported level of hearing loss. A breakdown of Robert's speech errors in the DEAP task is given in Table 7.2. The majority of Robert's errors were cluster reduction and final consonant deletion. For example, he omitted the final consonant of /bɹɛd/, and substituted the voiced alveolar approximant /ɹ/ with a

glide /w/, producing the word as [bweɪ]. He has a limited phonological repertoire that consists mostly of bilabial, labiodental and alveolar sounds. For example, he produced /swing/ as [fi], /watch/ as [fɔʃ], /feather/ as [pəbə]; and /kitchen/ as [ʊʃɪʃ].

Table 7.2: Speech errors and error patterns produced by Robert in the DEAP task. The table shows the number of errors made in the task and the error patterns (n=5 or more; except WSD* n=3), expressed as a percentage of the total number of errors made. The errors are coded according to the conventions outlined in the DEAP manual, which defines the following categories: D indicates delayed error patterns; U indicates unusual error patterns.

DEAP		
Percentage of error pattern/total errors		
Gliding	**D**	5.2
Cluster reduction	**D**	15.7
Final consonant deletion	**D**	17.8
Weak syllable deletion*	**D**	8.4
Initial consonant deletion	**U**	6.3
Medial consonant deletion	**U**	7.3
Other	**U**	17.8; E.g. initial and final sounds substituted with [b] and [ʃ]
Number of errors		**95**
Percentage of errors accounted for by error patterns		**78.5**

Robert's accuracy in producing speech sounds in different tasks is given in Table 7.3. He omitted a high proportion of sounds in the naming task (39.86%) and when imitating words of three syllables (38.73%). He made a high number of consonant errors (PCE-R) in imitating words and the percentage of consonant errors increased in line with the target syllable length.

Given the high number of errors made in the Syllables task as the target number of syllables increased, Robert was not asked to imitate phrases or non-words.

Table 7.3: Percentage of speech errors produced by Robert in core speech tasks. The table shows the percentage of errors in phonemes (PPE), consonants with habitual errors (glides) removed (PCE-R), vowel errors (PVE) and the percentage of sounds omitted (PEO). The number of words examined in each test and subtest (Syllables) is given.

Test no.	Test name	No. of words, n=	PPE	PCE-R	PVE	PEO
5	**DEAP**	**49**	59.22	68.12	40.58	39.86
11	**Syllables**	**27**	76.43	86.78	58.02	38.73
	1-syllable words	**9**	65.52	80.00	33.33	30.00
	2-syllable words	**9**	84.09	88.46	77.78	34.62
	3-syllable words	**9**	79.69	91.89	62.96	51.56

7.1.4 Voice and phonation

Robert was able to produce a range of 16 semitones when producing pitch glides on a /la/ sound (Table 7.1: test 22a). His range in speaking tasks was considerably lower (Table 7.4). His mean F0 was higher in spontaneous speech than in picture description, and both means were close to his minimum pitch range when producing glides.

Table 7.4: Mean pitch and range of Robert's vocal glides on 'la', and in connected speech. For each task, the table shows the mean, minimum and maximum values of fundamental frequency in Hertz, and the SD of F0 when speaking, in Hz and semitones (ST). The mean change in speaking pitch is also shown in ST. Mean F0 and SD of F0 is not given for glides, whose purpose was to examine vocal range.

Test No.	Test name	Mean F0, Hz	Min F0, Hz	Max F0, Hz	SD of F0, Hz	Pitch range, semitones
22a	**Pitch glides on /la/**		163.09	421.82		16.45
10	**Picture description**	172.93	158.09	197.37	9.75; 0.97	3.84
n/a	**Spontaneous speech, comment 2**	178.09	154.59	20.24	11.20; 1.09	2.45

In the medial section of sustained vowels, the mean HNR values were within typical range but shimmer (%) was elevated (Table 7.5). Eight of the eleven vowels were produced with interharmonics, but the onset of seven vowels overlapped with the demonstration, and it is not clear whether these also contained interharmonics. Eight vowels were produced with interharmonics in the medial portion and diplophonia was audible in three of these. Robert did not reproduce the target vowels, however, or the target pitches. Further details of his sustained vowels will be given in Section 7.3. The number of vowels in words that were produced with interharmonics was between 63% and 70%.

Table 7.5: The mean voice perturbation measures of Robert's sustained vowels and vowels in words. For sustained vowels, mean measures are given and separate values are given for the initial 100 ms (onset), medial duration (Total duration - (onset+offset), and the final 100 ms (offset). Mean MPT for sustained vowels was 1.5 seconds. Measures are given for the vowels in each syllable of words produced in DEAP, and Syllables tasks. The table shows the mean data for F0 (Hz), DUV (%), HNR (dB), jitter (%), shimmer (%) and CPP (dB). The table also shows the number of sustained vowels and vowels in the Syllables task that were produced with interharmonics. No data for interharmonics was generated for the DEAP task as the recording quality was low: Robert's voice was quiet on the recording and the DEAP data was also limited to 10 vowels that were not contaminated by extraneous noise.

Test no:	Test name	No. of valid vowels, n=	DVB, %	DUV, %	HNR, dB	Jitter, %	Shimmer, %	CPP, dB	No. and percentage of vowels produced with interharmonics
19	**Sustained vowels: total vowel duration (T)**	**11**	0.00	1.56	22.37	0.82	3.85	5.14	8/8= 100%
	Onset (t=100 ms)	*4*	*0.00*	*21.58*	*9.61*	*4.26*	*14.75*	*3.66*	*4/4 = 100%*
	Medial duration (T- (onset + offset))	*11*	*0.00*	*0.00*	*22.89*	*0.74*	*3.61*	*5.41*	*8/11 = 73%*
	Offset (t=100ms)	*11*	*0.00*	*9.72*	*14.95*	*1.93*	*15.54*	*4.50*	*0/11= 0%*
5	**DEAP**	**10**	0.00	6.64	12.61	0.71	13.39	4.48	7/10= 70%
11	**Syllables**	**48**	0.00	1.56	16.33	1.22	8.01	5.28	30/48 = 63%

7.1.5 Learning

During the baseline assessments Robert sang quietly in his upper register, with little melodic variation and no clear articulation of consonant sounds (see Appendix 3, p. 343 [17]). As a result, he was set the following learning goals, in addition to session goals (Appendix 3, p. 340):

- sustain phonation for 3 seconds using kazoo at any pitch;
- explore lower range of voice using kazoo with visual assistance; and
- articulate /w/ and vowels in the phrase *We Will Rock You* (May, 1977).

In the second session, Robert sustained phonation using a kazoo, and produced the /w/ consonant in the chorus of *We Will Rock You* with Makaton cues to prompt this (Appendix 3, 344 [36-41]). He began to produce greater contrast between vowels in the same song in Session 3 (Appendix 3, p. 344 [40]). In Session 3, he was able to produce vocal glides in his lower register using a kazoo, but did not use a lower register in any other context.

Section 7.3 will examine Robert's learning in more detail, based on the recordings of his initial and final assessments and on the recordings of his performance in specific songs.

7.2 Motor-music entrainment and imitation

7.2.1 Beat entrainment

Robert adapted his tempo (ITI, ms) in response to changes in the tempi of beat entrainment tasks (Table 7.6). The duration between his beats was variable (SD of ITI, %). The pace and regularity of his movements (ITI, ms, and SD of ITI, ms respectively) were most accurate at 160 bpm. Robert was unable to synchronise to the beat (AT, ms) and his timing was inconsistent (SD of AT, ms). On average, his timing to the target beat was delayed by 36% (mean AT, %). His timing to the beat (AT, %) was most accurate and least variable (SD of AT, ms) at 160 bpm, and at this tempo, he played between the target beats, close to anti-phase (AT, %).

7.2.2 Rhythm entrainment

Robert's ITI was variable in rhythm entrainment tasks and did not align to the target beats. The data are given in the bar charts in Figure 7.1 and 7.2. These show a difference in Robert's timing between beats when the stimulus was aural only (Figure 7.1) and when the aural stimulus was accompanied with a visual demonstration of the rhythm (Figure 7.2). Although the two rhythms are different in pattern, both patterns consisted of simple subdivisions (the quarter note, *ta*, and eighth-notes, *tee-tee*) and the tempo remained constant at 104 bpm. Robert's ITI varied between 481 ms and 637 ms when he performed the first rhythm to an aural stimulus (Figure 7.1). His range was 156 ms; half that of the stimulus (265-585 ms). The duration between one of his beats was close to

Table 7.6: The accuracy of Robert's timing in beat entrainment tasks. The table shows mean ITI and AT in milliseconds at 104 bpm, 132 bpm, and 160 bpm. The data are also expressed as percentage of the target interval, and the SD in milliseconds.

Tempo, bpm	**104**	**132**	**160**
Interval time, ms	**576**	**455**	**375**
No. of taps produced	**20**	**25**	**33**
ITI, ms	799	583	390
ITI, %	72	78	96
SD of ITI, ms	145	36	57
Mean range of ITI, %	16		
AT, ms	+184	+108	+199
AT, %	32	24	53
SD of AT, ms	191	159	73
Mean AT, %	36		

the target ITI (Figure 7.1: beat 10). The timing between his taps to the aural stimulus was less variable than the target rhythm: however, when Robert was instructed to copy a rhythmic movement whilst listening to the stimulus, he produced a wider range of ITI (255-854 ms), in excess of the stimulus (261-561 ms) (Figure 7.2). With ongoing visuo-spatial support, five of

Figure 7.1: Graph showing the target ITI and the ITI of Robert's production in part 1 of the Rhythm entrainment task - ta tee-tee ta ta. The chart shows the duration between beats during the final two bars of Rhythm 1 of the entrainment task, without visual support. The blue bars show the target duration, which consists of a pattern of ten beats in a 2:1 ratio. The purple bars represent the timing between Robert's beats. The green and blue lines show the 'moving average' of Robert's production and of the target stimulus, respectively.

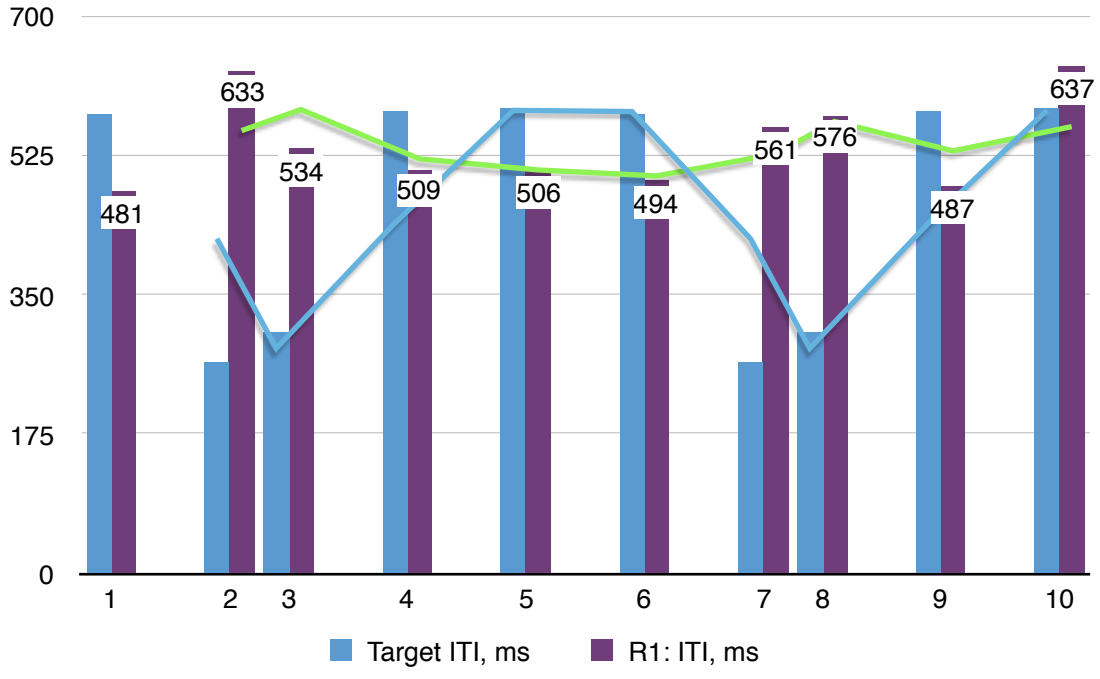
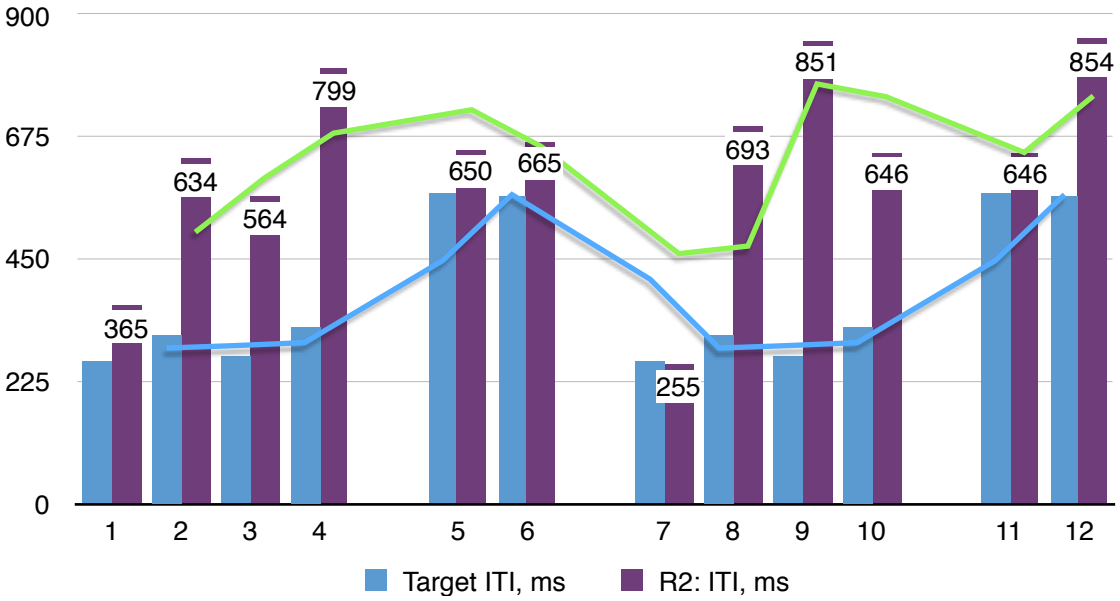


Figure 7.2: Graph showing the target ITI and the ITI of Robert's production in part 2 of the Rhythm entrainment task - tee-tee tee-tee ta ta, with visual support. The blue bars show the target duration. The purple bars represent the timing between Robert's twelve beats. There is variation in the duration between Robert's beats, and these are close to the target at beats 1, 4, 6, 7 and 11. The green and blue lines show the 'moving average' of Robert's production and of the target stimulus, respectively.



Robert's beats were close to the target ITI (Figure 7.2: beats 1, 5, 6, 7 and 11). The moving average of ITI was generated on each graph in 'Pages': in comparison to the aural-only production, Robert's moving average (in green) is more characteristic of the demonstration (in blue) in the rhythm that included visual demonstration (Figure 7.2).

7.2.3 Syllable segmentation and rhythm imitation

Robert was unable to reproduce clapped rhythms in the TRaCOL task (Table 7.1: test 4) or to spontaneously segment and clap word rhythms, with the exception of *cat*. However, he imitated word-based rhythms of up to three claps (Table 7.7). Only his imitation of the word *monkey* was produced with timing between claps that matched the target syllable duration when spoken.

Table 7.7: The number and relative duration of claps produced by Robert in syllable segmentation tasks. The table shows Robert's accuracy in producing the appropriate number of taps to represent the number of syllables in words, and his ability to produce these with appropriate relative duration between taps. All tasks were modelled for Robert, except for 'cat'.

Words	Target relative duration of claps	Spoken syllables reproduced	No. of claps/ taps produced	Relative duration of claps
Can you	**1:1**	0	1	n/a
Can you clap	**1:1**	0	3	2:3
Can you clap the beat	**1:1**	0	3	2:4:3
Monkey	**1:1**	2	2	1:1
Cat	**not demonstrated**	1	1	n/a
Caterpillar	**1:1**	3	1	n/a

7.3 Voice and prosody in imitation tasks

7.3.1 Sustained vowels

Data for all sustained vowels produced by Robert are shown in Table 7.8. The target pitch of the demonstrated vowels increased (196 Hz (G3); 261 Hz (C3); 392 Hz (G4)) but Robert's productions were all close to C3 (261 Hz), except for /a/1. DVB was 0.00% for all vowels and is not shown in the table. The SD of F0 was high in two vowels (/i/3 and /a/2), and the vowel /u/3 was produced with a high DUV. HNR was above 20 dB in all vowels except /a/1. Eight vowels were produced with interharmonics in either medial or initial position, and three (/u/4, /a/1, /a/2) were produced with

Table 7.8: The mean fundamental frequency and voice perturbation measures for Robert's sustained vowels. The table shows mean data (F0, DUV, HNR, jitter, shimmer, CPP) for nine sustained vowels, including onset and offset, but excluding any sections that were contaminated by extraneous noise. Robert's production of seven vowels overlapped with the demonstration and the contaminated portion ('Duration of overlap', ms: shown in red text) is excluded from these measurements. Robert did not produce the target /u/ vowel in his first two attempts; two additional vowels (/u/4 and /u/5) were included as he was asked to repeat the target vowels. The presence and duration (milliseconds) of single or complex (2+) interharmonics is shown, and their position in the vowel is indicated, as follows: onset (i; t=100 ms), medial (m: the period after onset, but excluding the final 100 ms), and final (t=100 ms). Where interharmonics are medial but are continuous with onset or offset, this is indicated with +. However, as a result of the overlap between the demonstration and Robert's production, it is not known if Robert produced interharmonics at or after onset in five vowels: these are indicated on the table as 'n/k - contaminated'. Bold type indicates values that lie within norms for the age range, using Sonneta (Goy et al., 2013); and vowels produced with audible diplophonia.

Target vowel (**production)	Duration of overlap/ Duration of vowel, ms	Mean F0, Hz	SD of F0, Hz	DUV, %	HNR, dB	jitter, %	Shimmer, %	CPP, dB	1 InterH: Duration (ms) and position	2+ InterH: Duration (ms) and position
/i/1*	236/**2250**	242.12	3.72	0.00	**24.93**	2.06	4.63	5.61	59 ms, 73 ms (m)	n/k- onset contaminated
/i/2	**1086**	298.08	2.89	1.91	**23.05**	0.72	4.07	4.35	0.00	100 ms (i) +2 ms (m)
/i/3	152/**1135**	277.96	0.68	0.00	25.33	0.54	3.17	5.16	0.00	n/k- onset contaminated
/u/1 (/i/)**	604/**1088**	238.76	0.92	0.51	**25.09**	0.47	5.05	5.11	0.00	n/k- onset contaminated
/u/2 (/a/)**	**1272**	257.95	5.41	3.61	20.48	0.59	4.09	4.88	0.00	100 ms (i)+ 252 ms (m)
/u/3	**1005**	271.99	3.08	6.22	**24.12**	0.64	**2.55**	5.30	0.00	100 ms (i) +126 ms (m)
/u/4	75/**2276**	250.69	13.37	0.76	21.09	2.14	4.58	5.47	**381 ms, 82 ms (m)**	118 ms (m)
/u/5	**1121**	279.11	2.79	2.67	24.12	0.79	3.46	4.74	0.00	100 ms (i) +3 ms (m)
/a/1	103/**2190**	177.91	1.94	0.19	18.87	1.04	4.42	5.97	**116 ms, 126 ms,** 77 ms (m)	107 ms (m)
/a/2	125/**1542**	252.70	12.35	0.82	20.47	0.50	4.82	5.00	**168 ms** (m)	n/k- onset contaminated
/a/3	331/**1228**	289.19	1.01	0.00	**23.21**	0.46	3.66	4.55	0.00	n/k- onset contaminated
Mean	**1472**		**4.38**	**1.52**	**22.80**	**0.90**	**4.05**	**5.10**		

audible diplophonia. Robert appeared to have difficulty in reproducing an /u/ vowel: the first two of these vowels were reproduced on an /a/, but the target sound is given in the table. Five of Robert's vowels were produced at or above the mean HNR values reported for young males (22.9 dB: Goy et al., 2013); all jitter and shimmer values were above these norms (<0.38% and 2.71%, respectively).

Details of two examples are given below. These were selected to illustrate two different types of vocal quality, and Robert's strategy in matching pitch (/a/2). The vowel /a/2 (Figure 7.3) was produced with minimal overlap with the demonstration and it illustrates one of the vowels that Robert produced with a wide pitch range: it was also produced with audible shimmer, and one of the highest shimmer values (Table 7.8). The vowel /a/1 (Figure 7.4, Figure 7.5) was produced with the lowest mean F0 and a high proportion of interharmonics, mid-vowel. It was also produced with a perceived tremolo and high jitter and shimmer values: however, the CPP value was also high (Table 7.8).

Example 1 - /a/2

Robert's imitation on second assessment of the note D4 on /a/ is shown in Figure 7.3. Although he did not match the pitch, he increased his pitch towards the target pitch (section 1, tier 2), and changed from an [ɪ] vowel to [a]. The waveform shows that his onset was harsh (section 2) but his

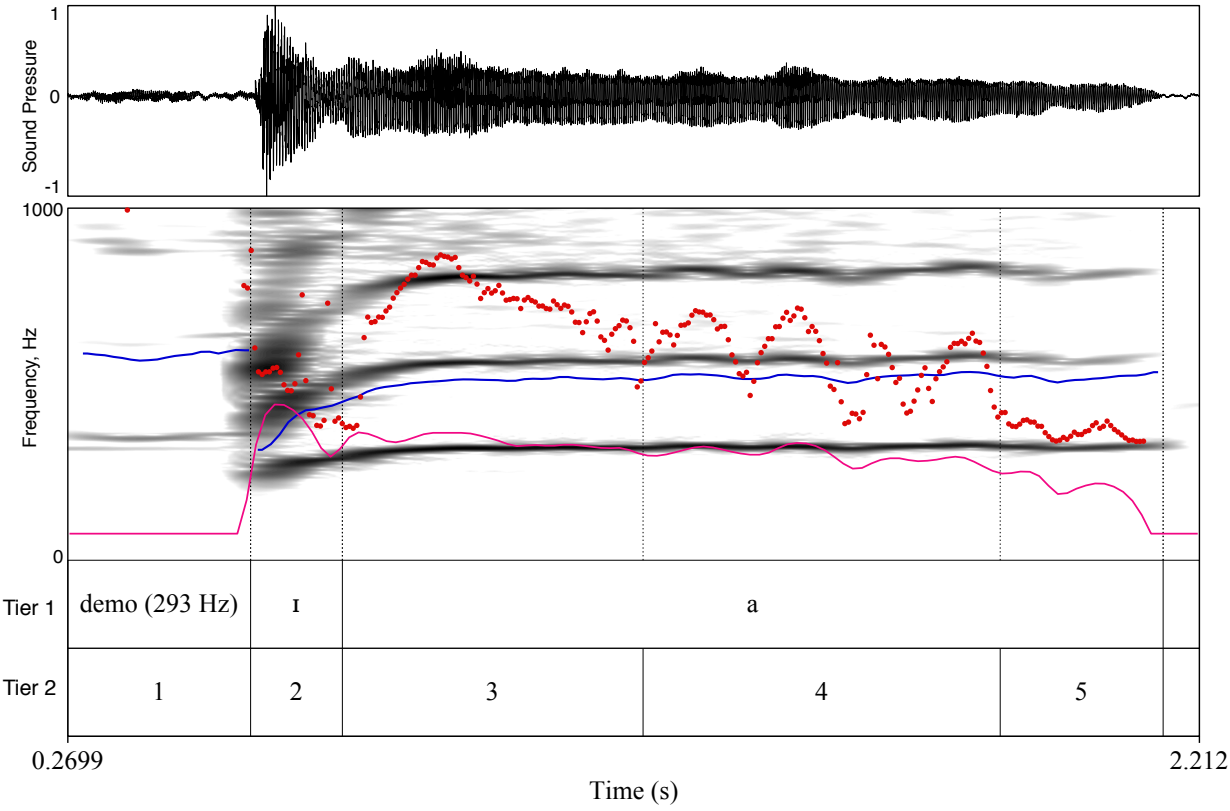


Figure 7.3: Waveform and annotated spectrogram showing Robert's sustained vowel /a/2 at a target pitch of D4. Robert's vowel was sustained at a mean F0 of 252.70 Hz. His production overlapped slightly with the demonstration (section 1) but the waveform shows that Robert's onset was harsh (section 2). The narrow-band spectrogram (0-1000 Hz; window length 150 ms, 40 dB) is overlaid with the pitch contour in blue (50-500 Hz), the intensity contour in pink (50-100 dB), and with the first formant frequency (red dots).

phonation was modal during the /a/ vowel. Robert sustained the pitch (in blue) as his intensity (in pink) declined, but the formant pattern (in red) indicates movement of the tongue body.

Example 2 - /a/1

Robert's /a/1 vowel (Figure 7.4) was produced just 16 Hz below target and his pitch remained steady in section 3, after onset (Figure 7.4). Two breaks in the pitch contour coincide with an overlap with the demonstration (section 2) and with diplophonic production (section 5). The vowel was produced with a gradual onset and decline in intensity, but the spectrogram shows additional interharmonics towards offset (Figure 7.4, sections 4-6).

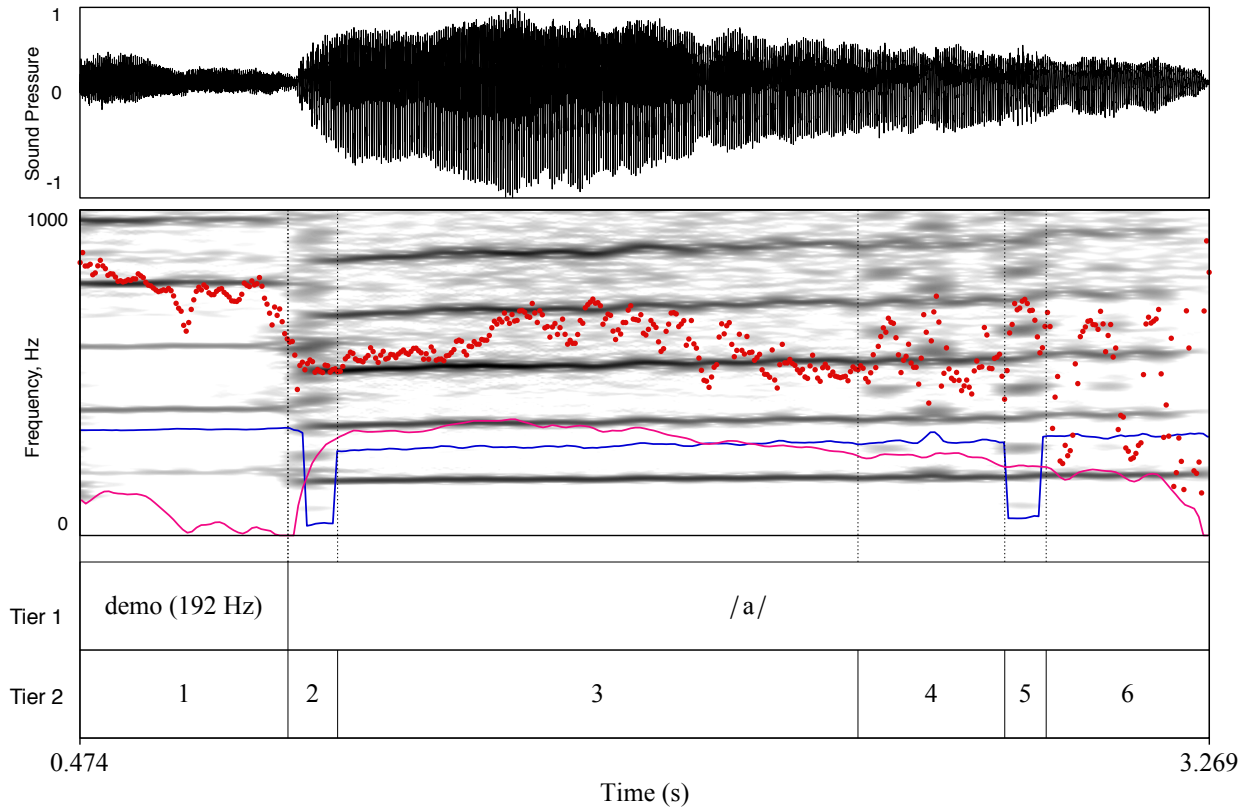
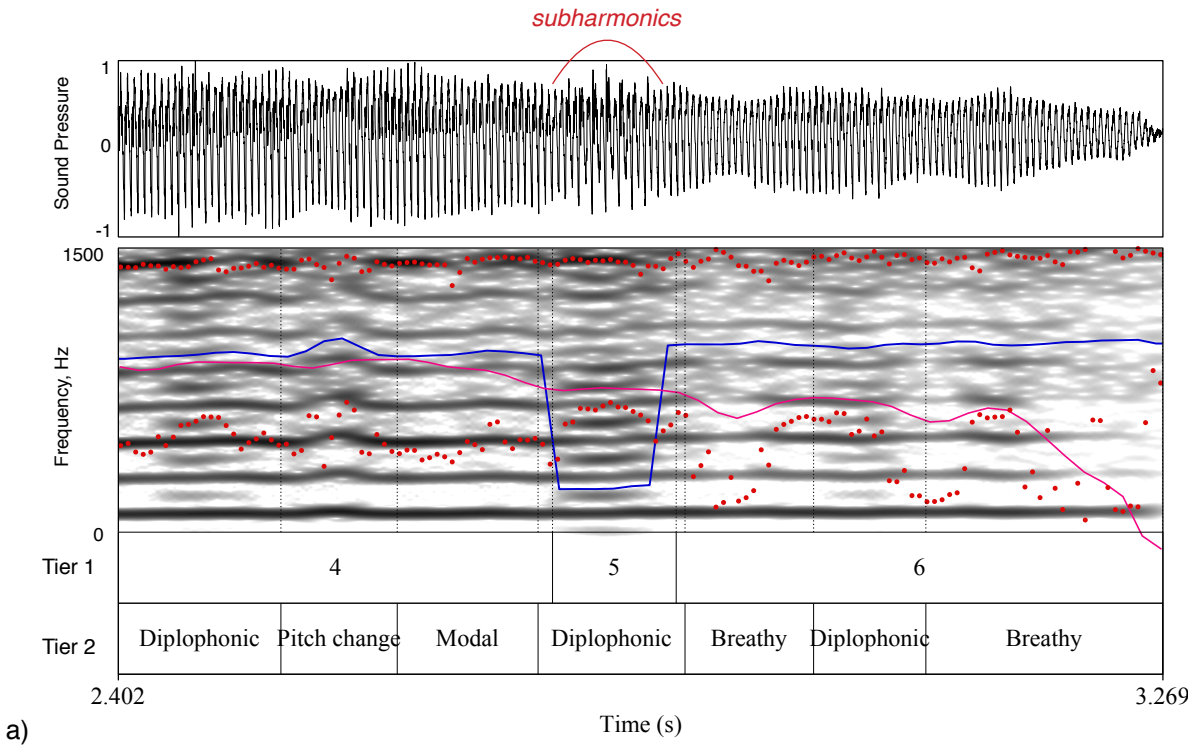
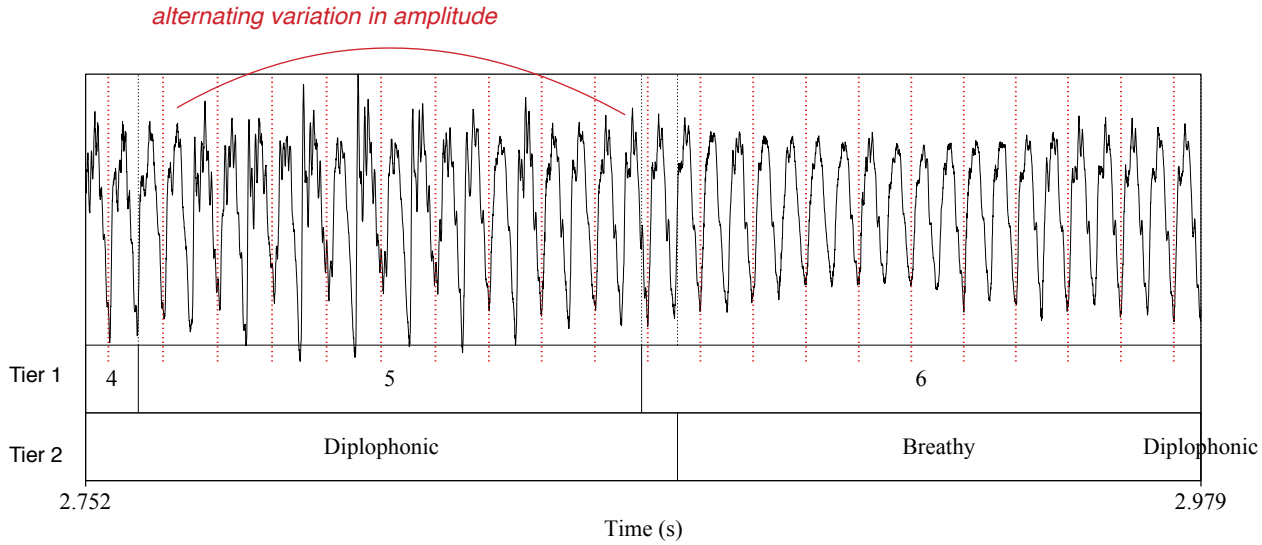


Figure 7.4: Waveform and annotated spectrogram showing Robert's sustained vowel /a/1 at G3. The image shows the waveform (top) and narrow-band spectrogram (0-1000 Hz; window length 150 ms; 40 dB). His production overlapped slightly with the demonstration (section 2). The blue contour indicates pitch (50-500 Hz), the pink shows intensity (50-100 dB) and the red dots show the first formant frequency. A subharmonic is visible on the spectrogram in sections 4, 5 and 6. Details of these sections are given in Figure 7.5.

Detail of the sections produced with interharmonics are shown in Figure 7.5, and the perceived vocal qualities are marked (Tier 2). Three sections were produced with audible diplophonia, each lasting approximately 125 ms in duration. The first formant frequency (in red) indicates instability in the tongue position during these sections. Detail of sections 4, 5 and 6 (Figure 7.5b) shows harmonics between the pulse lines on the waveform in both the diplophonic and breathy sections: the amplitude of the subharmonics are increased in the diplophonic section, and additional 'noise' is visible on the waveform (Figure 7.5 b: Tier 1, section 5).



a)



b)

Figure 7.5: Waveforms and annotated spectrogram showing a detail of Robert's sustained vowel /a/1. The image a) shows Sections 4,5 and 6 of the /a/ vowel (see Figure 7.4). The narrow-band spectrum is set to 1500 Hz in order to show the second formant (in red dots). For greater resolution, the window-length is 50 ms, and the pitch and intensity ranges are reduced (50-250 Hz; 55-70 dB). Perceived voice qualities are shown in Tier 2. The duration of each section that is associated with the different perceptual qualities is approximately 125 ms.

The image b) shows a detail of the waveform from section 5 and the start of section 6. There are two peaks visible between the pulse lines in all sections shown, but the quality of these are different in section 5, in which the relative amplitude alternates every two cycles.

7.3.2 Pitch and interval matching

Robert was unable to match absolute pitch when singing /la/ during either assessment (Figure 7.6). As a result of his pubophonic voice, the task was modelled in a high register, which may have affected his ability to accurately reproduce pitch. Despite this, during his second assessment he

produced a wider pitch range and he increased his pitch in line with the target pitch (Figure 7.6, R2). His imitation of notes G3 and D4 were 1.9 ST and 1.7 ST below target, respectively. His closest match during the initial assessment was 4.4 ST below target (G3). Robert was also unable to reproduce melodic intervals on /la/ (Figure 7.7). His imitation of the three-note pattern, *so-la-so*

Figure 7.6: Graph showing Robert's accuracy when matching pitches during initial and final assessments. The fundamental frequency of the target pitch (in Hz) is shown in blue and Robert's mean F0 is shown for productions in his first assessment (R1) and final assessment (R2), which were six weeks apart. The target pitch is shown in blue.

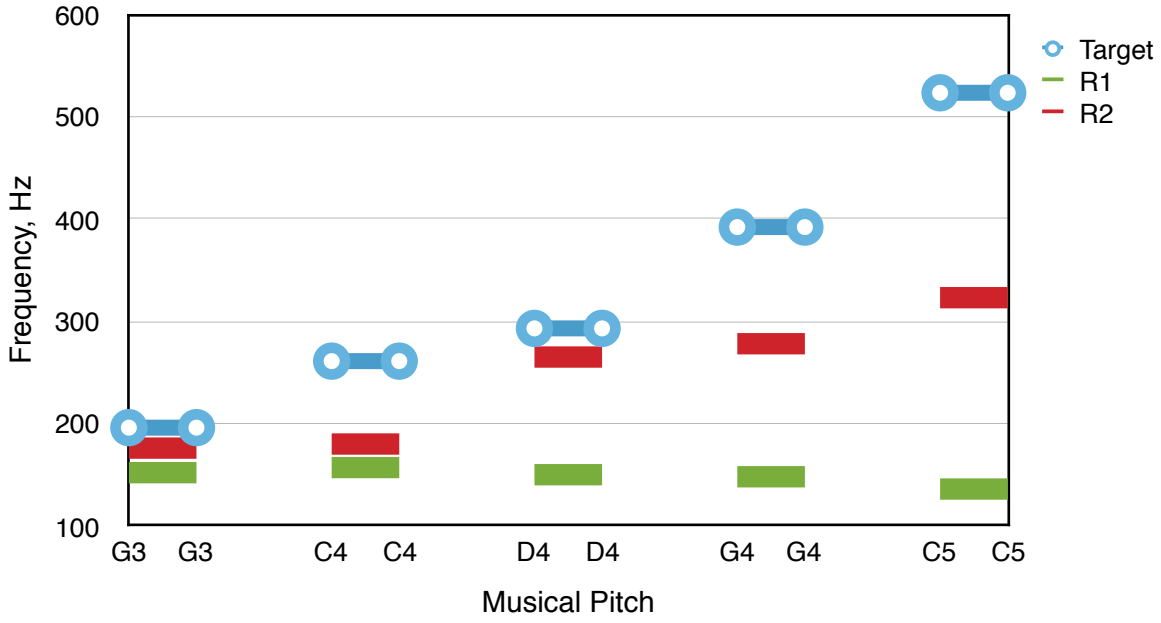
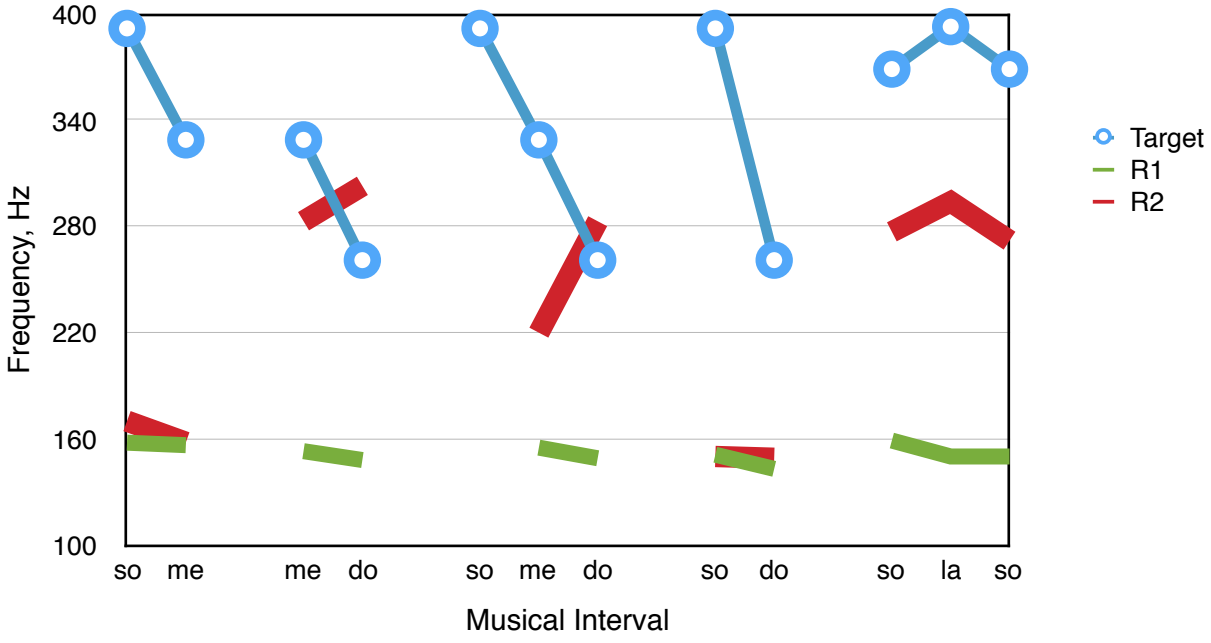


Figure 7.7: Graph showing Robert's accuracy when matching melodic intervals during initial and final assessments. The fundamental frequency of the target pitch (in Hz) is shown in blue and Robert's mean F0 is shown for productions in his first assessment (R1) and final assessment (R2), which were six weeks apart. The target pitch is shown in blue.



(Figure 7.7), improved on re-assessment, and he reproduced the interval to within 0.2 semitones of the demonstration.

7.3.3 Words of increasing syllable length

Table 7.9 shows the prosodic profile for words of one, two and three syllables as spoken by the researcher and as imitated by Robert. The relative contrast between successive vowel duration (nPVI-V) was close to the demonstration in words of one and two syllables. There was less contrast between vowel duration in words of three syllables, most of which were produced as two-syllable words. Robert's imitations were produced with limited pitch range or variation as measured by pitch range (ST), SD of ST, pitch trajectories and the percentage of rise or fall within vowels. However, Robert's mean pitch range did increase as the number of target syllables increased, and the mean rate of change within syllables (Traj inter- nucleus) also increased.

Table 7.9: The prosodic profile of words in the Syllables test, as demonstrated to Robert and as imitated by him. The table shows the mean values of nucleus duration, nPVI (nucleus and vowel) and pitch range (semitones), intra- and inter-syllable trajectory, rate of pitch change, and percentage of pitch rises and falls. The mean values are given for the researcher's production and for Robert's imitation of words of one, two and three syllables.

No. of syllables	No. of nuclei	nPVI: vowel	Pitch range, ST	SD of ST	Traj Intra-syllable	Traj Inter-nucleus	Rises, %	Falls, %
1: Robert	**9**	32	0.8	0.2	6.8	0.0	0.0	0.0
1: Demo	**9**	39	8.7	1.8	19.8	0.0	25.0	0.0
2: Robert	**18**	47	1.3	0.4	7.6	5.8	0.0	0.0
2: Demo	**16**	37	8.6	2.3	21.2	8.3	25.0	0.0
3: Robert	**22**	26	2.8	0.7	6.2	13.3	0.0	0.0
3: Demo	**27**	44	16.8	6.0	18.4	46.9	0.0	4.8
Percentage of demo		**87.5%**	**14.3%**	**12.8%**	**34.7%**	**34.6%**	**0.0%**	**0.0%**

Table 7.10 shows the key prosodic data for words of one, two and three syllables. The number of syllables he produced did not match the demonstration. He produced an additional syllable in the

Table 7.10: A summary of the segmental, prosodic and vocal features of Robert's imitation of words in the Syllables test, arranged by number of target syllables. The table shows the number of syllables Robert produced and the number and percentage of words he reproduced with the same pitch and intensity contours. It shows the number and percentage of the words affected by interharmonics. The table also shows the correct number of consonant and vowel sounds Robert produced.

No. of target syllables	No of words on the set	No. of syllables produced	Pitch changes/ direction reproduced	Changes in intensity contours reproduced	No. and percentage of consonants reproduced	No. and percentage of vowels correctly produced	No. and percentage of words produced with interharmonics
1	**9**	10/9 = 111%	n/a	n/a	4/20 = 34%	6/9 = 67%	5/9 = 56%
2	**9**	17/18 = 94%	8/9 = 89%	2/9 = 22%	3/26 = 11%	4/18 = 22%	6/9 = 67%
3	**9**	20/27 = 74%	6/9 = 67%	2/9 = 22%	3/37 = 8%	10/27 = 37%	7/9 = 78%
Mean percentage of target		**87%**	**78%**	**22%**	**14%**	**37%**	**67%**

word *thick* and produced *city* with one syllable: he produced all but two of the three-syllable words with just two syllables. He matched changes in the pitch direction of most words, but his accuracy reduced in words of three target syllables. He reproduced just 22% of the modelled changes in intensity between successive syllables. The number of consonants that Robert produced, and the accuracy of vowel sound he produced declined with syllable length. The incidence of interharmonics also increased in words of two and three syllables.

The voice data (Table 7.11) show little change in mean F0 or voice perturbation as target syllable length increased. Of the nine three-syllable words, Robert produced just four with three syllables, generating just twenty-two vowels. He omitted a syllable from one two-syllable word.

Table 7.11 Mean voice data for Robert's imitations of words in the Syllables test, arranged by number of target syllables. The table shows mean and SD values for F0 (Hz), DUV (%), HNR (dB) and CPP (dB).

No. of target syllables	Mean F0, Hz (SD)	DUV, % (SD)	HNR, dB (SD)	Jitter, % (SD)	Shimmer, % (SD)	CPP, dB (SD)
1	148.44 (27.19)	1.49 (1.50)	15.74 (3.26)	1.35 (0.92)	8.71 (4.87)	5.66 (0.79)
2	152.82 (21.06)	2.20 (4.42)	16.62 (3.65)	1.35 (2.05)	7.33 (2.67)	5.14 (1.38)
3	149.84 (6.89)	1.09 (3.42)	16.36 (2.86)	1.07 (0.71)	8.24 (2.87)	5.25 (1.32)

Robert's imitations of two-syllable and three-syllable words are shown in Figures 7.8 and 7.9, respectively. The pitch contours show that his responses were limited in pitch movement. Syllables were poorly defined by segmental changes but the syllables were perceptually salient as a result of changes in intensity.

Example: 'hope, hopeful, hopefully'

This set was chosen as an example because of all the three syllable words Robert imitated, the word *hopefully* was produced with the greatest clarity in terms of number of syllables, after three demonstrations.

hope: [həʊ]; hopeful: [hʊəɸ]; hopefully: [ˀbəiəi]

Robert's production of the word *hope* was breathy in quality and produced with low-intensity interharmonics throughout (Figure 7.10). He did not produce the final voiceless bilabial plosive, /p/. In *hopeful* and *hopefully,* Robert did not produce any consonant sounds in medial position and his initial and final consonants were indistinct (Figure 7.11, Figure 7.12). The syllables in *hopeful* and *hopefully* were defined by changes in intensity, as visible on the waveform and the intensity contours (Figure 7.11, Figure 7.12). The pitch contour shows a gradual rise of pitch in *hopeful* and a decline in *hopefully*. Comparison of the spectrograms of *hopeful* and *hopefully* shows that there is an increase in the complexity of the interharmonics during *hopefully*. There is additional noise in

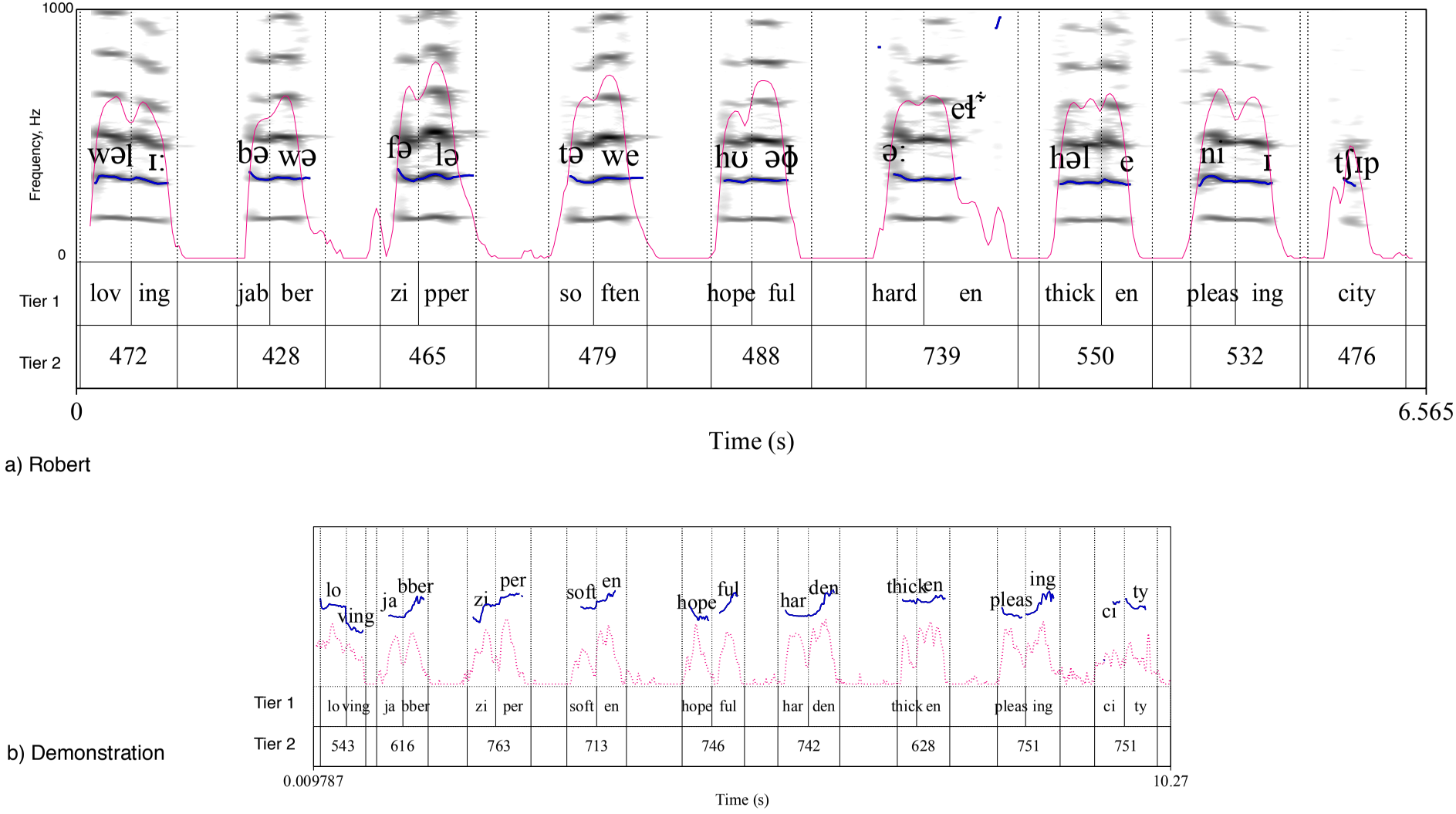


Figure 7.8: Annotated spectrograms showing Robert's imitation of words of two syllables (above) and the demonstration (below). The narrow-band spectrogram is shown for Robert's production (0-1000 Hz; window length 100 ms; 40 dB). The target syllables are shown in Tier 1 on both images and the duration (milliseconds) in Tier 2. Robert's phonetic production is shown alongside the pitch contour (range 50-300 Hz). The intensity contours are shown in pink: the intensity scale is 45-100 dB for Robert's image, and 40-100 dB for the demonstration.

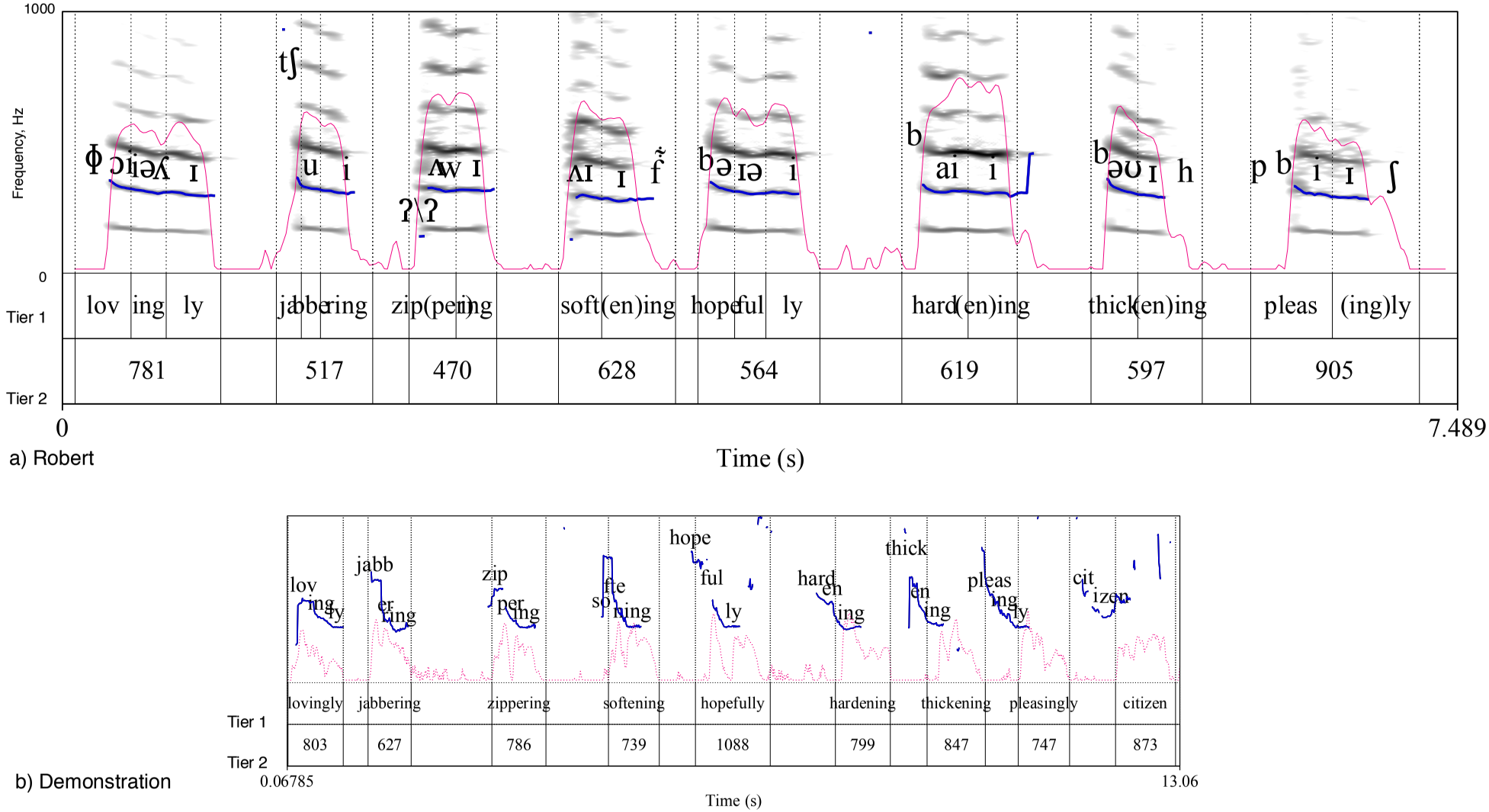


Figure 7.9: Annotated spectrograms showing Robert's imitation of words of three syllables (above) and the demonstration (below). The narrow-band spectrogram of Robert's production is shown (0-1000 Hz; window length 100 ms; 40 dB). Robert's phonetic production is shown alongside the blue pitch contour (50-500 Hz), The target syllables are shown in Tier 1 of Robert's production, and the syllables of the demonstration are placed against the pitch contour (75-500 Hz) in image b. The duration (milliseconds) of each word is shown in Tier 2 of each image and the intensity contours are shown in pink: Image a) is set at 45 dB-100 dB; image b) is set at 40 dB-100 dB.

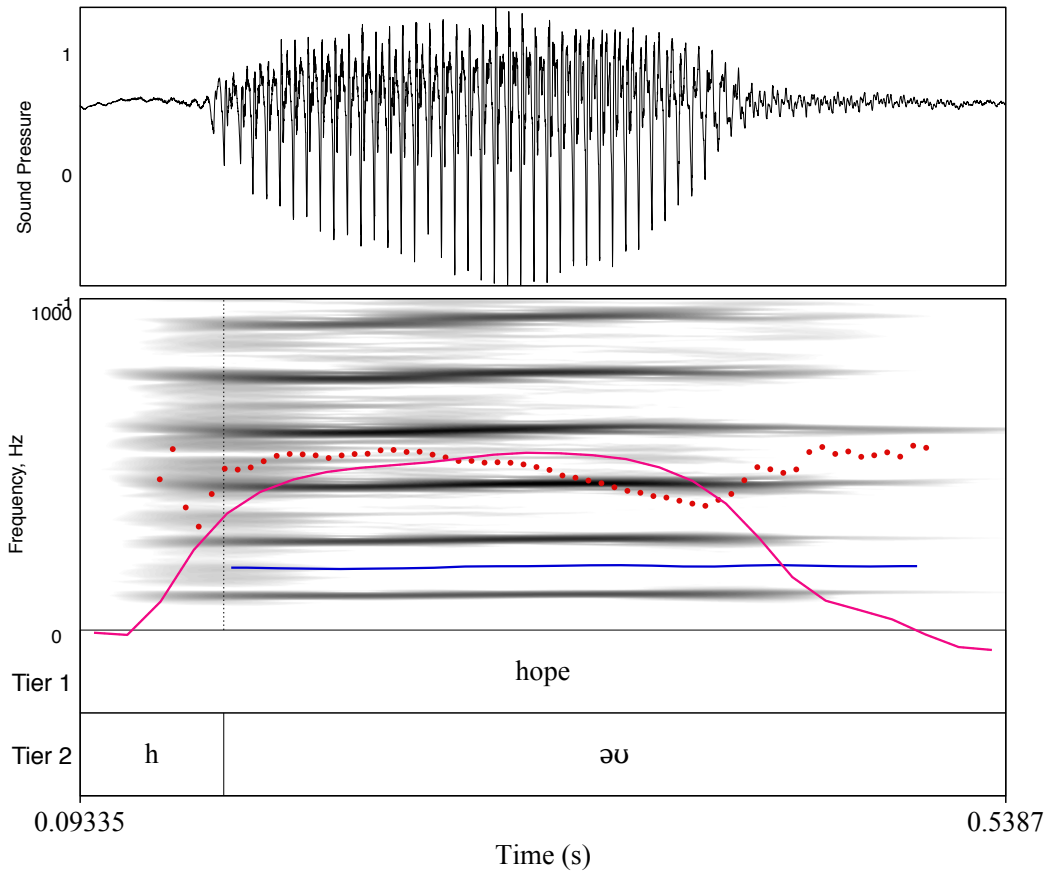


Figure 7.10: Waveform and annotated spectrogram showing Robert's imitation of 'hope'. The target word is shown in Tier 1 and Robert's phonetic production is shown in Tier 2. The narrow-band spectrogram (0-1000 Hz; window length 100 ms; 40 dB) is overlaid with the first formant frequency (in red dots), the pitch contour (blue: 50-500 Hz) and intensity contour in pink (40-100 dB).

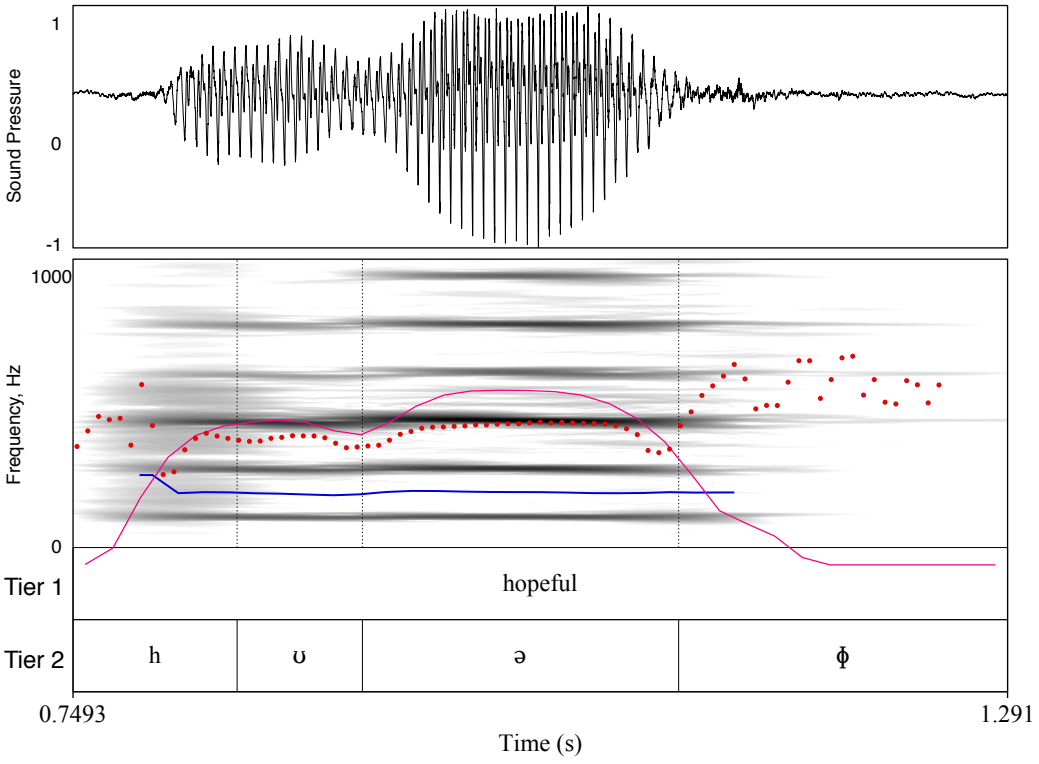


Figure 7.11: Waveform and annotated spectrogram showing Robert's imitation of 'hopeful'. The target word is shown in Tier 1 and Robert's phonetic production is shown in Tier 2. The narrow-band spectrogram (0-1000 Hz; window length 100 ms; 45 dB) is overlaid with the first formant (in red dots), the pitch contour (blue: 50-500 Hz) and intensity contour (pink: 50-100 dB). There is a rise in intensity contour that indicates the onset of the second syllable.

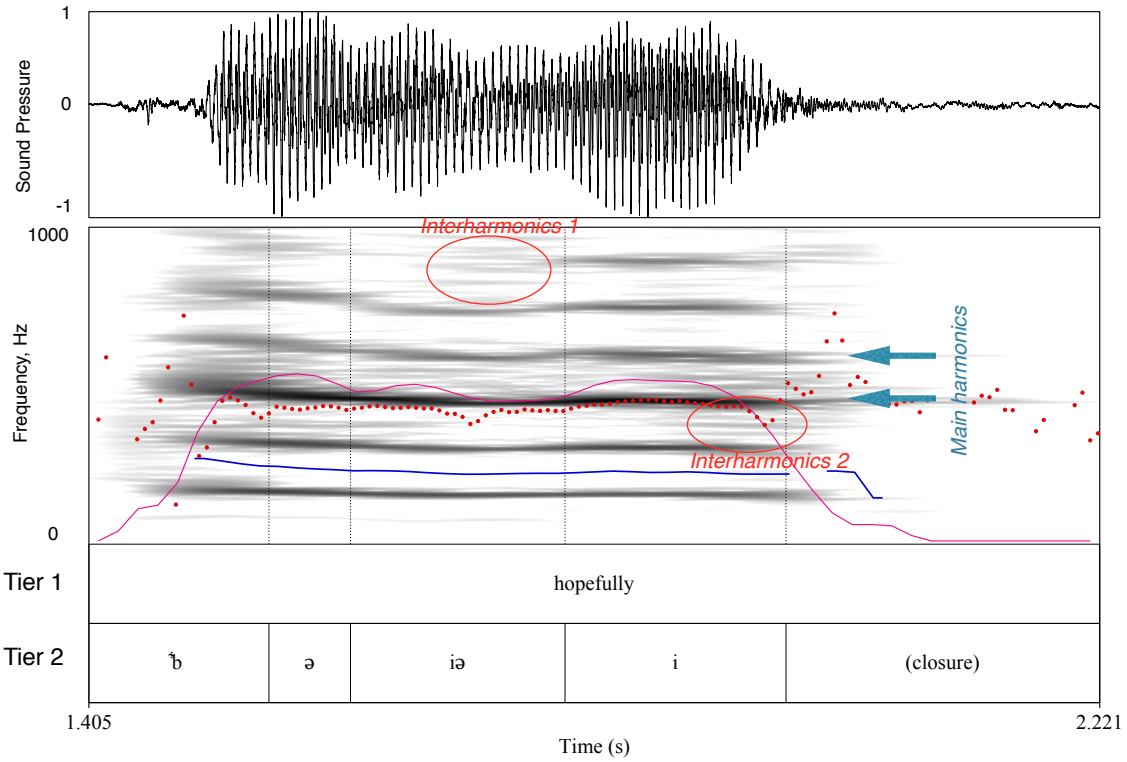


Figure 7.12: Waveform and annotated spectrogram showing Robert's imitation of 'hopefully'. The target word is shown in Tier 1 and Robert's phonetic production is shown in Tier 2. The narrow-band spectrogram (0-1000 Hz; window length 100 ms; 45 dB) is overlaid with the first formant frequency (in red dots), the pitch contour (blue: 50-500 Hz) and intensity contour (pink: 50-100 dB). There is a rise in intensity contour that indicates the onset of the second syllable. The red ellipses indicate the presence of multiple interharmonics in the upper frequency range (1) and lower frequency range (2). Examples of the main harmonics are indicated by blue arrows at offset.

the upper frequency range during the diphthong /iə/ (Figure 7.12: Interharmonics 1), and relatively higher-intensity interharmonics towards offset (Figure 7.12: Interharmonics 2).

7.3.4 Imitation of sung phrases

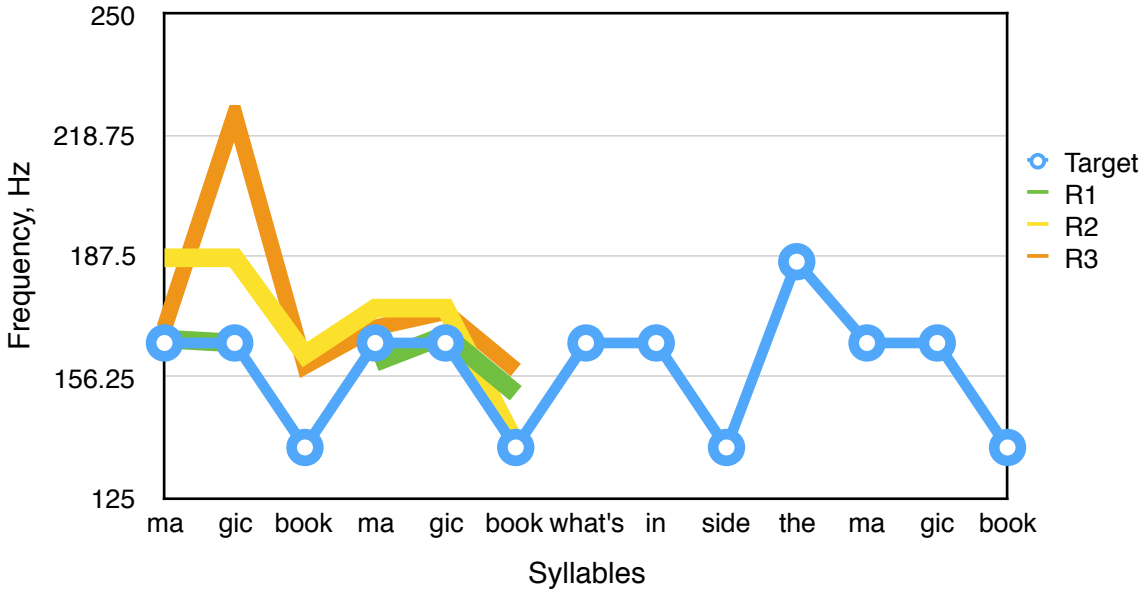
Magic Book

Robert's ability to reproduce the song *Magic Book* improved over a four-week period (Figure 7.13; Figure 7.14). During the first assessment (R1: week 2), Robert reproduced just two words and a syllable (Figure 7.13). During week four (R2) he recalled the words to the first phrase and

Target	R1 (week 2)	R 2 (week 4)	R3, solo (week 6)
1. **ma**gic book	1. {Ç **tdæ̰**ʃi Ç}	1. {V̤ mæ**bæ** ɪx bʊ	1. {V̤ ia **ʒɪʒɪ bʊ**
2. **ma**gic book	2. **bei**ʃi pə**bɔ**	2. pə**bə** ɪ̰ʃ V̤}	2. ʒæ**ʒɪx** bʊ V̤}
3. **what's** inside the			3.
4. **ma**gic book			4. **tʃæ**ʒɪx **bʊ**

Figure 7.13: Phonetic transcription and syllabic stress of Robert's production of 'Magic Book', over a four-week period. Robert' s recall of the song lyrics improved by week 6. The bold type shows the perceived primary stress, and the underlined text indicates secondary syllable stress. The Ext IPA symbols and VoQs indicate Robert's voice quality, which was intermittently whispery and creaky (in R1).

Figure 7.14: Graph showing the development of Robert's recall of the words and melody of 'Magic Book', over a four-week period. The figure shows the melodic contour of Robert's performance of the song, over a four-week period (R1, R2, R3). Robert produced just 3 words (R1) in his first imitation. The target contour is shown in blue.



reproduced the melodic shape more accurately (Figure 7.14, R2). In his solo performance (R3), he recalled part of the melodic shape (Figure 7.13) and his phonetic production was more consistent between phrases (Figure 7.13).

Table 7.12 shows the articulatory, melodic and rhythmic features of the first phrase of Robert's first (R1) and final (R3) productions of *Magic Book.* The data show improvements in the number of appropriate consonant sounds and his articulatory rate. The range of Robert's melodic changes increased from 0.1-1.6 ST (R1) to 0.5-4.7 ST (R3). In the final production he produced the words *magic book* with appropriate relative duration of syllables.

Monkey Man

The song *Monkey Man* (Hibbert, 1969, adapted) was taught with accompanying Makaton signs to assist and prompt word recall. Robert's phonetic production from the first teaching session is shown in Table 7.13. The song was taught as 'call and response', one phrase at a time. It was initially demonstrated verbally, then with Makaton signs, which were used to prompt Robert's word recall in his second, third and fourth production (Table 7.13). Robert produced more speech sounds with every production (Table 7.13) but his production of words was most accurate for the signs with which he was most familiar (*eye*, *monkey* and *man*). There was a delay between the presentation of the Makaton sign and Robert's production of the word.

Robert's voice quality in the first production (R1) was very quiet and breathy in quality (Table 7.13). His voice production in R3 and R4 was louder but the final /a/ in line 1 (R3) was produced with

Table 7.12: The segmental, articulatory, melodic and rhythmic features of the first phrase of 'Magic Book' as demonstrated, and as produced by Robert during his first and final productions. Segmental production details the number of consonant sounds (C) Robert produced in the phrase ('magic book, magic book'), and their relevance to the target sound in terms of manner and place. Articulatory features measure the rate of production of consonants and vowels and the relative duration of these. The table also shows the mean duration and relative duration of syllables in the words (rhythmic features).

Demonstration: key features		**Target**	**First production**	**Final production**
Segmental production	**Number of C sounds**	**10**	5	7
	Number of relevant sounds	**10**	3	5
Articulatory features	**Total duration C, ms**	**856**	795	805
	Articulatory rate: sounds per second	**11.8**	6.3	8.7
	Total duration V, ms	**1577**	1181	1833
	Average duration of C:V	**100:184**	100:148	100:227
Melodic features	***Magic*: Mean ST change**	**0**	0.1	2.1
	***Book*: Mean ST change**	**3**	n/a	2.3
	***Magic*: Mean ST change**	**0**	0.3	0.5
	***Book*: Mean ST change**	**3**	1.6	4.7
Rhythmic features	**Mean duration of *magic*, ms**	**398**	749	729
	Mean duration of *book*, ms	**797**	441	505
	Relative duration of syllables *ma:gic*	**1:1**	1:2	1:1
	Relative duration of the words *magic:book*	**1:2**	2:1	1:2

Table 7.13: Robert's phonetic production of 'Monkey Man'. The table shows the changes in Robert's phonetic production of the song in four different learning conditions. ExtIPA symbols and VoQS are used to indicate Robert's vocal quality, which was mostly whispery with intermittent whispery creak. Stressed syllables are in bold text. Lyrics are reproduced with permission.

Phrase no. and words		**R1: phrase imitation, verbal only**	**R2: imitation, with Makaton**	**R3: supported with verbal prompts and with Makaton**	**R4: supported with Makaton only**
1	**aye aye aye**	{V̩ a	{V̩ a la la	{V̩ ai a̰	{V̩ ai ai a
2	**aye aye aye**	a	ai a a	ai ai a̰	a̰i ai a̰
3	**everybody**	ewi əu	ai eβə ɔ	ʌwi̯ə bɔ:	ʌ̰wɪ bɑɛ̰**:**
4	**talk about the**	deu	tɔ	tɔ̰x	gʌɛ uə V̩} θə
5	**big monkey**	wɛ V̩}	mʌgi	mʌŋki	**mʌŋ**ki
6	**man**		mæ V̩}	bæ̰ V̩}	bæ̰

audible shimmer. There was an increase in the incidence of a 'creaky' quality (e.g. /a̰/) in each subsequent production: for example, in R3, two vowels in lines 2 and 3 (R4) were produced with creaky voice quality; and in R4, four vowels were creaky in quality (Table 7.13). In R4 Robert produced more sounds in the fourth phrase but these were whispered. After the first imitation (R1) the word *monkey* was produced with near-modal quality and with an increase in volume (Table

7.13, R2 - R4). His final production (R4) was also produced with appropriate stress on the first syllable.

Robert's initial production was a monotone, but he began to echo the melodic shape after further demonstration and teaching of the Makaton signs (Figure 7.15). When supported with pitch prompts at the start of phrases and with Makaton signs, his third production (R3) was accurate in the first phrase (*aye aye aye*) (Figure 7:15, R3). When cued with Makaton signs only, (Figure 7.15, R4) he produced the melodic shape and most words.

Figure 7.15: Graph showing Robert's melodic production of 'Monkey Man', with verbal prompts (R3) and with Makaton prompts only (R4)

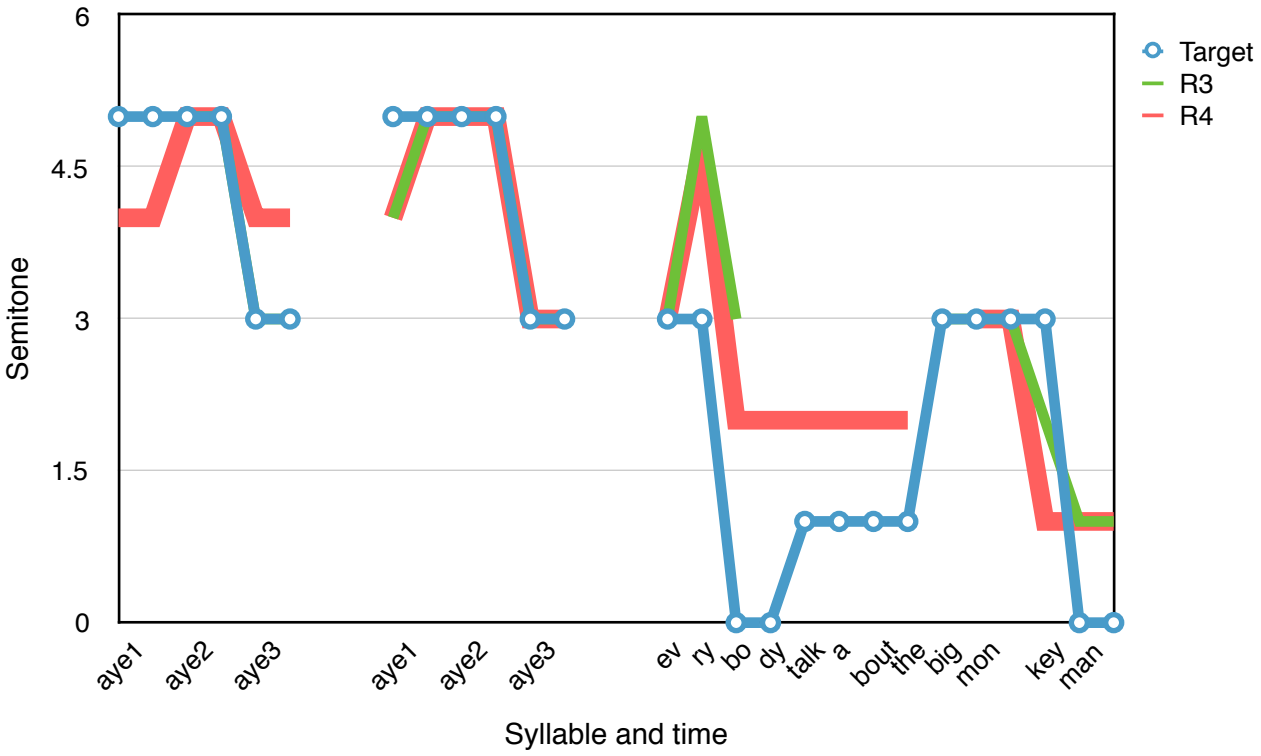


Figure 7.16 shows the final phrase of the song, which was supported with Makaton signs throughout and two verbal cues (Tier 3: indicated by T). The recording had been filtered in Praat to remove much of the background noise (see Chapter 3, Section 3.7.4), but the image retains the key features of the pitch contour and relative intensity levels. The pitch contour shows a descending pitch in the last phrase (*big monkey man*), that echoes the melody of the song. The waveform and the intensity contour also reflect Robert's relatively louder vocal production. His production of the word *man* was creaky in quality, and this resulted in a pitch break (Figure 7.16: *2).

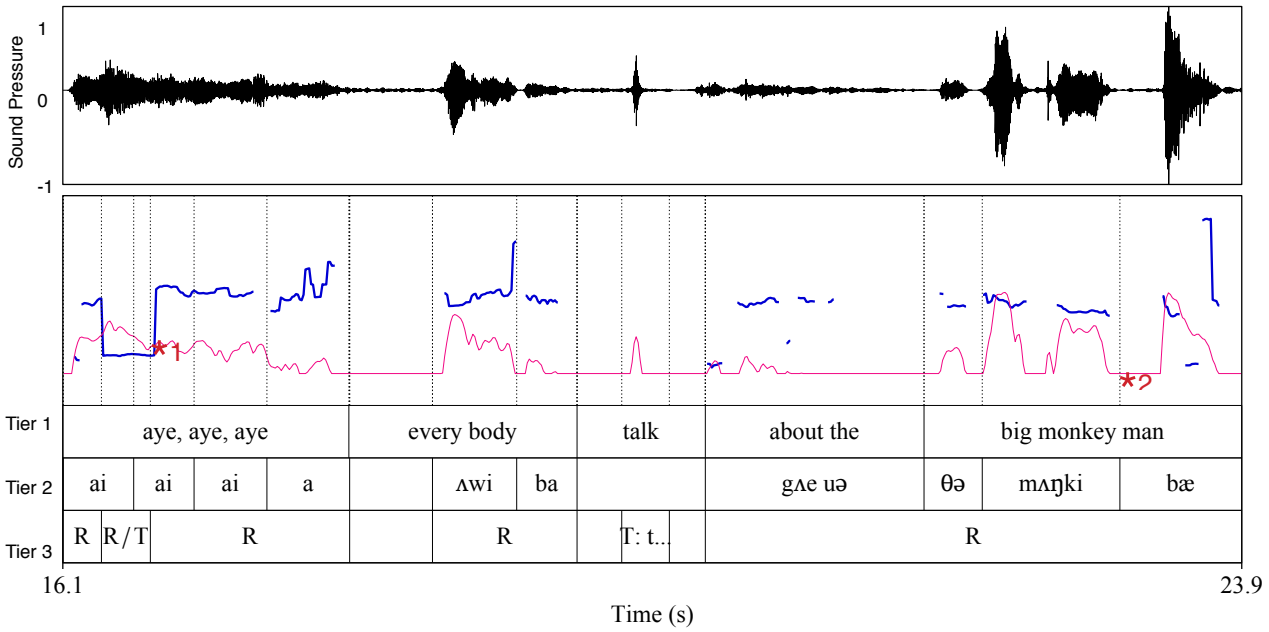


Figure 7.16: Waveform and image showing details of Robert's production of 'Monkey Man', phrase 4. The image shows the filtered waveform (see Chapter 3). Tier 1 shows the target words, Tier 2 shows Robert's phonetic production, and Tier 3 indicates where the researcher interjected (T) with verbal prompts, and indicates which sections were produced by Robert (R). The pitch contour was adjusted to 60-300 Hz for improved resolution: the first pitch break (*1) shows overlap between Robert's production and a pitch prompt from the researcher (Tier 2). The second pitch break indicates an atypical vocal quality. The pink intensity contour shows the range 50-100 dB: the contour and waveform both show an increase on Robert's words *monkey* and *man*.

7.4 Prosody and voice in production of speech and song from LTM

7.4.1 Comments to peers

Robert's spontaneous speech when commenting on his peers' singing was not always intelligible and his phonetic production of known words was inaccurate (Figure 7.17). Comment 1 was produced with slight pitch movement across the phrase as a whole (Figure 7.18): the comment *very good* was produced at a higher mean pitch than the unintelligible phrase (188.47 Hz and 169.04 Hz, respectively) but there was minimal variation of pitch within each phrase (4.9 Hz and 4.4 Hz, respectively). Comment 2 was produced with less pitch movement (Figure 7.19). In both phrases, the word *good* was emphasised (Figure 7.17). The word *very* was produced with one syllable in comment 1, and the word *singing* in Comment 2 was also monosyllabic.

Comment	Interpretation
1. e **gʊəʰ**…βiɔiʌ**wʌ**bwʌ	1. very **good** (unintelligible)
2. fɪu**gʊθ** enai sɪ	2. very **good**, very nice singing

Figure 7.17: Phonetic transcription and syllabic stress of Robert's comments to peers: comments 1 and 2. The image shows Robert's comments (IPA transcription). Not all words were intelligible and possible interpretations are given. The syllables that were produced with perceived stress are shown in bold. Robert's voice quality was close to modal.

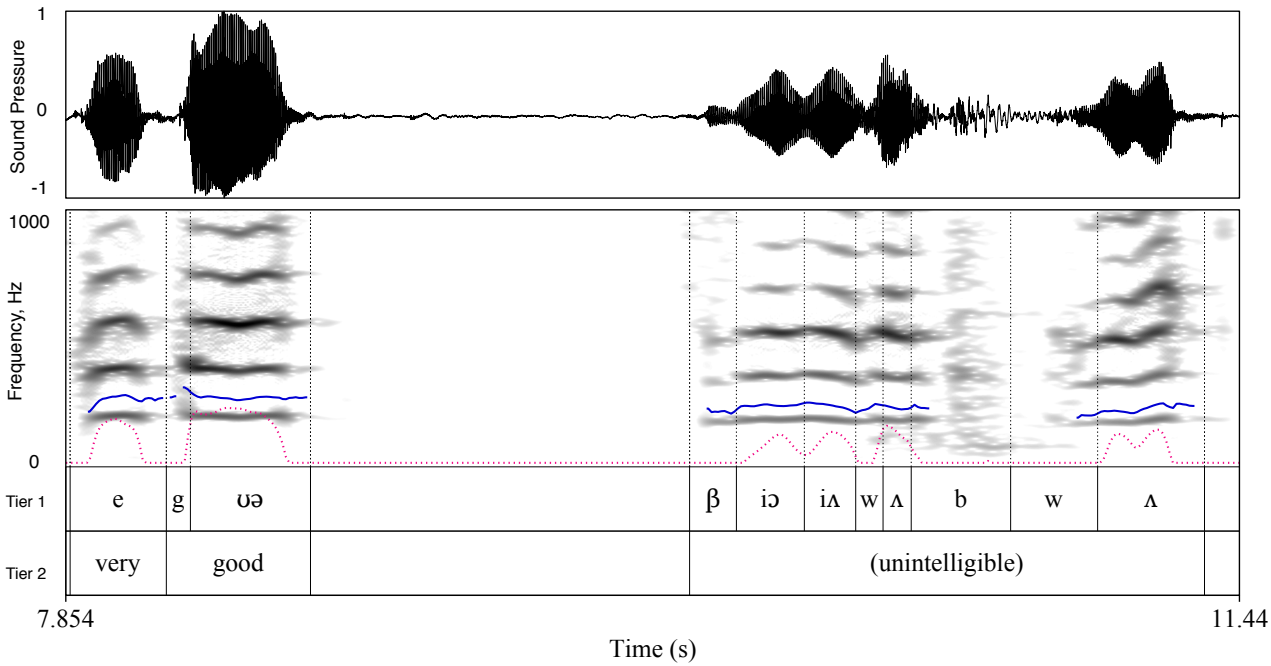


Figure 7.18: Waveform and annotated spectrogram showing Robert's 'Comment 1'. The narrowband spectrogram (0-1000 Hz; window length 100 ms; 40 dB) is overlaid with the pitch contour (blue: 50-500 Hz) and intensity contour (pink: 40-100 dB). Robert's phonetic production is given in Tier 1 and an interpretation is given in Tier 2.

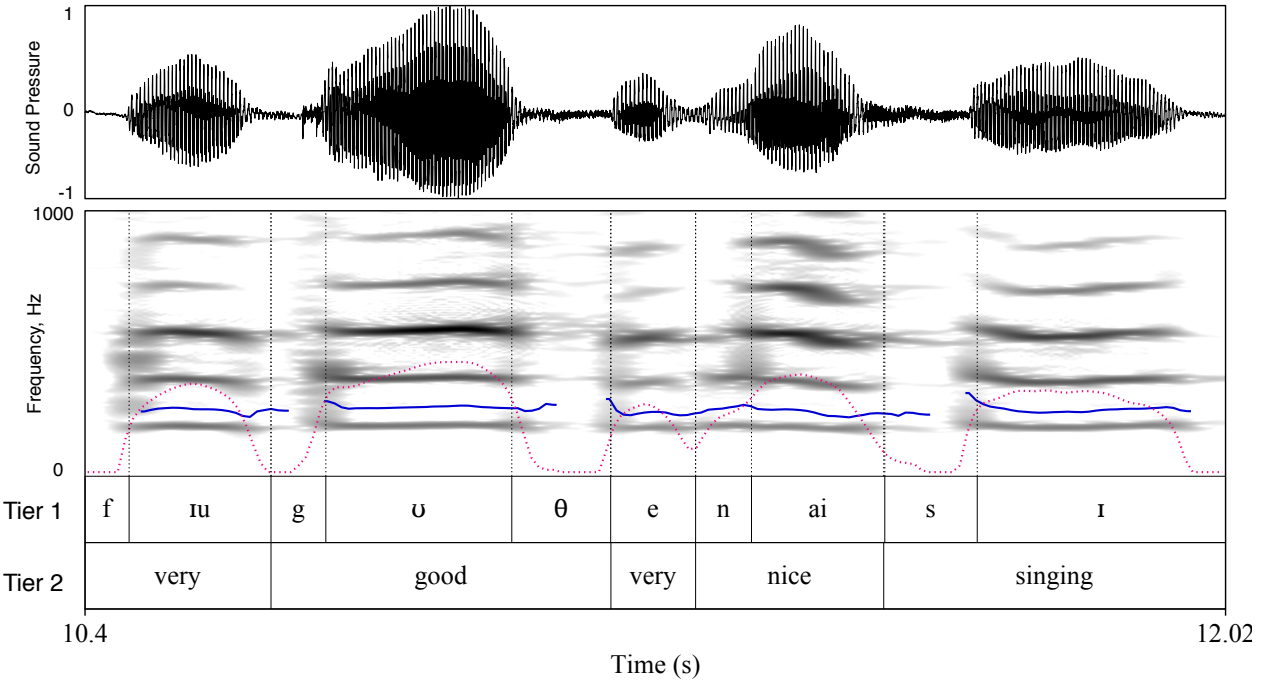


Figure 7.19: Waveform and annotated spectrogram showing Robert's 'Comment 2'. The narrowband waveform spectrogram (0-1000 Hz; window length 100 ms; 40 dB) is overlaid with the pitch contour (blue: 50-500 Hz) and intensity contour (pink: 40-100 dB). Robert's phonetic production is given in Tier 1 and a translation in Tier 2.

The third comment (Figure 7.20) was partly imitated. The full transcript is given in Figure 7.20 and the imitated and spontaneous sections of Robert's speech (edited to remove the researcher's speech) are indicated on Figure 7.21 (Tier 3). Robert produced the words *October* and *Hallowe'en* with just two syllables and with omitted phonemes. His voice quality was perceptually breathy throughout. The break in pitch contour (Figure 7.21) shows the onset of creaky phonation at the

Transcript	Robert's production	
T1: What day is it today?		
	R1: Monday	R1: {V̥ **mʌ**nde
T2: Monday the thirty-first of…		
	R2: (thirty) first	R2: **pes**
T3: of October:		
	R3: (Oct)ober	R3: **bəbe**
T4: it's…		
	R4: hallowe'en… Oh yeah!	R4: **hɔwɪ̰ jaije** V̥}

Figure 7.20: A phonetic transcription of Robert's 'Comment 3', showing the interaction between the researcher (T) and Robert (R). Robert's phonetic production is shown (IPA transcription) and the syllables that he produced with perceived stress are shown in bold. Ext IPA symbols and VoQS indicate his breathy production and a creaky vowel.

transition between the end of the word *hallowe'en* and *oh*. Detail of the word *Monday* and of the researcher's subsequent production of the word, are shown in Figure 7.22. Robert's production of the word *Monday* was articulated clearly, and initial stress was placed on the first syllable: it was elongated and produced with similar intensity across the /m / and the vowel (Figure 7.19). The intensity dropped sharply on the second syllable, and the pitch and intensity contour are different to the researcher's subsequent production.

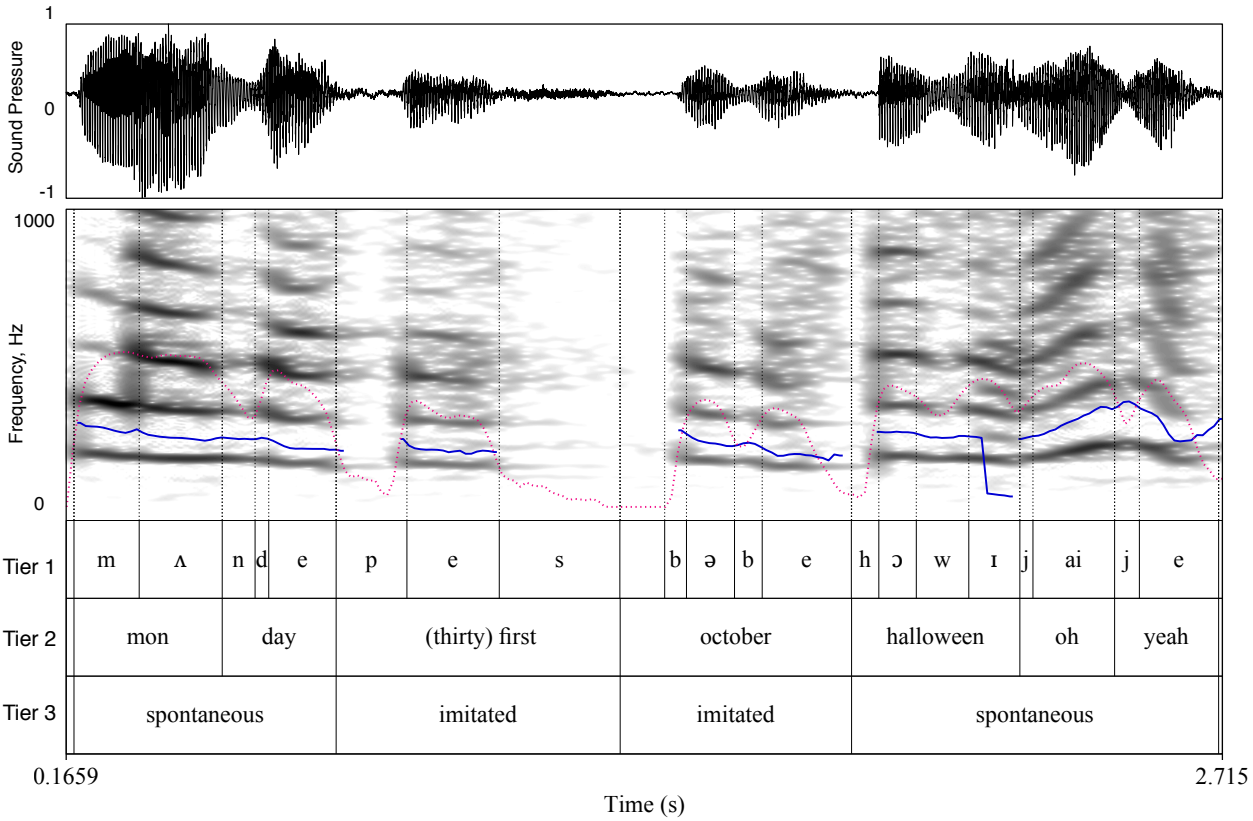


Figure 7.21: Waveform and annotated spectrogram showing Robert's 'Comment 3'. The image shows the narrowband waveform (0-1000 Hz; window length 100 ms; 40 dB) waveform and spectrogram of Robert's partly imitated phrase. The pitch contour (blue: 50-500 Hz) and intensity contours (pink: 35-100 dB) are shown. Robert's phonetic production is given in Tier 2 and a translation in Tier 3. Tier 3 indicates the sections that were imitated. The researcher's prompts have been edited out of the image: a detail of *Monday* is given in Figure 7.22.

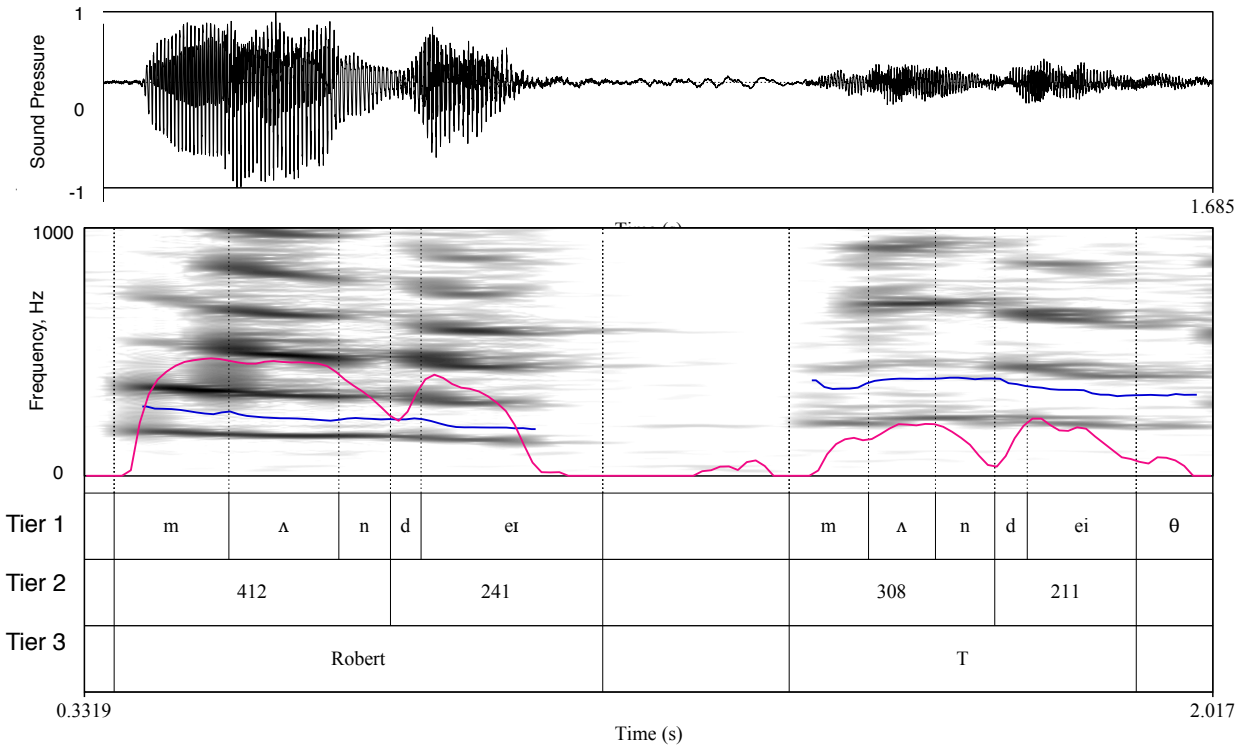


Figure 7.22: Waveform and annotated spectrogram showing a detail of Robert's 'Comment 3'. The image shows the narrowband waveform (0-1000 Hz, window length 100 ms, 40 dB) waveform and spectrogram of Robert's production of the word 'Monday' and the researcher's repetition of the word (indicated by T, in Tier 3). The pitch contour (blue: 50-500 Hz), intensity contours (pink: 35-100 dB) and duration of syllables are shown (Tier 2). The phonetic productions are given in Tier 1.

7.4.2 Happy Birthday: transfer from song to speech

Robert was unable to speak or sing the words to *Happy Birthday*, unaided. He required prompts at the start of each phrase (Figure 7.23). After speaking the song, Robert required more prompts to sing (Figure 7.23: Singing 2) than he had needed in an earlier session (Singing 1).

Target	Singing	Spoken	Singing 2
1. happy **birth**day **to you** 2. happy **birth**day **to you**	1. (*happy*) **bə**bei dʊ **jʊ** 2. (*ha*) æwə bə:dei **hʊ jʊ** *Interharmonics: n/k*	1. (*happy birthday*) æwə **bə:**te 2. (*happy birthday to you*) {V̤ **pə** ju V̤} *Interharmonics: 75%*	1. (*ha*) {Ç̤ **æ**bə **bə:**de 2. (*to*) bə **bə:**dei 3. **æ**bə **bə:**de 4. (*to you*) wə Ç̤} **ju:** *Interharmonics:*

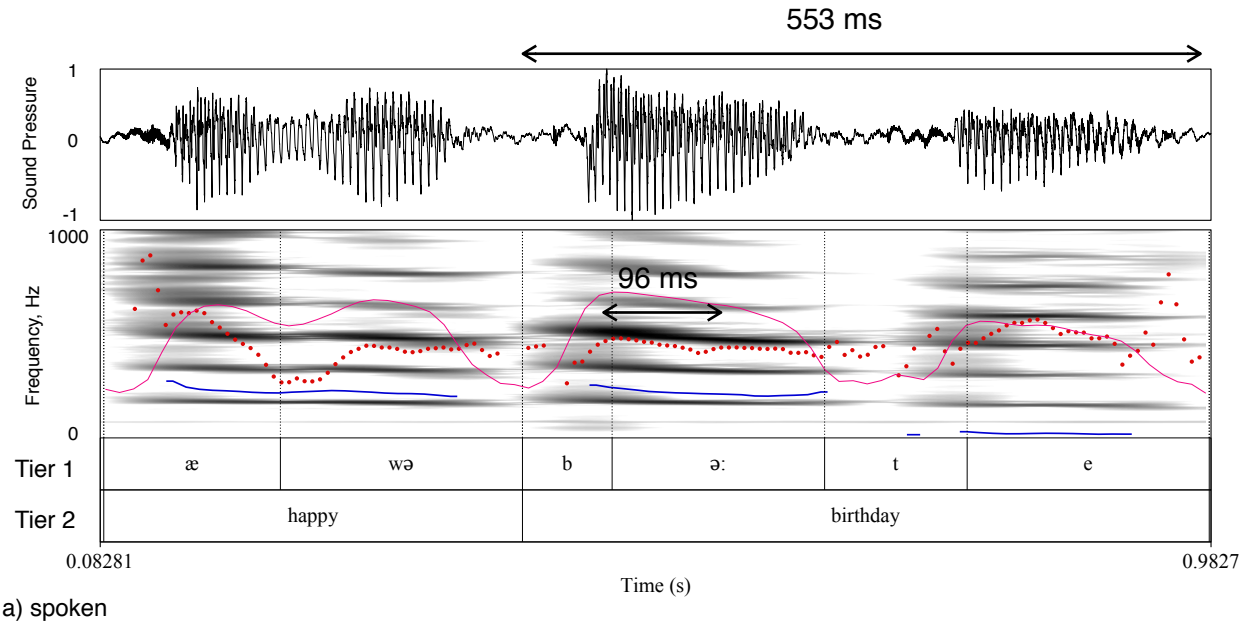
Figure 7.23: Phonetic transcription and syllabic stress of Robert's sung and spoken productions of 'Happy Birthday'. The target words and the phonetic transcription of Robert's sung and spoken productions are given. Primary stress is indicated in bold type: secondary stress is underlined. Robert produced the first phrase only, and needed a prompt to complete the spoken phrase: the prompts are shown in italics. The percentage duration of voiced sections that were produced with interharmonics is shown: however, it was not possible to calculate this for Singing 1, which was contaminated by background noise. Ext IPA and VoQs indicate the voice qualities for Robert's spoken production, and for his second sung production. Both were mostly breathy, with some creaky production.

The spoken and sung performances from the second session were produced with similar pitch ranges as measured in Hz (Table 7.14); but both sung productions were wider in range, as measured in semitones (Table 7.14). Excluding pauses for prompts, Robert's spoken production was faster, but the relative duration of Robert's syllables was closer to the target in the sung condition (Table 7.14). Robert emphasised the syllable *birth* in his sung versions, whereas his spoken production was produced with less rhythmic variation between syllables.

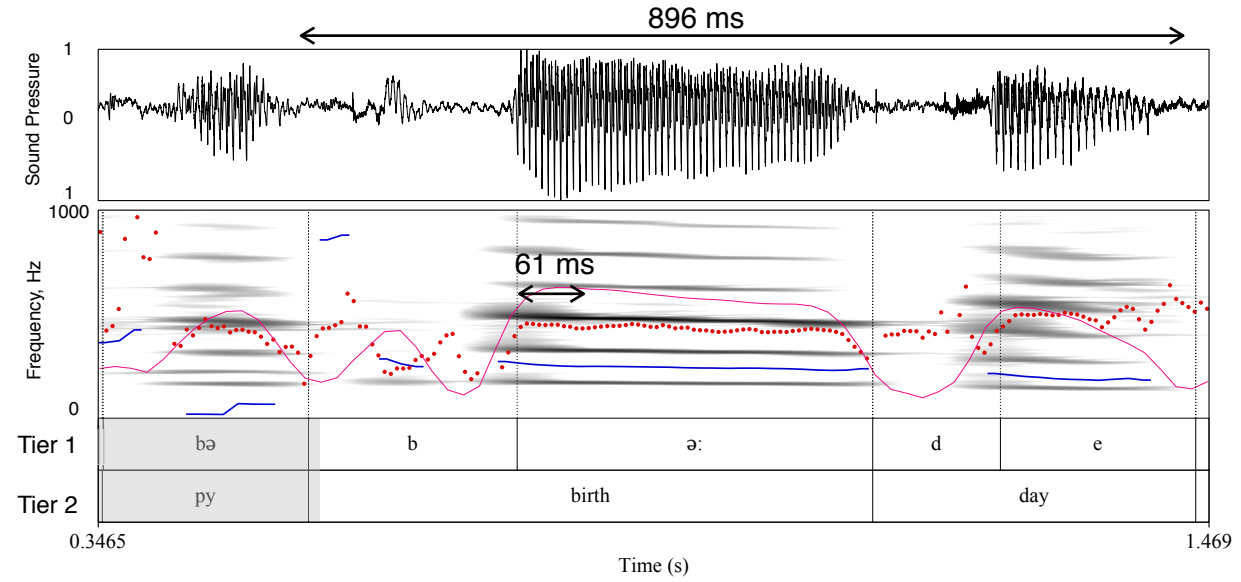
Figure 7.24 shows the waveform and spectrogram of the phrase *happy birthday* in both conditions. The sung phrase (Figure 7.24b) immediately followed phrase 2 of Robert's spoken production (Figure 7.23). Robert's spoken production (Figure 7.24a) was produced more quickly, as indicated by the duration of the word *birthday.* The slower speaking rate and the elongated vowel in the first syllable distinguish his sung production from his spoken production. The pitch and intensity contours are similar for both productions, but the sung production (Figure 7.24b) was produced with less intense interharmonics (e.g. at the onset of 'birth'). The duration of these from the onset of the vowel was longer in the spoken condition.

Table 7.14: The relative duration and pitch of Robert's syllables in 'Happy Birthday', in sung and spoken conditions. The table shows the change in mean pitch (in semitones) between syllables (ST), and the relative duration of syllables (RD), in comparison to the shortest syllable's duration in the phrase (100 ms). The (+) indicates a rising melodic interval the (-) indicates a falling melodic interval. The asterisks marked the sections that were prompted.

Syllable	**Target pitch change, ST**	**Target RD: 1=100 ms**	**SInging 1 (phrase 2)**		**Spoken (phrases 1 & 2)**		**Singing 2 (phrases 3 &4)**	
			Pitch change, ST	**RD: 1=100 ms**	**Pitch change, ST**	**RD: 1=100 ms**	**Pitch change, ST**	**RD: 1=100 ms**
ha	**0**	**3**	0.0*	1	0.0*	1	0.0	1
py	**0**	**1**	+0.7	1	-0.7	1	-0.7	1
birth	**+4**	**2**	+0.4	2.5	+0.1	3	+0.6	2
day	**-4**	**2**	-1.4	2	-2.4	3	-1.7	1
to	**+5**	**2**	+0.8	1	+2.0*	3	+2.4*	1.5
you	**-2**	**4**	+3.6	3	0.0	2.5	+0.2	2
Pitch: mean and range (Hz)			**164 Hz; 23 Hz**		**146 Hz; 31 Hz**		**158 Hz; 34 Hz**	
Pitch range (max-min) ST			**5.00**		**3.10**		**4.10**	
Mean rate, syll/second			**2.67**		**4.37**		**3.10**	



a) spoken



b) sung

Figure 7.24: Happy Birthday, spoken and sung. The images show the narrowband waveform (0-1000 Hz; window length 100 ms; 40 dB) waveform and spectrogram of Robert's production of the phrase *Happy Birthday* in spoken (a) and sung (b) conditions. The spectrograms are overlaid with the pitch contour (blue: 70-500 Hz), intensity contours (pink: 35-100 dB) and first formant frequency contour (red dots). Tier 1 shows the sounds Robert produced and Tier 2 shows the syllables. The duration, in milliseconds, of the word *birthday* is shown above the waveform. The shaded box in image (b) shows where Robert's production overlapped with the verbal prompt, which contaminates the spectrogram. Comparison of the spectrogram in the vowel of 'birth' shows a difference in the intensity and duration (marked with an arrow) of interharmonics at the onset of the vowel.

Discussion

The information from Robert's college records and the data from core tasks (Section 7.1) indicate that Robert's speech and musical production are affected by his severe hearing loss. Therefore, this section will begin with a discussion of the possible impact of his hearing impairment (HI) on his motor movement, his perception and production of speech, and his perception of music (Section 7.5). Section 7.6 will discuss his motor timing in beat and rhythm entrainment and imitation tasks, and his rhythmic production of speech and song. His abilities in imitating melody and prosody will also be discussed in Section 7.6. Section 7.7 will consider the evidence that Robert can develop musical skills and the extent to which he can transfer learning from song to speech. Section 7.8 will summarise Robert's abilities and difficulties and will discuss the potential benefits of using singing activities to develop aspects of his speech perception and production.

7.5 The implications of Robert's moderate hearing loss for motor control and auditory perception

7.5.1 Motor movements

Hearing loss impacts upon postural balance, the speed of motor movements and upon motor co-ordination (for a review, see Rajendram, Roy, Jeevanantham, 2012). The evidence suggests that people with severe-profound HI are most at risk of motor difficulties (Walicka-Cuprys, Przygoda, Czenczek, Truszczyńska, Drzał-Grabiec, Zbigniew and Tarnowski, 2014; Wong, Leung, Poon, Leung and Lau, 2013). However, those with any degree of HI show less than optimal performance in tasks measuring postural control and motor skill (Rajendram et al., 2012). The effects of HI on gross-motor abilities in those with DS are unknown: those with HI were excluded from studies investigating postural control in DS (Shumway-Cook and Woollacott, 1985; Rigoldi et al., 2011) and were not assessed by Malak et al. (2013) in their study of balance and motor skills. For Robert, any difficulties in gross-motor functioning and balance that are associated with DS may be exacerbated by his hearing impairment.

7.5.2 Speech perception

Hearing loss can affect speech perception through reduced sensitivity to acoustic cues (Blamey, Sarant, Paatsch, Barry and Bow, et al., 2001; Revoile, 1999). Individuals with HI of over 45 dB may have a reduced ability to detect frequency, that is necessary for the identification of vowels (Vatti, Santurette, Pontoppidan, and Dau, 2014). According to Robert's audiogram, he may have a reduced ability to identify vowels, although the residual hearing in his right ear may be adequate. Robert had difficulties in identifying two vowels in the minimal pairs test. He had greater difficulty in perceiving the consonant contrasts in the minimal pairs test (Table 7.1: test 13): his bilateral hearing loss of over 40 dB in the 250-1000 Hz range would have masked the initial consonants of

/b/ and /m/ in the items that he failed to discriminate in the minimal pairs task (Table 7.1: test 13). However, he correctly discriminated the higher-frequency sounds /v/, and /ʃ/ in this task, which would be expected to be above his hearing threshold (50-70 dB at 2000-4000 Hz): only the low-frequency sound /g/ (correctly discriminated) was within his expected hearing abilities at normal conversational speech levels. With the exception of /g/, the initial consonants of all test items given were articulated towards the front of the mouth. Potentially, Robert used visual cues to identify place of consonant, which limited his choices, allowing him to discriminate *shoe f*rom *two*, *zoo* and *shoe*. However, acoustically, Robert's moderate hearing loss would mask the stop-burst of /b/ in *bat* (Revoile, 1999). This would make it difficult for Robert distinguish *bat* from the target word *mat* in the minimal pairs test (Appendix 2, p. 328).

Hearing impairment has consequences for speech development, depending upon its severity and age of onset (Blamey et al., 2001; Laws and Hall, 2014). Robert's repertoire consists primarily of forward-placed speech sounds, whose movements are visible to an observer, which can support perception of speech sounds (Massaro and Light, 2004). For sounds whose movements are less visible, his HI will have affected his ability to distinguish the acoustic cues that are necessary to learn to articulate a wider repertoire of speech sounds (Revoile, 1999). Furthermore, this will have affected his ability to form stable representation of words. As a result, as for children with hearing loss, he will be slower to learn new words and require repetition (Stelmachowicz, Pittman, Hoover and Lewis, 2004).

7.5.3 Music perception

Reduced frequency resolution in people with bilateral hearing loss affects their ability to perceive aspects of pitch (Santurette and Dau, 2007). They may also struggle to separate the melody from competing sounds, especially if distracting sounds are intense (Marozeau, Innes-Brown and Blamey, 2013; Vatti et al., 2014). Research by Santurette and Dau (2007) shows that people with HI are less able to detect binaurally-presented melodies than those without HI (Santurette and Dau, 2007). It is probable that Robert's HI affected his ability to match pitches (Figure 7.6) and intervals (Figure 7.7). It is also likely that the high levels of background noise in the assessments and during teaching sessions made it more difficult for Robert to complete tasks.

7.6 Rhythm, voice and melody in speech and song

It is assumed that Robert's hearing loss affects his performance as a result of reduced hearing acuity. However, the types of difficulties associated with DS will also affect his performance. The following sections will therefore focus on how these factors may interact or affect different aspects of his performance, in order to determine the extent to which different factors may affect his performance in rhythmic motor tasks, speech and song.

7.6.1 Temporal perception and production in music tasks

Robert's difficulties in producing rhythmic gross-motor movements (Table 7.1:18) indicate some difficulties in maintaining balance. His poor co-ordination, slow movements and unstable posture are consistent with research into DS (Latash, 2007; Rigoldi et al., 2011; Shumway-Cook and Woollacott, 1985), and may be affected by his HI (Section 7.5.1). Robert's difficulties in producing well-co-ordinated motor movements may have consequences for his perception of beat and rhythm tasks. Experience in movements that engage the vestibular system has been shown to affect the perception of beat and rhythm in infants and adults (Phillips-Silver and Trainor, 2007, 2008). In addition, motor timing in tapping tasks is linked to motor development (Kuhlman and Schweinhart, 2000; Phillips-Silver, 2009). Robert was unable to synchronise to the beat: his ITI (%) and SD of AT (Table 7.6) exceed reported norms (Repp, 2005; Repp and Su, 2013), and could reflect a general motoric delay. However, Robert's improved ability to reproduce a rhythm when supported visually (Figure 7.2) indicates that reduced motor skills alone cannot account for his poor performance on the beat entrainment task.

The ability to maintain a steady beat is dependent upon the ability to perceive and lay down an accurate temporal template (Thaut, 2008). In beat entrainment, Robert was able to respond to tempo changes (Table 7.6) which indicates that he heard the stimulus. However, despite Robert's inaccuracy in the rhythmic tasks, he was more accurate and consistent in his timing between taps (ITI %, SD of AT: Table 7.6) at faster tempi. This characteristic is typical of younger children (Drake et al., 2000; Patel, 2008; Reifinger, 2006; Repp, 2005; Repp and Doggett, 2007). Children of 4 years are more accurate in tapping tasks when the ITI is between 300-400 ms, and they may fail to attend to or discriminate tempi outside of these limits (Drake et al., 2000). The tempo at which Robert played most accurately was 160 bpm (ITI: 375ms), which is close to the perceptual threshold for young children reported by Drake et al. (2000). This is consistent with both his verbal MA and nonverbal MA (Table 7.1) and indicates that his difficulties are associated with a general development delay. However, as a result of his reduced digit span (Table 7.1), it is possible that Robert has a difficulty in perceiving the slower temporal patterns (Flaugnacco et al., 2014; Grahn and Schuit, 2012; Saito, 2010; Snyder, 2001). In TD adults, sensory memory for auditory information lasts between 1.5 and 4 seconds (Grahn and Schuit, 2012), and the data suggest that Robert's memory may be close to the lower limit. For example, at 160 bpm, one bar of four beats is 1.5 seconds in duration, but the duration of a bar at the slower tempi is close to two seconds. For Robert, a digit span of 2 (Table 7.1) may mean that he is unable to perceive and lay down a stable temporal template if the duration of one bar exceeds a certain threshold.

Evidence from the rhythm imitation task (Table 7.1: test 4) and rhythm entrainment task (Figure 7.1, Figure 7.2) also indicate that memory difficulties affect Robert's rhythmic production. Robert

reproduced just two beats in TRaCoL, which is consistent with his TAPS score (Table 7.1: test 3). The entrainment task required Robert to perceive and reproduce rhythmic patterns of five (Figure 7.1) and six (Figure 7.2) notes which is likely to have exceeded his digit span. With an ongoing visual demonstration, Robert copied the movements with some appropriate variations in rhythm (Figure 7.2). In combination with auditory stimuli, visual representation can support perception of complex rhythms (Grahn, 2012; Su, 2014). Additionally, visual information is known to assist motor performance in DS individuals (Bunn et al., 2007; Ringenbach et al., 2006; Sacks and Buckley, 2003). The visual mode may clarify the task expectations and may help him focus attention which is a pre-requisite for rhythmic motor entrainment (Repp, 2005).

Robert's improvement with visual support highlights a difficulty in linking the auditory signal to motor movement. Reduced perception of the auditory signal has been found in hearing adolescents who have variable tapping rates (Tierney and Kraus, 2013). Differences in how the brain responds to sound have been identified in people with DS (Groen et al., 2008; Pekkonen et al., 2007) and slow processing time for auditory signals has been demonstrated in DS children (Groen et al., 2008; Porter, Grantham, Ashmead and Tharpe, 2014). For Robert, any processing difficulties may be exacerbated by his bilateral hearing loss (Marcell et al., 1990; Marcell, 1995). Potentially, this could interfere with his performance in beat entrainment tasks, which requires integration of auditory signals with proprioception and motor output (Repp, 2006; Thaut, 2008). Thaut (2008) attributes negative AT (tapping after the beat), as was noted in Robert's playing at all tempi (Table 7.6), specifically to disturbances to the auditory and proprioceptive feedback that is required in order to correct timing for the following tap.

7.6.2 Temporal perception and production in speech

In word imitation tasks the number of sounds and syllables omitted (Table 7.3: PEO) or incorrectly imitated (Table 7.3: PPE, Table 7.10) indicate that Robert's digit span of 2 (Table 7.1) affected his speech production. It is difficult to isolate the effects of reduced STM from his hearing loss, but during the TAPS task Robert repeated numbers (e.g. the word *three*) in two-digit sequences that he failed to reproduce in sequences of three digits. This indicates that the test did successfully measure his memory, rather than his perception of words. Repeated listening and use of Makaton to support his memory for words allowed Robert to develop accuracy in producing words in songs (Figure 7.13, Table 7.12). Furthermore, Robert was inconsistent in his production of repeated words in new and familiar songs (e.g. Figure 7.13: *magic*, *book*) and in spontaneous speech (Figure 7.17: *very*, *good*). Severe hearing loss can affect consistency of production (Bowen, 2014; Northern and Downs, 2002), knowledge of speech sounds, and categorisation of sounds (Delage and Tuller, 2007; Northern and Downs, 2002). It is therefore likely that the majority of Robert's errors in speech production (Table 7.3) are the result of his hearing loss.

There is also a strong link between the severity of hearing loss and atypical rhythmic production of speech (Bowen, 2014; Markides, 1983; Osberger and McGarr, 1982). The rhythm of speech in hearing-impaired populations is affected by a high incidence of phonemic omission, slow production, pauses, and prolongation of speech segments (Bowen, 2014; Osberger and McGarr, 1982). It is likely that Robert's hearing loss contributes to his poor production of the syllabic rhythm: for example, he was inconsistent in the number of syllables he produced in the words *magic* and *book* (Figure 7.13); and in his spontaneous comments, he produced the word *very* as a monosyllable (Figure 7.17, Figure 7.18, Figure 7.19). However, Robert was able to imitate rhythms of two and three beats when the clapped rhythm accompanied a spoken word (Table 7.7: *monkey*, *can you clap*). Presenting a rhythm with a word, rather than in isolation, may assist production in people with learning disabilities (Jackson et al., 1997; Peters, 2000). Therefore, Robert may have been able to perceive the rhythm more clearly when it was spoken and clapped. His inability to produce more than three rhythms in the syllable segmentation tasks (Table 7.7) is likely to reflect a poor memory for the stimuli, rather than a motor difficulty or a failure to understand the task: he was able to produce one clap for *cat*, which demonstrates understanding, and he was able to produce and sustain rhythmic movements with visual support (Figure 7.2). The data from the Syllables task also indicate the effects of STM on his rhythmic production: with two exceptions (*lovingly*, *hopefully*: Figure 7.9), Robert produced just two syllables in words of three target syllables. It is likely that his hearing loss, in combination with his limited memory, leads to poor representations of syllabic rhythm.

7.6.3 Phonation and voice quality

Robert's voice qualities in speech tasks are consistent with those of people with hearing impairments. These include breathiness, nasality, reduced velopharyngeal control, reduced F0 variation, phonatory disturbances, and the presence of fluctuations in the amplitude of the waveform (Bolfan-Stosic and Simunjak, 2007; Coelho et al., 2015; Dehqan and Scherer, 2011; Monsen, 1979). All of these were observed in Robert's speech production.

Although his voice production was usually higher than those reported for DS adult males (Lee et al., 2009), his mean minimum F0 when describing pictures (Table 7.4) was closer to vocal fry range (Blomgren, Chen, Ng, and Gilbert, 1998; Childers and Lee, 1991; Fawcus, 2009; Mathieson, 2001). Robert's speech in individual tests was also at times perceived as creaky (e.g. Table 7.13). Vocal fry and creaky voice overlap somewhat (Keating et al., 2015) and are also common features in the speech of people with hearing impairments (Coelho et al., 2015; Das, Chatterjee and Kumar, 2013). The acoustic data also correspond to data from HI populations. In sustained vowels, Robert's mean HNR values (22.37 dB) and his mean shimmer values (3.85%: Table 7.5) are

similar to those in hearing-impaired populations (Coelho et al., 2015; Deqhan and Scherer, 2011). Furthermore, eight of the eleven sustained vowels (Table 7.8) and 63% (n=17) of the vowels in speech (Table 7.5, Table 7.10) were produced with additional interharmonics. Diplophonia at onset and medially, as observed in Robert's sustained vowels (Table 7.8), is also reported to result from hearing loss (Monsen, 1979; Das et al., 2013).

In people with HI, disturbed phonation can be attributed to reduced auditory feedback that leads to poor control of the vocal folds (Dehqan and Scherer, 2011; Fourcin and Abberton, 2008) and reduced regulation of sub-glottal pressure (Bolfan-Stosic and Simunjak, 2007; Monsen, 1979). Robert's productions show that he had difficulties in regulating sub-glottal pressure: for example, the intensity and formant contours of some sustained vowels were unstable towards the offset of vowels (e.g. Figure 7.3: section 4, Figure 7.4: sections 4-6, Figure 7.5). During these sections, the F1 contour was also unstable, which indicates a difficulty maintaining stable lingual postures for vowels. Changes in the volume of the vocal tract can directly affect sub-glottal pressure (Lucero et al., 2012; Stager, 2011; Tao and Jiang, 2008; Titze, 2008), and this may be affected by the position of the tongue (Gick et al., 2013; Honda, 1983; Honda et al., 1995; Stevens, 1997), with consequences for phonation (Honda et al., 1995). In vowel /a/1 (Figure 7.4, Figure 7.5), these changes in F1 also coincide with the onset of a subharmonic and audible diplophonia (section 4). Theoretically, at this point Robert may have attempted to regulate sub-glottal pressure as his volume dropped by moving his articulators: for example, a lowering of the tongue would lead to a lowered larynx and a reduction in the rate of loss of air pressure (Stevens, 1997). This will result in a higher F1 (Gick et al., 2012; Honda, 1983) as observed in Robert's /a/ vowel (Figure 7.4: section 5). However, loss of sub-glottal pressure destabilises phonation (Titze, 1994, 2008), and the maintenance of phonation therefore requires additional muscular adjustment to compensate for this (Gick et al., 2012). Detail of Robert's sustained /a/ vowel (Figure 7.5: sections 4-6) shows three instances of diplophonia that coincide with a raised F1 and a relatively stable intensity contour. Periods of phonation between the diplophonic sections are slightly breathy in quality, and the waveform and intensity contour show a concurrent drop in intensity. In order to sustain pitch, Robert may be compensating for loss of breath pressure and intensity with the physical activation of either his supralaryngeal articulators or the muscles within his larynx. Although Robert's voice was produced with continuous phonation in most tasks, any difficulties in co-ordinating voice and intensity will have an impact on his ability to use pitch and intensity to signal prosody.

Despite intermittent periods of atypical phonation, Robert's voice varied little in terms of its perceived vocal quality between tasks. However, comparison of the incidence and quality of interharmonics in words of one to three syllables (Figures 7.10-7.12) suggest that his production may have been affected by the demands associated with producing longer words, in line with the

other participants (Chapters 4-6). If so, the effect of this on mean perturbation measures is less clear (e.g. Table 7.11).

7.6.4 Pitch production and perception in speech and song

The results from the word imitation tasks and songs do not indicate a difficulty in perceiving pitch direction (Table 7.9, Figure 7.9, Figure 7.14, Figure 7.15) but Robert's imitation of words and songs was narrower than the demonstration in terms of pitch range (Table 7.9, Table 7.12). In comparison, Robert's spontaneous comments 1 and 3 (Table 7.14) and later productions of *Magic Book* and *Monkey Man* were produced with a wider pitch range and movement (Figure 7.14, Figure 7.15). This suggests that Robert can learn to reproduce melodic changes in speech, and that his reduced intonation in imitation tasks stems from a lack of opportunity to practice, rather than an inability to perceive melodic changes.

Robert's ability to replicate single pitches and intervals was inaccurate during his initial and final assessments (Figure 7.6, Figure 7.7). All but two notes (D4 and G5) were at least two semitones from target (Figure 7.6). The majority of Robert's pitching errors on singing tasks fall outside of the norms for TD adults: most nonsingers are accurate to within 1.3 semitones when matching pitches (Dalla Bella, Giguére and Peretz, 2007) and 60% of adults can reproduce a song they know well within two semitones (Levitin, 1994). According to the vocal sensorimotor loop (VSL) model, errors in singing can occur at three levels above physiological production: perception, sensorimotor mapping, or memory (Berkowska and Dalla Bella, 2009b; Dalla Bella et al., 2011; Dalla Bella et al., 2012). In order to reproduce a pitch, interval or song, the stimulus must be perceived and retained in working memory, the target must be mapped onto the articulatory and vocal system, and the output must be monitored and compared to the target (Dalla Bella et al., 2011). Robert demonstrates an ability to self-correct when pitch-matching (Figure 7.3: sections 2-3) and to improve in accuracy when reproducing relative intervals (Figure 7.7: *so-la-so*). His best reproduction of *so-la-so* shows that he can perceive changes of frequency to within one semitone and can monitor his output in order to sustain a pitch. Typical pitch perception is commonly reported in people with HI (Oxenham, 2008) and it is therefore unlikely that he has difficulty in perceiving the target frequencies, at least to within the minimum that is required to detect half a semitone (6%: Oxenham, 2008, 2013). Furthermore, poor perception of pitch change is not necessarily the cause of poor singing (e.g. Pfordresher and Brown, 2007; Pfordresher and Mantell, 2009; Berkowska and Dalla Bella, 2009a, 2009b; Dalla Bella et al., 2011, 2012). It therefore remains that his difficulties in pitch production in music and in speech lie within memory, auditory feedback, or sensorimotor domains.

In TD adults, memory for single pitch is retained for up to one minute (Krumhansl, 2000) but memory for intervals relies on working memory (Deutsch, 2013; Pfordresher and Brown, 2009; Williamson et al., 2010). In non-musicians, accuracy in reproducing intervals deteriorates as the number of target notes in a sequence increases (Dalla Bella et al., 2011; Pfordresher and Brown, 2007). Robert's poor auditory memory (Table 7.1) may therefore affect his performance in imitating intervals and melodies. Any difficulties in serial recall may also affect the order of his pitch changes: this could explain his transposition of notes during imitation of *me-do* and *so-me-do* (Figure 7.7). However, auditory feedback is necessary when learning to pitch in order to lay down sensorimotor traces for those pitches and intervals (Keough and Jones, 2009). In non-professional TD singers, the artificial generation of delays and distortions to auditory feedback result in transposed notes and rises in pitch (Keough and Jones, 2009). Examples of these behaviours in Robert's productions (Figure 7.7, Figure 7.3) may therefore indicate a difference in auditory processing, as is reported in people who have DS and hearing impairments (Marcell et al., 1990; Marcell, 1995). However, Marozeau et al. (2013) found that experience improved perception of melody, irrespective of the degree of hearing loss. Once a song is learned, motor-memory can take over and auditory feedback is less essential (Keogh and Jones, 2009).

7.7 Learning and potential links between musical and speech domains

7.7.1 Learning of speech sounds, word rhythm and pitch

Robert's performance in *Magic Book* (Figure 7.13, Figure 7.14) and *Monkey Man* (Figure 7.15, Figure 7.16, Table 7.13) show that with repetition and with visual support, Robert was able to reproduce more speech sounds, syllables and words. For example, in his later production of *Magic Book* (Figure 7.13: R3), he produced a voiceless velar fricative, [x], in two of his three productions of *magic.* This was an appropriate substitute for the target voiceless velar plosive, /k/, in terms of place. In learning *Monkey Man*, Robert also produced additional segmental content on each subsequent imitation (Table 7.13). Although his production of consonants did not develop between R3 and R4, he produced additional words (phrase 4) and syllables (phrase 3), and was able to place appropriate syllabic stress on the word *monkey* in his final production. It seems likely that with repeated listening, Robert was able to perceive additional segmental information, despite his hearing impairment. The results are comparable with previous research that shows that people with HI acquire new words slowly but make progress with repetition (Stelmachowicz et al., 2004). Stelmachowicz and colleagues also found that children with HI have poorer retention of words than non-HI peers. Furthermore, people with DS require novel phonological forms to be 'maintained' as a result of reduced STM capacity (Jarrold et al., 2009). For Robert, his reduced STM (Table 7.1) will interfere further with his reduced ability to learn new words. Repeated opportunity to listen will

help Robert form and maintain a more stable representation. Additionally, the slow, exaggerated production in song can support processing of speech (Wan et al., 2010) and might have further supported Robert's learning in this task.

Robert was able to clap the rhythm of familiar words (Table 7.7) and his rhythmic production of the words *magic book* increased with practice (Table 7.12). These indicate that he has an awareness of binary rhythm, and the conceptual understanding necessary to clap words of one and two syllables, but that his ability to imitate these is restricted by his poor auditory STM. Robert's performance in beat entrainment tasks confirm that his perception may be limited by the duration of the auditory stimulus. This may explain his unstable motor performance, but may also contribute further to his poor imitation of speech sounds in imitation tasks. The data indicate a reduced ability to process words, speech sounds, and rhythm, rather than an inability. The quality of his processing of temporal stimuli in music and speech domains may both be affected by his hearing impairment (Marcell et al., 1990; Marcell, 1995).

In learning songs, Robert developed quickly in musical accuracy, as his segmental accuracy also developed. Over two weeks, Robert's segmental production in *Magic Book* was produced more quickly and with improved relative vowel:consonant duration (Table 7.12), as well as with improved pitching accuracy. In learning *Monkey Man* (Table 7.13), Robert produced the repeating intervals accurately in his second week (Figure 7.14). It is possible that once he had mastered segmental content, he was better able to reproduce the melodic and rhythmic structure.

In most occasional singers, reduction of lyrical content to a vowel results in improved musical performance, which is attributed to a reduction in memory load (Dalla Bella et al., 2012). Several studies have reported that augmenting memory during imitation tasks can facilitate output, which is taken as evidence of memory deficit (Dalla Bella et al., 2012; Wise and Sloboda, 2008). Robert's pitch production in *Magic Book* and *Monkey Man* was most accurate when supported with pitch cues (Figure 7.14: R2, Figure 7.15: R3) which indicates that Robert had a weak memory trace of the pitch details. For Robert, reduced memory for auditory stimuli and reduced acuity during learning may have resulted in poor memory traces and made learning slow and imprecise.

However, the ability to extract, recognise and categorise salient features in an auditory stream does not just depend on acuity of perception or immediate memory capacity. According to Snyder (2000), it also depends upon previous exposure which creates a 'semiactivated' or unconscious memory trace within the brain. Snyder (2000) also states that further exposure to similar events re-activates the subconscious memory traces and 'at some point, they move into the focus of conscious awareness'. It is possible that Robert had limited previous experience in listening to and

repeating pitches and melodic intervals in isolation: in this case, his relatively rapid pace of development (Figure 7.6, Figure 7.7, Figure 7.14, Figure 7.15) may reflect the process of 'activation' within the brain (Snyder, 2000). This is comparable to the VSL model's 'covert' or 'overt' production of pitch (Berkowska and Dalla Bella, 2009b; Dalla Bella et al., 2011, 2012). According to the model, production may be an automatic response or it may be deliberate, but learning is most efficient when attention is conscious and when errors are noticed and corrected. With conscious learning, basic musical skills can be improved in 4-8 hours in TD populations (Oxenham, 2013). The changes observed in Robert's performances may therefore reflect heightened awareness as a result of experience, coupled with the use of a visual medium to focus his attention on one salient aspect at a time. Direction of his attention to one feature at a time would allow conscious, focussed learning and also reduce memory load. This has implications for the teaching of songs, but also implications for long-term learning of words in songs and, potentially, for learning words for speech.

7.7.2 Happy Birthday: long-term learning and transfer between song and speech

Robert's production of *Happy Birthday* in spoken and sung conditions was limited in pitch range and movement, but the direction of his pitch changes indicate recall and transfer of much of the melodic contour (Table 7.14). Recent research with TD adults has shown that the pitch contour of a song tends to be preserved in the spoken production unless the target production is markedly different from the stimulus in contour (Wisniewski et al., 2013). However, Robert's production of the song's intervals was limited in scope and the direction of some intervals was incorrect (Figure 7.7: *me-do*, *so-me-do*). This suggests that Robert had retained the global melodic features but had not processed or recalled the interval details. As discussed above, if the song had not been explicitly taught, Robert's memory for all details may be weak.

Robert required verbal prompts for the words in the spoken condition of *Happy Birthday* (Figure 7.23). Following this task, he also required prompts to sing the song. He had independently recalled more of the text in an earlier production of the song (Figure 7.23: Singing 1). There is evidence that words in songs are recalled serially and that the metrical structure can cue words (Racette and Peretz, 2007): so if the first line is not recalled, the second line will be more difficult to access. As a result of task demands associated with transferring songs to speech (Racette and Peretz, 2007; Wisniewski et al., 2013), it may have been especially difficult for Robert to transfer to the spoken mode. This may explain his poor recall in the spoken production, in comparison with his earlier sung version (Figure 7.23: Singing 1). His poor spoken production may in turn have affected his confidence for subsequently singing the song (Singing 2), which was reduced in pitch range (Table 7.14). A reduction in pitch range, rhythmic variability and volume are associated with uncertainty of knowledge in adults (Pon-Barry, 2008; Pon-Barry and Shieber, 2011). Tremolo, as

observed in Robert's sustained vowel /a/1, can also indicate uncertainty or nervousness (Baker and Lane, 2008). In contrast, Robert's 'best' productions in terms of prosody were the phrases *very good* (Figure 7.18) *oh yeah* (Figure 7.21) and the words *monkey* and *man* (Table 7.13, Figure 7.16). These examples were produced with changes in relative volume and pitch and with relatively more volume than other parts of the phrase (e.g. Figure 7.18). The louder, rhythmic production of the words *monkey* and *man* (Figure 7.16), prompted by Makaton, indicate that his confidence may be supported when his memory is supported, with consequences for his prosodic production.

7.8 Summary and implications

It is likely that Robert's hearing impairment is the primary reason for his poor production of speech sounds and rhythmic production of words. However, Robert's poor auditory memory affects the amount of detail he can reproduce in all imitation tasks. Although he was able to improve in the accuracy of his timing of motor-rhythms with visual support, his movements in this and in beat entrainment tasks were poorly timed. Difficulty in entraining at slower tempi indicate the potential effects of memory limitations on his ability to form a mental representation, but do not rule out an underlying difficulty in processing temporal information, which may be exacerbated by his hearing impairment. Reduced auditory feedback is also implicated in errors from interval-matching tasks, pitch-matching tasks, and phonatory control. However, his oro-motor limitations also affect his production in speech and song, and it is not know to what extent his gross motor abilities affect his production in musical-motor tasks.

Auditory-motor control and rhythmic music-motor production both depend primarily on opportunities to practise musical skills (Bispham, 2006; Corrigall and Schellenberg, 2016; Hargreaves, 1986; Reifinger, 2006; Lamont, 2009). Robert's inaccurate performance of *Happy Birthday* and his initial inability to match pitch or intervals are likely to reflect a previous lack of conscious learning, rather than an inability to perceive or learn. He demonstrates an ability to improve rapidly when perceptual/hearing and memory limitations are supported, and when teaching is focussed on specific elements. Despite his cognitive impairment, his results are comparable to similar programmes of music and singing with TD hearing impaired children that have shown improvements in aspects of auditory perception and production (Rochette, Moussard and Bigand, 2014; Welch et al., 2015). For Robert, there may be several benefits in developing his musical and singing skills.

Poor singing is associated with poor rhythmic skills (Dalla Bella and Peretz, 2003; Dalla Bella et al., 2015); likewise, the ability to perceive rhythm is strongly linked to speech perception (Hausen, Torppa, Salmela, Vainio, and Särkämö, 2013). For Robert, teaching should focus primarily on his rhythmic skills and initially on the development of whole-body movements to the beat. Robert's

movements are poorly co-ordinated (see Section 7.1.2) and there is evidence that how an adult moves in time to music has direct consequences for their perception of the beat (Phillips-Silver and Trainor, 2007). Robert may therefore benefit from whole-body rhythm activities, as advocated by Kodaly (Houlahan and Tacka, 2008) and Dalcroze (Findlay, 1971) approaches. Developments would be beneficial at the whole-body stage, initially, to support his awareness of beat and rhythm; to engage his vestibular system (Phillips-Silver and Trainor, 2008) and also to support development of his motor co-ordination. Given his hearing impairment, visual support or other multisensory support (e.g. tactile feedback, such as through a resonance board) may support his perception of auditory beat and rhythm.

Robert's data from the beat timing tasks indicate the effects of reduced auditory processing. A growing body of research links music training with enhanced auditory processing of pitch for music and language (Schön et al., 2008; Besson et al., 2007; Magné et al., 2006; Parbery-Clark, Anderson, Hittner and Kraus, 2012; Tierney and Kraus, 2013) and for rhythmic structure of words (Jakobson, Cuddy and Kilgour, 2003; Marie et al., 2011; Hausen et al., 2013). Recent evidence suggests that musical training may account for improvements in auditory perception (Parbery-Clark, Anderson and Kraus, 2013) and for improved phonetic discrimination in children with hearing impairments (Rochette et al., 2014). In people with HI, musical experience enhances rhythmic tapping performance in adults (Matsubara, Terasawa and Hiraga, 2014). Tierney and Kraus (2013) argue that practice in beat timing may improve the brain's response to processing auditory signals. Musical training may therefore enhance Robert's perception of sound, and may also benefit his speech perception and production. For example, learning words in the context of a song may enrich processing (Racette and Peretz, 2007); the slower production of speech sounds in song can lead to enhanced perception of the sounds (Kilgour, Jakobson and Cuddy, 2000; Schlaug et al., 2008); and the synchronisation of a metrical beat with linguistic stress can help the processing of lyrics (Gordon et al., 2010). A focus on speech rhythms may support Robert's perception and production of speech sounds: rhythmic priming can enhance the speed of phonemic processing in adults (Cason and Schön, 2012; Cason, Astésano and Schön, 2015) and can improve phonological production in children with moderate to profound hearing impairment (Cason, Hidalgo, Roman and Schön, 2014). Such activities may directly assist coding in memory if segmental content is made the focus of Robert's attention. Slower production in singing also assists accuracy in output (Keough and Jones, 2009). Robert's production of speech sounds in *Magic Book* improved with practice and with repetition. Although repetition alone will support Robert's ability to learn words (Jarrold et al., 2009; Stelmachowicz et al., 2004), slow, accurate singing with visual support will reduce the processing load. This may help Robert to develop stronger, more accurate memory traces for words and pitches, even with his hearing loss.

Memory plays a crucial role in rhythm perception (Grahn and Schuit., 2012; Dalla Bella et al., 2012), in pitch perception (Dalla Bella et al., 2012) and in singing (Pfordresher and Mantell, 2009). The data indicate that Robert's digit span affects his performance, so in order for him to benefit from musical training, his auditory memory would need to be supported for rhythm, pitch and words. A reduction in memory load during teaching would support his learning and may also support his performance in singing tasks. In particular, reducing lyrical content will support vocal accuracy in pitch (Welch, 2009; Dalla Bella et al., 2012). There are associations between musical experience and working memory (Chan, Ho and Cheung, 1998; Lee, Lu and Ko, 2007; Parbery-Clark et al., 2009) and specifically with auditory working memory (Kraus et al., 2012; Schulze, Zysset, Mueller, Friederici and Koelsch, 2011). Robert would therefore benefit from a multisensory approach to support his reduced auditory STM.

In sum, a range of musical activities may be appropriate to support Robert's ability to learn and produce words, but teaching would need to address his current levels of perception, hearing and motor co-ordination. Robert demonstrated rapid learning in songs and his response to visual and gestural instruction indicates increased awareness of target pitch for single notes and some interval shapes, as well as for rhythm. In the present study, this technique was used regularly to support conceptual understanding, but the use of visual feedback to correct performance was also trialled (Appendix 3, p.347 [19,]; p350 [25-31]. For poor pitch singers, such approaches are recommended to support the singer's understanding of their own performance in relation to the intended goal, and to enable them to adjust performance (Welch, 1985). This may be especially important for Robert, whose internal methods of feedback (auditory and proprioceptive) may be impaired as a result of both his HI and of the syndrome. Additionally, the research evidence indicates that movement-based teaching may support his musical perception. Therefore, a programme that integrated whole-body rhythmic motor movements and visually-supported singing techniques would be recommended.

Chapter 8: Summary and conclusions

This study aimed to discover whether rhythmic and song-based musical activities have the potential to support or develop the production of speech and voice in adults with DS and SLD. Previous research has identified correlations between musical abilities and cognitive processes that support speech, and has indicated that skills may transfer between domains, especially if certain conditions are met during learning (Patel, 2011). People with DS have well-documented difficulties in speech production and some perceptual difficulties. They are historically reported to have a specific fondness for music (Lense and Dykens, 2011), which means that music may be a useful tool for developing or supporting skills required for speech. However, research into the musical and song-based abilities of people with DS is limited to a handful of studies (Edenfield and Hughes, 1991; Lense and Dykens, 2011; Ringenbach et al., 2006; Stratford and Ching, 1983; Stratford and Ching, 1989). Therefore, this study sought to discover to what extent a group of four young adults with DS, SLD and unspecified HI were able to produce rhythmic motor-movements and to develop skills in singing over a six-week programme of song-based music activities.

The case studies (Chapters 4-7) addressed the research questions in terms of individual abilities. However, in order to address the research aims it is also necessary to consider the abilities of the group as a whole. This chapter will directly address the research questions given in Chapter 2 and below, in order to consider the applicability of the findings to the wider DS population:

1. What are the verbal and non-verbal abilities of the individuals in relation to each other and to the DS population as a whole? (Section 8.1);
2. What factors affect the individual's and group's musical and speech skills in terms of rhythm, melodic intonation and phonation? (Section 8.2);
3. To what extent can the individuals and the group learn to improve their musical ability in terms of perception and production of musical features through song-based teaching? (Section 8.2); and
4. For people with DS, what is the potential for rhythmic and song-based music tuition to develop aspects of auditory processing, voice and speech production? (Section 8.3)

This chapter specifically addresses the hypotheses posed in Chapter 2:

1. For adults with DS, musical development will be in line with their overall non-verbal cognitive development and vocal abilities.;
2. Given the shared physiology between singing and speech and correlation between certain speech and musical abilities in terms of perception and production, difficulties that affect the speech domain will correspond to difficulties in the musical domain.;
3. Adults with DS and SLD are capable of learning songs and of developing their melodic-rhythmic skills in singing. In line with other populations with developmental disabilities,

individuals with DS can develop in both awareness and production of musical features when these skills are explicitly taught and when teaching is tailored to strengths.; and

4. If the group is able to transfer learning between domains, and if the appropriate functional, emotional and cognitive conditions are met, then the development of musical skills may be employed to develop aspects of speech perception and production in adults and children with DS, of all abilities.

The section begins with a summary of the empirical findings and a profile of the group's abilities and difficulties (Section 8.1). Section 8.2 discusses the findings of the group as a whole and addresses the research questions and hypotheses, in order. Section 8.3 considers the implication of these findings for the wider DS population. It recommends a programme of musical activities that may support the learning needs that were found to be common to the group in this study, and which are common to the DS phenotype. Section 8.4 presents a critical evaluation of the study and makes recommendations for future research. The chapter concludes with an outline of the contribution of this study to the research area.

8.1 Summary of the key findings

This main findings were presented and discussed in the case studies (Chapters 4-7). This section will summarise the key findings and themes that are common to the group in this study, in terms of: their performance in core tests; key difficulties in tasks; phonation; hearing, processing and attention; memory demands and cognitive load; learning abilities; and transfer of words from song to speech. It will also summarise any differences that were observed in speech and musical domains[1].

8.1.1 Key abilities (Table 8.1)

Table 8.1 summarises the results from the assessments that were conducted before teaching sessions commenced, allowing comparison between individuals. Participants are ordered left to right according to their BPVS MA-equivalent score (1). The table shows a low spread of scores for BPVS. A difference in overall ability is evident between Robert and the other three participants for TAPS (2), DAP (3), DEAP (4), CNREP (5) and all music tasks (6-12). Of the group, Andrew scored the highest on TAPS (2) and in phonemic accuracy (4, 5), and his performance in TAPS suggested a digit span of between 3 and 4 (see Chapter 4). Vocally, Rachel and Robert produced wider vocal pitch ranges on vowel glides (13, 14). When speaking, their voices were also less harsh in quality (17) than Andrew's or Kerry's. The voice qualities of Andrew and Kerry were less harsh when singing (18) than when speaking (17). All but Robert produced some dysfluent speech (19). When imitating words of two and three syllables, Rachel and Robert were able to reproduce the

1 The numbers in brackets in the text refer to the task or item number in each table.

demonstrated contrasts in the relative duration of most syllables (nPVI: test 20). There were no data available for Andrew or Kerry as a result of the high degree of voice perturbation in the task (see Chapter 3, and Chapters 4 and 5). Andrew was the most able to reproduce changes in intensity contours, and Robert was the least able (21). Robert was also the least able to reproduce pitch contours (22).

Table 8.1: Group profile: a summary of data from key tasks from the battery of assessments (pre-teaching). The table shows individual scores and indicative levels of ability that were measured during initial assessments, and the group mean for applicable tasks.

		Subject/ Test	Andrew	Kerry	Rachel	Robert	Mean
Cognitive ability	**1**	BPVS: Verbal MA equivalent	3;4 (2;11 – 3;9)	3;1 (2;9 -3;7)	3;0 (2;8 -3;5)	3;0 (2;8 -3;5)	3.1
	2	Digit span (TAPS)	3+	3	3	2	2.75
	3	Draw–a-person (DAP)	5 years	7-3 years	7-6 years	4-3 years	6.05
Phonemic accuracy	**4**	DEAP: PPC	78.26%	62.11%	69.32%	40.78%	62.62%
	5	CNREP: PPC	48.94%	28.61%	30.50%	no data (15.90% PPC in 2 syll words)	36.02%
Musical ability	**6**	Mean Music Age equivalent	**6.2**	**6.3**	**6.2**	**4.5**	**5.8**
	7	*Gross motor: march to beat*	5-7	5-7	5-7	<4	
	8	*Pulse & meter*	7-9	7-9	7-9	<5	
	9	*Rhythm*	5-7	5-7	5	<4	
	10	*Melodic shape*	<5 (but dysphonic)	5-7	5- 7	<5	
	11	*Tonality*	n/a	7	7	3-5	
	12	*Small intervals*	n/a	5	5	<5	
Voice	**13**	F0 range on vowel glides, Hz	108.94 - 238.84	221.20 - 324.10	152.40 - 472.40	163.09-421.82	
	14	Pitch range on vowel glides, (semitones)	13.59	6.61	19.58	16.45	14.06
	15	Speaking range; Mean F0 Hz (describing a picture)	53.71 - 221.21; 118.57	206.08 - 293.27; 245.55	110.31 - 295.32; 215.12	158.09 - 197.37; 172.93	216.97 (m); 230.34 (f)
	16	SD of speech from mean (describing a picture): Hz; semitones	30.17; 4.43	19.52; 1.37	37.69; 13.04	9.75; 0.97	24.28; 4.95
Perceived voice quality	**17**	Speaking	mostly harsh/ tense/ventricular	rough/whispery creak	partly rough, partly modal	high register, mostly breathy	
	18	Singing	variable: falsetto voice is more modal	mostly modal	mostly modal	high register; mostly breathy	
	19	Fluency	some dysfluency in spontaneous speech	mostly dysfluent	rare dysfluency, some cluttering and festination	fluent	
Prosodic features	**20**	Reproduction of variability of vowel duration (nPVI) in words (1-3 syllables)	n/k: lack of measurable voicing	n/k: lack of measurable voicing	87%	87%	
	21	Accuracy in reproducing intensity changes in words (2-3 syllables)	81%	70%	61%	22%	
	22	Accuracy in reproducing Pitch contour in words (1-3 syllables)	85%	89%	89%	78%	

8.1.2 Key difficulties (Table 8.2)

All participants had difficulties in speech and in song in terms of: articulation of speech sounds (3); producing and sustaining modal phonation (4); reproducing the demonstrated number of speech sounds, irrespective of the accuracy of the reproduced sounds (12); and reproducing words consistently (18). Andrew and Kerry had difficulties in producing controlled changes in intensity in speech and song (2). In sustained vowels, all four had difficulties in sustaining phonation (5) and no participant was able to sustain phonation for a mean duration of more than two seconds (5). The spectrograms indicated that all but Andrew were unable to maintain a stable articulatory configuration when sustaining some vowels, as indicated by variations in the contour of the first formant (6).

Robert was the only participant with confirmed hearing loss (1: and see Chapter 7). Kerry's responses to some tasks of *what?* and *what's that?* (Chapter 5) could indicate inadequate hearing. However, these could also reflect attentional difficulties (14-16), or difficulties in auditory processing (19). Rachel (Chapter 6) responded to all task instructions, but her unconfirmed hearing loss (1) may explain her reduced ability to discriminate speech sounds in the minimal pairs test (8, 9) and in imitation tasks (11, 12). Andrew's behaviour and performance in tasks did not suggest hearing loss and he was able to perceive speech sounds (8, 9) and to recall words in speech and song, *Magic Book* (11). However, his recall of the speech sounds in this task was comparable to that of the other participants (12). In speech and song, the group had difficulties in imitating speech sounds (12), weakly-stressed syllables in words (10), and words in phrases (11).

All participants demonstrated some difficulties in attending to some tasks (14-16). Andrew and Robert did not listen to demonstrations before imitating them (14), and all displayed some difficulties in sustaining attention to some tasks and inhibiting distractions (15). Andrew and Rachel both avoided taking part in some tasks (16). In speech and song, Kerry, Rachel and Robert appeared better able to imitate one aspect of speech and song at a time: either segmental aspects, or rhythmic and melodic. Andrew did not seem to have this difficulty, but his melodic production was limited in comparison to the other participants (see Table 8.1: 10-12). Andrew and Kerry both seemed to process phrases in speech and song at a semantic level: they were able to substitute nouns and pronouns in phrases and songs, and synonyms when reading (13). Rachel and Robert did not demonstrate this behaviour, and Rachel substituted the word *what* with *why* in the song *Magic Book*, indicating a failure to process the meaning of the phrase (see Chapter 6).

The data reveal a few differences in observed behaviour and difficulties between speech and song. None of the participants were dysfluent when singing (7). Andrew, Kerry and Rachel recalled a higher percentage of words in songs than in spoken phrases (11), but all but Robert produced

Table 8.2: Group profile of key physiological, perceptual and behavioural difficulties demonstrated in speech and music-related tasks. The presence or absence of observed difficulties is indicated by Y (yes) and N (no); a suspected difficulty is indicated by (?). Additional information and scores are given, where data are available. Red type is used to allow easy comparison between individual abilities.

				Andrew		Kerry		Rachel		Robert	
			Domain:	Speech	Song	Speech	Song	Speech	Song	Speech	Song
Area of difficulty		**Difficulty**	**Task/Source**	Difficulties indicated (Y/N) and level of ability if known: (% correct)							
Physiological	**1**	Hearing loss	*Behaviour/SaLT report*	N: 'adequate for speech, except whisper'		Y: responded with 'what'?		?: 'unconfirmed hearing loss'		Y: moderate bilateral hearing loss	
	2	Difficulty in controlling co-ordination of voice and intensity for prosody *(All)*		Y	Y	Y	Y	N	N	N	N
	3	Difficulty in articulating speech sounds	*DEAP*	Y	Y	Y	Y	Y	Y	Y	Y
	4	Difficulty producing efficient/modal phonation	*All tasks*	Y	Y	Y	Y	Y	Y	Y	Y
	5	Difficulty in sustaining phonation (MPT)	*Sustained vowels*		Y: 1.121 s		Y: 0.580s		Y: 1.191s		Y: 1.472s
	6	Difficulty in maintaining steady F1	*Sustained vowels*		N		Y		Y		Y
	7	Evidence of dysfluency	*Spontaneous speech/ song*	Y	N	Y	N	Y	N	N	N
Perceptual	**8**	Difficulty in discriminating consonants	*Minimal pairs*	N: 100%		Y: 40%		Y: 80%		Y: 60%	
	9	Difficulty in discriminating vowels	*Minimal pairs*	N: 100%		Y: 80%		Y: 20%		Y: 80%	
	10	Difficulty in discriminating weak syllables	*CNREP/Magic Book 1*	Y	N	Y	Y	Y	Y	no data	Y
Cognitive	**11**	Difficulty in recalling words	*Imitated phrases / Magic Book 1*	Y: 84%	N	Y: 65%	Y: 70%	Y: 33%	Y: 50%	no data	Y: 30%
	12	Proportion of consonants recalled in imitated words	*Syllables/Magic Book 1, phrases 1-4 (based on number, not accuracy)*	Y: 88%	Y: 41%	Y: 71%	Y: 54%	Y: 67%	Y: 32%	Y: 11.9%	Y: 22%
	13	Difficulty in processing meaning of words in phrases	*Phrase imitation/ songs*	N	N	N	N	Y	Y	no data	Y
Attention and EF	**14**	Difficulty in listening to demonstrations		Y	Y	N	N	N	N	Y	Y
	15	Difficulty attending fully to tasks		Y	Y	?	N	Y	Y	Y	Y
	16	Evidence of task avoidance		Y	Y	N	N	Y	Y	N	N
Implicated difficulties	**17**	Difficulty dividing attention between segmental and melodic features of phrases	*Imitated phrases/ Magic Book*	N	N	Y	Y	Y	Y	no data	Y
	18	Poor representation of words	*repeated words (songs, spontaneous speech, reading*	Y	Y	Y	Y	Y	Y	Y	Y
	19	Slow or impaired auditory processing	*learning of speech sounds in song; slow beat entrainment*		Y		Y		Y		Y

fewer consonants when imitating sung phrases (12). Andrew perceived the weakly-placed syllable in the song *Magic Book* but not in CNREP (10)

Voice and phonation

Sustained vowels (Figure 8.1)

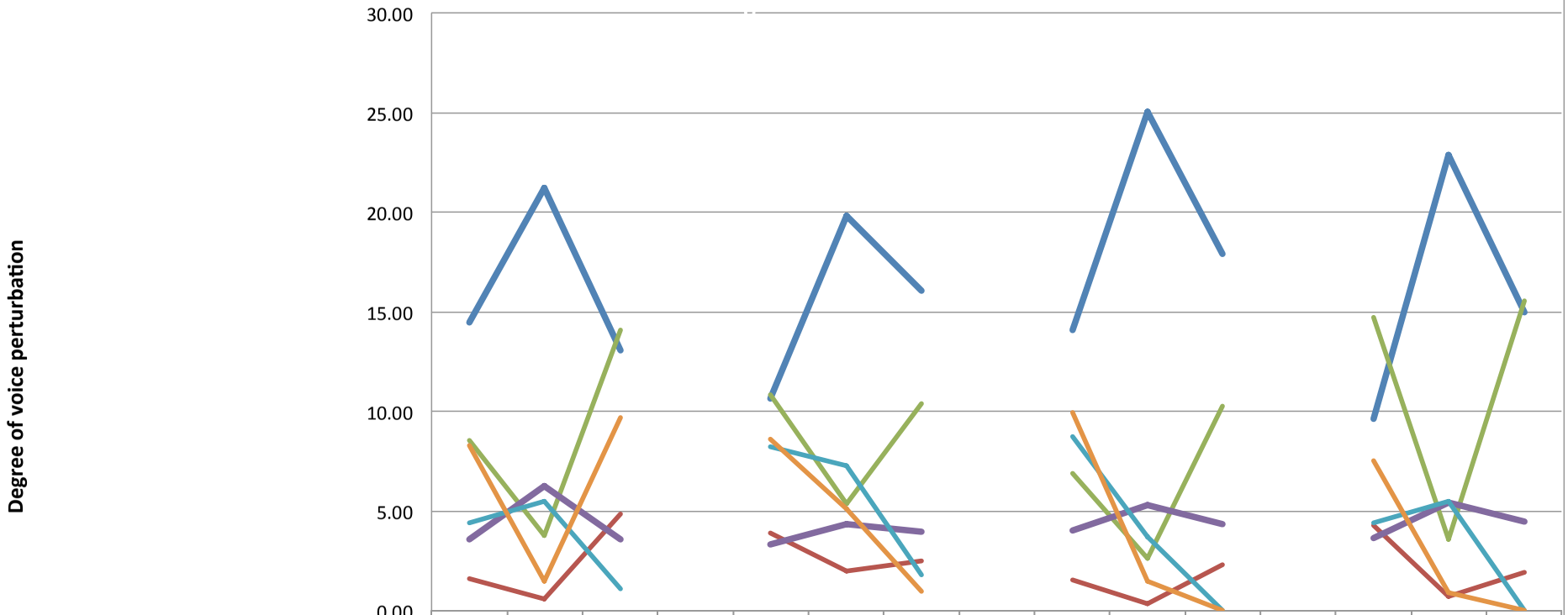
All participants produced a higher degree of perturbation at onset and offset of sustained vowels (Figure 8.1: all measures). Mean HNR (dark blue contour) and CPP (purple contour) values were higher for Rachel and Robert than for Andrew and Kerry. There was no clear trend for jitter (red) or shimmer (green). Of the group, Kerry produced the highest number of vowels with interharmonics (pale blue contour), and Robert produced the fewest. The duration of interharmonics (orange contour) produced at onset was on average at least 75% of the onset (100 ms) in all measurable vowels (see Chapters 4-7). Kerry produced a higher number of vowels with interharmonics in the medial portion, and these were the highest of the group in terms of vowel duration (orange). Only Andrew and Kerry produced interharmonics at offset. For Andrew, the duration of interharmonics was greater at offset than at onset.

Increasing segmental demand (Figure 8.2)

In comparison to mean values for whole sustained vowels (including onset and offset), vowels in words were produced with a lower mean HNR for all participants (Figure 8.2). However, Robert produced lower jitter values (red contour) for vowels in words, and he and Rachel produced lower shimmer values (green contour) than for sustained vowels. Some parameters changed as the target length of syllables increased (Figure 8.2): mean HNR (dark blue contour) increased in longer words for all but Andrew; the incidence of vowels produced with interharmonics (pale blue contour) increased with syllable length for all participants but Rachel; and mean DUV (grey contour) and shimmer (green contour) values increased in multisyllabic words for Andrew and Kerry. Mean CPP values (purple contour) were less variable. For Kerry and Rachel, a marginal improvement was observed in words of two and three syllables for CPP and HNR: the possible explanations for these were discussed in the case studies (Chapter 5: Section 5.5.4; Chapter 6: Section 6.6.1).

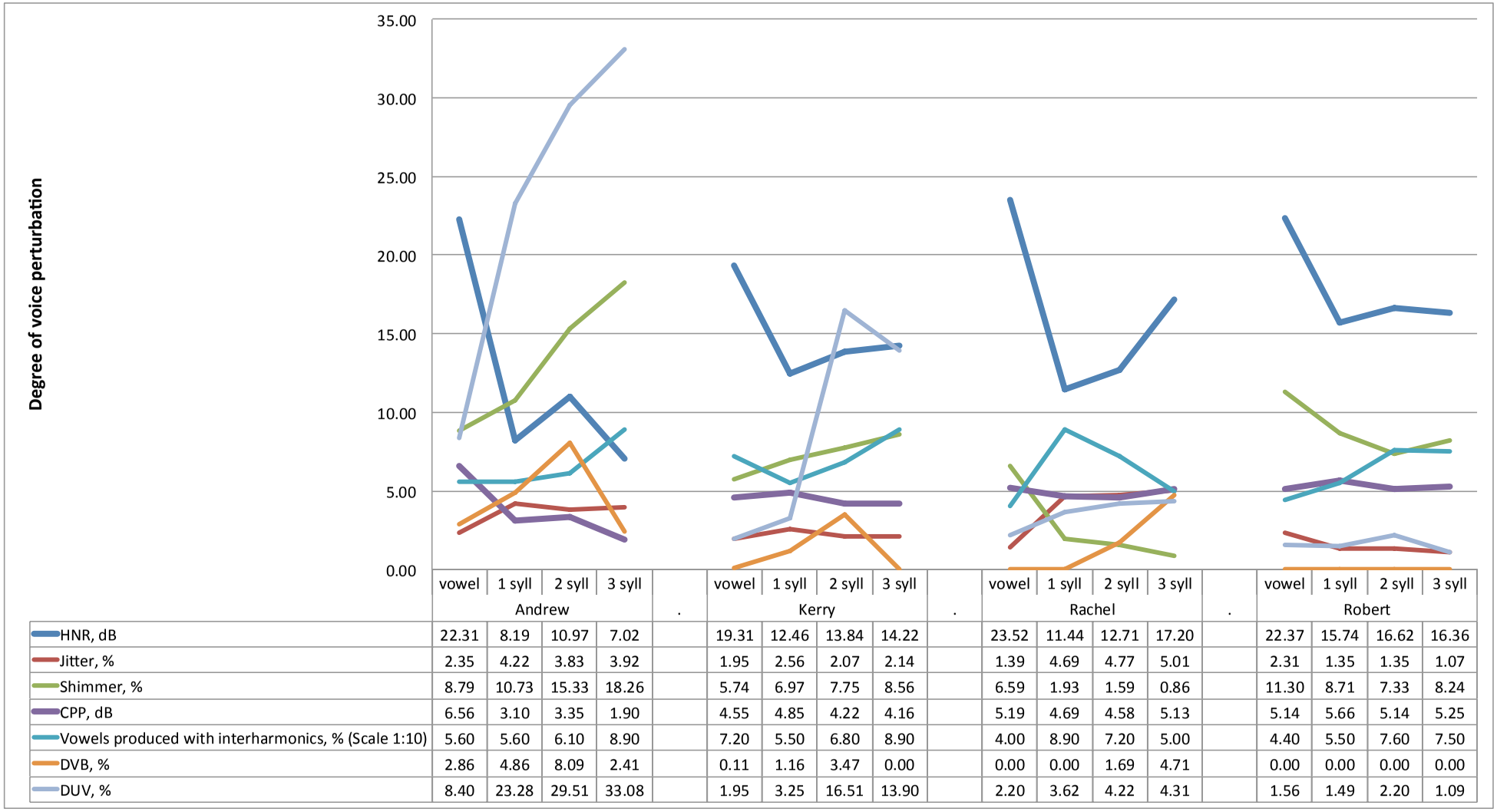
Memory demand and processing load (Table 8.3)

Table 8.3 shows the level at which speech, song and music imitation tasks appeared to be affected by auditory memory demands (1-11). It also specifies which observed behaviours in individuals may be related to the memory demand inherent in tasks (12-17). Of the group, in imitation tasks Andrew reproduced a higher number of syllables in words (1) and nonwords (2), and a higher number of word-based rhythms (4). Andrew and Kerry were similar in their abilities to clap word-based syllabic rhythms (6), and to reproduce pitch contours in nonwords (8). Kerry was the only group member to partially or correctly imitate pitch contours in phrase imitation tasks (9). However, she failed to reproduce the correct pitch change in melodic intervals of three notes (10).



	Andrew				Kerry				Rachel				Robert		
	Onset	Mid	Offset	.	Onset	Mid	Offset	.	Onset	Mid	Offset	.	Onset	Mid	Offset
HNR, dB	14.45	21.23	13.10		10.65	19.81	16.04		14.07	25.07	17.91		9.61	22.89	14.95
Jitter, %	1.59	0.60	4.86		3.90	1.96	2.52		1.52	0.35	2.30		4.26	0.74	1.93
Shimmer, %	8.52	3.77	14.09		10.81	5.39	10.41		6.88	2.62	10.28		14.75	3.61	15.54
CPP	3.61	6.24	3.59		3.30	4.36	3.95		4.01	5.29	4.33		3.66	5.41	4.50
Vowels produced with Interharmonics, % (Scale 1:10)	4.40	5.50	1.10		8.20	7.30	1.80		8.75	3.70	0.00		4.40	5.50	0.00
Duration of interharmonics, % (Scale 1:10)	8.30	1.48	9.70		8.64	5.14	0.95		9.95	1.49	0.00		7.50	0.91	0.00

Figure 8.1: A graph showing the mean perturbation measures produced by individuals in sustained vowels. The image and table show mean values of HNR (dB), Jitter (%), Shimmer (%) and CPP (dB). Data for vowels are given for the onset (t=100ms), offset (t=100 ms) and the middle section (t= total vowel duration - (onset+offset).The percentage of vowels that were produced with interharmonics at onset, medially and at offset is also shown, but is scaled to a factor of 1:10. The duration of interharmonics is also shown, expressed as a percentage of the duration of onset (100 ms), medially (T-(onset + offset)) and at offset (t=100 ms), scaled to a factor of 1:10.



	vowel	1 syll	2 syll	3 syll	.	vowel	1 syll	2 syll	3 syll	.	vowel	1 syll	2 syll	3 syll	.	vowel	1 syll	2 syll	3 syll
	Andrew					Kerry					Rachel					Robert			
HNR, dB	22.31	8.19	10.97	7.02		19.31	12.46	13.84	14.22		23.52	11.44	12.71	17.20		22.37	15.74	16.62	16.36
Jitter, %	2.35	4.22	3.83	3.92		1.95	2.56	2.07	2.14		1.39	4.69	4.77	5.01		2.31	1.35	1.35	1.07
Shimmer, %	8.79	10.73	15.33	18.26		5.74	6.97	7.75	8.56		6.59	1.93	1.59	0.86		11.30	8.71	7.33	8.24
CPP, dB	6.56	3.10	3.35	1.90		4.55	4.85	4.22	4.16		5.19	4.69	4.58	5.13		5.14	5.66	5.14	5.25
Vowels produced with interharmonics, % (Scale 1:10)	5.60	5.60	6.10	8.90		7.20	5.50	6.80	8.90		4.00	8.90	7.20	5.00		4.40	5.50	7.60	7.50
DVB, %	2.86	4.86	8.09	2.41		0.11	1.16	3.47	0.00		0.00	0.00	1.69	4.71		0.00	0.00	0.00	0.00
DUV, %	8.40	23.28	29.51	33.08		1.95	3.25	16.51	13.90		2.20	3.62	4.22	4.31		1.56	1.49	2.20	1.09

Figure 8.2: A graph showing the mean voice perturbation measures for whole sustained vowels and vowels in words of 1, 2 and 3 target syllables. The image shows the mean values of, HNR (dB), CPP (dB), jitter (%) shimmer (%), DVB (%), DUV (%) and the percentage of vowels in each task that were produced with one or more interharmonics (%, scaled 1:10) in whole sustained vowels (including onset and offset) and in words of increasing target syllable length (1-3). The data show a mean increase from left (Andrew) to right (Robert) in HNR (dark blue contour), and a decrease in DVB (%) (orange contour) and DUV % (grey contour). Most participants show an increase in one or more measure of perturbation as the target syllable length increased.

The case studies showed that participants displayed behaviours associated with stress in imitation tasks that were absent in tasks relying on speech production from long-term memory. Rachel (Chapter 6) and Robert (Chapter 7) were better able to perform musical tasks when visual aids were used, and Robert's vocal performance appeared more confident after repetition and when supported by Makaton. For Kerry, Andrew and Rachel, an increase in cognitive load was also linked to a rise in the incidence of dysfluency (14). A rise in interharmonics was observed for Andrew, Kerry and Robert in multisyllabic words (Figure 8.2) and in tasks of increasing demand (15). In tasks that placed greater demands on phonological memory or on linguistic processing, Andrew (Chapter 4) and Kerry (Chapter 5) appeared to use a greater degree of muscle tension in their vocal production, and the incidence of Rachel's vocal expressions of uncertainty increased.

The total number of tasks that were *not* presumed to be affected by memory demand suggest that Andrew's performance was least affected by cognitive load and that Rachel's was the most affected (Table 8.3: 'number of tasks presumed to be unaffected'). Kerry demonstrated an improved retention of rhythm when words were part of a song (3), and Andrew and Rachel reproduced a higher number of syllables in imitated sung phrases than in speech imitation tasks (3). Kerry and Rachel were better able to clap the syllabic rhythm of words in song than in speech (6). Kerry's performance was the most accurate of the group in music tasks: she produced a higher number of clapped syllables in sung phrases than in spoken (6); produced sung phrases fluently (14); and she was the only participant whose inaccuracies in beat and rhythm tasks were not attributed to reduced memory of the stimulus (12, 13).

Table 8.3: The relationship between auditory memory demand and performance in speech and music tasks. The table shows whether the data indicated a difficulty in the recall of rhythm or pitch in tasks requiring the use of auditory short term memory. The presence or absence of observed difficulties is indicated by Y (yes) and N (no). The table also indicates where an observed difficulty might be linked to STM. Red type is used to allow easy comparison between individual abilities. The table also shows the total number of tasks that were presumed to be unaffected by memory demand for each individual.

			Andrew		**Kerry**		**Rachel**		**Robert**	
		Domain	speech	song/ music	speech	song/ music	speech	song/ music	speech	song/ music
Digit span			3+		3		3		2	
Task		**Impaired recall of:**	**Difficulty implicated (Y or N), and max no. recalled correctly:**							
Rhythm	**1**	No. of syllables in words of 3 syll	N		Y: 2		Y: 2		Y: 2	
	2	No. of syllables in nonwords (CNREP)	Y: 4		Y: 3		Y: 3		no data	
	3	No. of syllables in phrases (up to 6 words)	Y: 6 syll; 4 words	Y: 9 syll; 5 words	Y: 6 syll; 4 words	Y: 6 syll; 4 words	Y: 2	Y: 6 syll; 4 words	no data	Y: 1
	4	No. of claps in TRaCoL rhythm (up to 7)	N: 7		Y: 5		Y: 4		Y: 2	
	5	Relative duration between claps	Y: 4		Y: 4		Y: 4		Y: 2	
	6	No. of claps in syll segmentation (up to 4 syll spoken/5 syll sung)	N: 4	Y: <4	N: 4	N: >5	Y: 1	N: <4	Y: 3	Y: 3
Pitch	**7**	Pitch contour: words of 1-3 syll	N		Y: 2		N		N	N
	8	Pitch contour: Nonwords of 2-5 syll	N		N		Y: 2 syll		no data	
	9	Pitch contour: phrases (max. no. of words 5)	Y: all	Y: 4	N	Y: 4	Y: all	Y: 4	no data	Y: 2
	10	Melodic Intervals		N: 3		Y: 2		N: 3		N: 3
	11	Absolute pitch		n/k MTD		N		N		Y
Possible effects of Auditory STM demands on:										
Perception	**12**	Beat entrainment		Y (104 bpm)		N		Y (104 bpm)		Y (132 bpm)
	13	Rhythm entrainment		Y		N		Y		Y
Voice	**14**	Dysfluency	N	N	Y	N	N	N	N	N
	15	Voice Perturbation/ interharmonics	Y	Y	Y	Y	Y	Y	N	N
	16	Emotional response: tension, confidence	Y	Y	N	N	Y	Y	Y	Y
Divided Attention	**17**	Reduction in segmental processing	N	N	Y	N	Y	Y	N	N
Total number of tasks presumed to be unaffected by memory demand			**7**	**3**	**4**	**7**	**2**	**4**	**4**	**5**

8.1.3 Progress in song-learning (Table 8.4)

The final two sessions of the six-week programme were cancelled by the host setting (see Appendix 3, p 341). The quality of teaching and learning was also affected by changes in group dynamics and by individual emotions (Appendix 3, pp. 336-349). However, when the participants were engaging well with the tasks and with each other, learning was affected positively – for example, dominant individuals were willing to demonstrate tasks, which allowed opportunities for discussion regarding what each participant 'did well'. Despite the loss of teaching sessions, all participants made some progress between their initial and final assessments in at least one aspect

Table 8.4: A summary of the group's progress in learning 'Magic Book'. The table shows how each participant improved between their first and final productions of the song. It also shows whether their performance in rhythmic tasks improved when given visual support and whether their singing performance improved with vocal cues when singing. The presence or absence of observed ability is indicated by Y (yes) and N (no). Additional information and scores are given, where data are available: the + and - symbols indicate an increase or decrease in accuracy, respectively. The magnitude of changes is indicated in numbers and the total number of improvements is shown. Red type is used to allow easy comparison between individual abilities.

		Improvements in:	Andrew	Kerry	Rachel	Robert
Pitch and interval	**1**	**Accuracy of pitch matching**	N	N (-3)	Y (+2/5)	Y (+3/6)
	2	**Range of pitches attempted**	Y (+3/5)	N (-3)	Y (+2: G4, C5)	Y (+3)
	3	**Accuracy of Intervals**	N	N	N	Y (+1/5)
	4	**Absolute pitch in intervals**	N	Y	Y	Y (2/5)
Magic Book	**5**	**Word recall (before any prompt)**	N (n/a)	Y (+4)	Y (+4)	Y (+2)
	6	**Segmental production**	N	Y (+4)	N (less accurate)	Y (+2)
	7	**Rate of articulation**	Y (x1.18)	Y (x2.17)	Y (x1.54)	Y (x0.72)
	8	**Accuracy of melodic intervals or shape**	N (3/4)	N (3/4 accurate)	N (3/4 accurate)	Y- increased range/melodic shape
	9	**Number of accurate melodic intervals**	N	N	N	N
	10	**Relative duration of syllable rhythm**	Y	n/a - accurate in first performance	Y	Y
	11	**Voice quality (no. of interharmonics)**	N	Y	N (n/a)	N (n/a)
		Number of improvements	*3*	*4*	*6*	*9*
Improvements observed with:						
Visual support	**12**	**Words in song (gesture/ Makaton)**	N	N	N	Y
	13	**Rhythm of words (cat, monkey)**	Y	N	Y	Y
	14	**Beat/rhythm eEntrainment**	Y	n/a- accurate in first performance	n/a - not tried	Y
	15	**Melody**	N	N	N	Y
Vocal cues	**16**	**Initial sound or syllable and pitch**	Y	Y	Y	Y

of melodic, segmental and rhythmic production (Table 8.4: *Number of improvements*). On reassessment, Andrew, Rachel and Robert produced a wider vocal range when matching pitches (2), and Rachel, Kerry and Robert reproduced more accurately the absolute pitch of some notes in intervals (4). Andrew recalled all words to the song on his first attempt (5), but the remaining participants independently recalled more words in their final performance (5). Kerry and Robert also produced a higher number of target speech sounds in their final assessment (6). However, Rachel produced fewer sounds, and less accurate sounds, with each repetition (6). In their final assessments, all participants articulated the words *magic book* at a faster rate (7) and all were able to independently reproduce the relative duration of the words and syllables in the phrase *magic book* (10). Kerry's voice production in her final performance contained fewer interharmonics (11). None of the group improved in the number of intervals they reproduced accurately (9: as measured to within half a semitone), but Robert sang with a wider pitch range and appropriate melodic shape (8).

Robert's performance improved the most in terms of the number of parameters that changed, and he also benefitted in all tasks from visual support (12-15). Andrew and Rachel benefitted from visual presentation of rhythms (13), and Andrew and Robert played rhythms and beats more accurately in response to a visuo-spatial gesture. All sang with greater accuracy when given vocal cues for pitch (16).

Table 8.5: Transfer of knowledge from song to speech of the words to the song 'Happy Birthday'. The table summarises the lexical, segmental, prosodical and vocal features features that were observed in the group's spoken and sung productions. The presence or absence of observed ability is indicated by Y (yes) and N (no). Additional information and scores are given, where data are available. Red type is used to allow easy comparison between individual abilities.

		Andrew		Kerry		Rachel		Robert	
		Song	Speech	Song	Speech	Song	Speech	Song	Speech
1	**Max. no. of words recalled without subsequent prompt**	6/16 (37.50%)	11/16 (68.75%)	16 (100.00%)	4 (25.00%)	11/16 (68.75%)	8/16 (50.00%)	4 (25.00%)	2/16 (12.5%)
2	**Approximate syllabic stress of the song produced**	Y	Partial-retained some lengthened vowels	Y	N	Y	N	Y	N
3	**Relative duration of syllables in song produced/retained**	N	N	Y	N	N	N	N	N
4	**Pitch shape produced/ retained**	N	Y (final phrase)	Y	N	Y	N	N	N
5	**Pitch range (ST)**	5.30	7.30	6.76	3.24	12.80	17.00	4.55	3.10
6	**Speech rate**	0.96	1.12	1.99	2.49	4.14	4.95	2.67	4.67
7	**Interharmonics (visible at 35 dB, 1000 Hz, 15 ms)**	100%	73%	56%	41%	67%	70%	100%	75%

8.1.4 Transfer of learning between song and speech (Table 8.5)

The productions of the group when speaking the words to *Happy Birthday* were qualitatively different to their sung versions. Kerry, Rachel and Robert independently produced more words when singing than when speaking (1). Only Kerry sang with the appropriate rhythmic and melodic features (2-4). The whole group sang with appropriate word stress, but they produced spoken words with a different stress pattern (2) and at a slower rate (6). Andrew partially reproduced the song's rhythm and melody when speaking the final phrase (2, 4). Kerry's spoken production was fluent (6). Andrew, Kerry and Robert produced fewer interharmonics when singing (7).

8.2 General discussion

This section will discuss the key difficulties and barriers faced by the study group, as outlined in Section 8.1, and will compare these to the wider DS population. Section 8.2.1 discusses the common abilities and difficulties in areas of perception and cognition, singing, voice production and prosody. It links to the first aim of the study; w*hat are the verbal and non-verbal abilities of the individuals in relation to each other and to the DS population as a whole?.* Section 8.2.2 examines the potential perceptual, physiological and cognitive factors that affect the abilities of the group, drawing on evidence from the DS population. This links to the second aim of the study: *what factors affect the individual's and group's musical and speech skills in terms of rhythm, melodic intonation and phonation?* Section 8.2.3 discusses the ability of the group to learn songs. This links to the third aim: *to what extent can the individuals and the group learn to improve their musical ability in terms of perception and production of musical features through song-based teaching?* The final section, 8.2.4 discusses the evidence that the group is able to transfer learning from the musical domain to the speech domain. This provides preliminary discussion of the fourth aim, *what is the potential for rhythmic and song-based music tuition to develop aspects of auditory processing, voice and speech production for people with DS?* This aim will be explored in greater detail in Section 8.3.

8.2.1 Key abilities and difficulties within the group

Perception and cognition

A spectrum of cognitive, motor and speech abilities was observed in this group (Table 8.1) which is typical of the DS population (Capone, 2004; Silverman, 2007; Naess et al., 2011). The mean scores of the group (Table 8.1) are lower than those reported for verbal MA based on BPVS by Laws (1998), who reported a mean MA of 4 years (spread: 1;11-7;2 years) in 33 children with DS and SLD (CA = 11;6). This spread could be attributed to the wide age-range of participants (5-18 years), or by the acknowledged heterogeneity of people who are classed as having SLD across different groups (Katz and Lazcano-Ponce, 2008; Giangreco, 2006). The mean PPC score for DEAP (62.62%: Table 8.1) is lower than those reported by Wood et al. (2008: 67%) and in Kent

and Vorperian (78%), but is comparable to a group of six DS subjects assessed by Flipsen (1999: 60.2-82.3%). A mean digit span of 3-4 digits is typically reported in people with DS (Kay-Raining Bird et al., 1994, 2000; Laws, 1998; Seung and Chapman, 2000) irrespective of the CA of participants (Purser and Jarrold, 2005). The mean digit span of this group seems to lie within the lower limits of previously reported scores. The spread of scores in CNREP (Table 8.1: 5) may highlight subtle differences between participants in phonological working memory (Laws, 1998). The data from both tests (Table 8.1: 1, 5) suggest that Andrew has the highest phonological memory span.

Nonverbal MA is close to verbal MA in DS subjects although some studies have found higher nonverbal MA (see Naess et al., 2011 for a review). The results for this group (Table 8.1: 3) indicate superior abilities in nonverbal skills for DAP (Table 8.1: 3) and for overall musical ability (Table 8.1: 6) than for verbal ability (Table 8.1: 1). DAP may be an inappropriate indicator for non-verbal MA, especially in relation to modern psychometric testing (Laws and Lawrence, 2001). However DAP may be indicative of motor development, which is correlated with MA (Sacks and Buckley, 2003; Sourtiji et al., 2010). The data suggest a close link between music-motor skills and DAP for this group (Table 8.1: 6, 3).

Singing and rhythmic-motor movements

Although the mean music age-equivalent scores are similar for Andrew, Kerry and Rachel (Table 8.1: 6), the individuals are at different phases of development in specific sub-skills (Table 8.1: 7-12). Such a spread of ability in singing (e.g. Davidson et al., 1981; Lamont, 2009; Welch, 2016) and in rhythmic performance (Drake et al., 2000; Gooding and Standley, 2011) is typical of TD children. However, as a result, the broad range of age-equivalent scores for some skills (e.g. pulse and meter, rhythm: Table 8.1: 9, 10) do not reflect the more developed abilities of individuals within certain tasks. Of the group, Kerry (Chapter 5) was the most accurate and consistent in beat rhythmic entrainment tasks (see also Table 8.4: 10,14; and Chapter 5: Table 5.6); Andrew was the most accurate in rhythm imitation tasks (e.g. Chapter 4: Table 4.1, TRaCoL, test 4); and Rachel (Chapter 6) was the most accurate in matching gross motor movements to music (Table 1 in each case study, test 18).

As a group, the ability of the participants in rhythmic tapping tasks (Table 8.1: 9) is consistent with available research in DS that indicates a qualitative difference relative to TD controls of equivalent MA (Stratford and Ching, 1989) and age-matched TD controls (Ringenbach et al., 2006). The rhythmic-motor abilities of this group are in line with their DAP scores (Table 8.1: 3), which indicates a motor delay for all participants, as is observed in DS subjects (Sacks and Buckley, 2003; Silverman, 2007). However, as reported in other studies of rhythmic motor movement in DS (Ringenbach et al., 2006; Chen et al., 2015), visuo-spatial support led to improvements in motor

timing (Table 8.4: 13, 14). It is likely that much of the variance in rhythmic motor timing can be explained by differences in auditory memory capacity and in their focus of attention, as is reported in TD children (Drake and Gérard, 1990; Drake et al., 2000).

As far as the author is aware, there are no data available to compare the group's singing abilities with other DS participants.

Voice range, pitch and quality

The vocal ranges produced in vowel glides by Andrew, Robert and Rachel (Table 8.1: 14, 15) are larger than the means reported by Lee et al. (2009) for their DS participants of a similar age, and are approximately half of the range produced by their TD controls. The pitch range produced by Kerry (Table 8.1: 14, 15) is considerably lower than those of Andrew, Rachel and Robert, and of the male and female DS participants in the study by Lee et al. (2009). Her reduced range may stem from her difficulties in maintaining efficient phonation (Chapter 5: Section 5.5.4) or from a failure to understand the objective of the task. The data indicate that the DS group as a whole may have a reduced ability to produce phonation across a wide vocal range. This supports the findings of Lee et al. (2009). This difficulty might relate to any number of physiological factors, such as laryngeal anomalies (Venail et al., 2004), the degree of hypotonia within the larynx (Howell, 2010), or atypical mucosa affecting the vocal folds (Beck, 1988). Such factors may account for the variation between participants within this study and for differences in findings between this study and that of Lee et al. (2009). However, methodological differences might also account for differences between these studies. For example, Lee et al. (2009) do not state whether they demonstrated the vocal glide to their participants, the range and type of which may have influenced production in either study.

In conversational speech, the group's mean deviation around the mean (Table 8.1: 16) is lower than that reported by Lee et al. (2009), whose findings were based on the readings of *The Rainbow Passage* by four male and five female DS participants. The SD for Rachel is comparable to the mean reported by Lee et al. (2009) for their female participants. However, there are significant differences between the studies that make it difficult to draw clear comparisons. These include the use of conversational speech in the present study rather than reading, as neither Rachel nor Robert were able to read. The known hearing loss of Robert, and suspected hearing loss for Kerry may have further limited their prosodic output (Coelho et al., 2015).

Relative to norms for each sex, the participants in this study produced conversational speech at a higher mean F0 (Table 8.1: 15) (208 Hz, f; 108 Hz, m — in Goy et al., 2013). The data agree with findings by Albertini et al., (2010) and Lee et al. (2009), who reported a higher mean F0 in DS subjects. However, as with previous studies, although the acoustic data report near-normal measures, the perceived voice qualities are atypical (Table 8.1: 17). The nature of the

interharmonics produced in sustained vowels (Figure 8.1) and speech (Figure 8.2) suggest possible reasons for the inconclusive relationship between perceived voice quality and acoustic measures. Kramer, Linder and Schönweiler (2013) investigated the relationship between perceived voice quality and the number and intensity of interharmonics that were produced by 145 speakers with dysphonia and rough voices. They reported a correlation between high interharmonic content and their raters' perceptions of low F0. They also found that the intensity of subharmonics, and the percentage of low F0 values were also linked to perceived roughness. They conclude that the percentage of interharmonics in speech increases the perception of low-pitched noise and contributes to the perception of roughness. The data from the present study suggest that the incidence and duration of interharmonics at offset of sustained vowels (Figure 8.1) may distinguish the typically rougher voice qualities of Andrew and Kerry (Table 8.1: 17). Furthermore, the intensity of Andrew's interharmonics distinguished his production from that of the other participants (Chapter 4: Section 4.5.3), and may reflect his habitually harsh, ventricular production. The data in the present study suggest that the intensity of interharmonics may usefully distinguish between sub-types of dysphonia within the DS population. Furthermore, Kramer et al. (2013) reported an inability of the algorithm used to measure F0 when interharmonics were present, as also found in this study. As the lack of periodic voicing leads to inaccurate or null readings for acoustic voice measures, a high incidence of interharmonics may account for the acknowledged difficulty in aligning perceived quality with acoustic data in people with DS (Albertini et al., 2010; Lee et al., 2009; Moura et al., 2008).

Despite agreement that the voice quality of people with DS is atypical, previous studies have reported mixed findings for acoustic voice data of speech or sustained vowels (Albertini et al., 2010; Lee et al., 2009; Moura et al., 2008). In this study, the mean acoustic data for vowels in speech do indicate a high level of pathology, in contrast to the findings of Albertini et al. (2010) and Lee et al. (2009). Methodological differences may account for these, such as the use of Sonneta in this study, and different speech stimuli. However, not all speech samples were produced by participants in this study with atypical values (see Chapters 4-7). Furthermore, the mean perturbation data from the central portion of sustained vowels in this study (Figure 8.1) do not conclusively indicate pathological production in this task. The mean HNR in the medial portions of vowels is higher for Rachel and Robert, whose voices are typically less rough-sounding in conversational speech, than for Andrew and Kerry. However, in sustained vowels, the mean HNR data for Andrew are within norms for younger male adults (22.9 dB); and those for Kerry are just below norms for female adults (25.3 dB)(Goy at al., 2013). For Rachel, mean measures of jitter (0.35%) and shimmer (2.62%) in the central vowel are within norms (Goy et al., 2013); but mean values of jitter and shimmer for Andrew (0.60%, 3.77%) and Robert (0.74%, 3.61%) are similar to each other, despite markedly different voice qualities in their speech. The data partially support the

findings of Moura et al. (2008), who reported higher levels of pathology in sustained vowels in Portuguese children with DS, aged 3-8 years. However, the findings also support evidence from studies that report normal values of jitter and shimmer in children and adults with dysphonia and perceived rough voices (Kramer et al., 2013; Nuñez-Batalla, Nieto, Pinto and Ferreras, 2000).

The findings support suggestions that dysphonia is prevalent in the DS population (Beck, 1988; Kent and Vorperian, 2013; Moran, 1986; Pryce, 1994), and that acoustic measures of sustained vowels and speech do not necessarily reflect this. However, in this study there was a clear indication that voice production differed in response to task demands (Table 8.1: 17, 18; Table 8.3: 15; Figure 8.2). In particular, tasks requiring increased demands on phonological memory were produced with an increase in perturbation; whilst those drawing on long-term memory were produced with lower perturbation. It is therefore possible that in the wider DS group, acoustic measures may vary according to the inherent task demand and to the individual's response to tasks. This has not been previously investigated in DS, although research into TD populations does show that cognitive load affects voice production (Giddens et al., 2013; Quatieri et al., 2013; Yap, 2012). The types of difficulties associated with voice production in the participants of this study will be discussed in further detail in Section 8.2.2.

Prosody

The findings demonstrate atypical prosody, in line with previous research in DS (Bray, 2008; Pettinato and Verhoeven, 2009; Setter and Stojanovik, 2011; Stojanovik, 2011; Zampini et al., 2015). The group was unable to reproduce all changes of rhythm, pitch and that marked the syllabic stress of 1-3 syllable words (Table 8.1: 20-22). For all participants, prosodic production would be affected by physiological difficulties in sustaining phonation and in regulating pitch and intensity. In addition, the data suggest that the speech production of all participants but Robert are affected by dysfluency. Dysfluency is believed to affect 47% of the DS population (Bray, 2008): this group shows a higher incidence rate (Table 8.1: 19). The presence of stuttering was task-dependent (Table 8.2:7), and for Kerry, there was a notable reduction in singing tasks, and when speaking from LTM. Nevertheless, the findings indicate a common motor-deficit for speech. This could have implications for the ability of the participants to control the temporal aspects of speech production in a range of tasks.

The evidence suggests that the group has difficulties in prosodic output. However, it was also suggested that reduced memory affected the participants' abilities to reproduce the number of syllables in words of 3 syllables (Table 8.3: 1), and in nonwords (Table 8.3: 2) which would further affect the prosodic output. The data also suggest perceptual deficits for weak-initial words in nonsense words (see Table 8.2: 10), in line with findings by Pettinato and Verhoeven (2009). Data from Phrase 4 of the song *Magic Book* suggested that weakly-stressed words in song were also

less readily perceived by all participants (Table 8.2: 10). Evidence from Robert's case study (Chapter 7) suggested that difficulties in auditory perception or auditory processing would further affect the accuracy of prosodic production in imitation tasks. This would have consequences for the ability of individuals to acquire stable representations of words.

Summary

The data show a strong link between musical ability and nonverbal cognitive ability, as indicated by the DAP scores. The group results for perception, cognition and motor abilities are representative of people with DS, but suggest that this group lie within the lower ability of the spectrum, in line with their pre-existing diagnosis of severe learning disabilities. The voice quality and prosodic abilities of the groups are representative of the DS group, but their voice production is affected by task demands. For most of the group in this study, voice production is more melodic and is produced with fewer perturbations and interharmonics when singing. The data suggest that the musical and speech domains within the group have features in common, but that speech production is less well developed than limb-motor or vocal musical production. The potential reasons for this are discussed in Section 8.2.2. Overall, the data support hypothesis 1: *for adults with DS, musical development will be in line with overall non-verbal cognitive development and vocal abilities*.

8.2.2 Key factors affecting rhythm, melodic intonation and phonation

Hearing, perception and auditory processing

Hearing impairment accounted for much of Robert's difficulties on all tasks (Chapter 7). Reduced hearing acuity may also account for Rachel and Kerry's poorer performance in speech perception tasks (Table 8.2: 8, 9). Furthermore, HI affects the ability to regulate respiratory and laryngeal control (Bolfan-Stosic and Simunjak, 2007; Coelho et al., 2015; Monsen, 1979). This results in higher levels of voice perturbation (Coelho et al., 2015; Deqhan and Scherer, 2011), and diplophonia (Monsen, 1979; Das et al., 2013). Robert's confirmed HI is likely to explain his phonatory difficulties. Given the high incidence of HI in people with DS (Alexander et al. 2015; Maris et al., 2014), it is possible that HI also affects the voice production of other participants. Also, given that HI is prevalent in DS children (Laws and Hall, 2014; Maris et al., 2014), and is associated with hyper- or hypo-functional voice disorders in HI populations (Das et al., 2013), it is a possibility that earlier, intermittent HI may have contributed to the development of hyper- or hypo-functional habits.

Although poor speech production is associated with the degree of hearing loss in young children with DS (Laws and Hall, 2014), speech production difficulties in DS cannot be solely explained by HI (Chapman et al., 2001; Laws, 2004). The performance of the group in some tasks (Table 8.1: 19) indicate reduced or atypical auditory perception or processing. Auditory processing may be

slow in people with DS (Marcell and Cohen, 1992; Marcell, 1995) and neurological studies demonstrate atypical processing of speech sounds (Chua et al., 1996; Elliott and Weeks, 1987; Elliott and Weeks, 1993; Heath et al., 2005; Groen et al., 2008). A deficit in the coupling between auditory information and motor response was suggested to explain Robert's inconsistent motor production in beat entrainment tasks (Chapter 7: Table 7.6), which would be exacerbated by his hearing loss (Marcell et al., 1990; Marcell, 1995). Kerry's performances on rhythmic tasks did not suggest poor perception of the stimuli (Table 8.2: 1 and 14), but auditory processing responses can develop in response to previous training in TD children (e.g. Kraus and Chandrasekaran, 2010; Chandrasekaran et al., 2015; Parbery-Clark et al., 2009, 2012). Furthermore, there is increasing evidence that musical training leads to superior auditory processing in those with HI (Parbery-Clark et al., 2013; Rochette et al., 2014). Therefore, Kerry's superior timing in beat entrainment (Table 8.3: 12) may be partly the result of her training.

Reduced processing and hearing acuity lead to poor mental representations and inconsistencies in word recall, as observed in the group's productions of familiar words (Table 8.2: 18). However, poor mental representations are also associated with poor phonological memory (Jarrold et al., 2009; Gathercole, 2006; Gathercole and Baddeley, 1993; Laws, 1998). Effective processing also requires attention to the salient features. Any perceptual difficulties will therefore be confounded further by reduced memory capacity and behaviour.

Physiological and neurological factors

Articulatory difficulties were evident in all participants in speech tasks (Table 8.2: 3). Articulatory difficulties and nasal speech were evident in all productions: these were not examined in the case studies because these are common features of DS speech (Beckman et al., 1983; Kent and Vorperian, 2013; Montague and Hollein, 1973; Rodger, 2009), and because there was no evidence of this being different in singing than in speech. However, hypernasality can affect air pressure within the glottis and impact upon phonation (Gick et al., 2012; Mathieson, 2001). Thus, the group's physiological difficulties in producing speech sounds may have consequences for temporal aspects of their speech production and for their phonation. Additionally, in sustained vowels, fluctuations in the first formant contour coincided with fluctuations in the intensity contours for Kerry (Figure 5.4), Rachel (Figure 6.3 a) and Robert (Figure 7.3, Figure 7.4); and coincided with changes to the number of interharmonics for Andrew (Figure 4.4) and Robert (Figure 7.4). Their difficulties in regulating sub-glottal pressure may be linked to their ability to regulate articulatory configurations, with consequences for the stability of phonation (see Chapter 4 and Chapter 7 for discussions).

No participant was able to complete the DDK task, which is in common with previous findings (Brown-Sweeney and Smith, 1997; Rosin, Swift, Bless and Vetter, 1998; Cleland et al., 2010).

Despite the prevalence of hypotonia in DS (Cleland et al., 2010; Kent and Vorperian, 2013) an inability to sequence /p,t,k/ might indicate motor-timing and sequencing difficulties (Cleland et al., 2010). For the participants in this study, such difficulties might be common to the oral-motor movements, limb-motor movements and phonation. For example, motor timing difficulties are apparent in the dysfluent productions of Andrew, Kerry and Rachel, and in the variability of their timing between claps in rhythm imitation and syllable segmentation tasks. Previous research with TD children and adolescents has identified a connection between stuttering and poor performance in rhythm tasks (Falk et al., 2014; Falk et al., 2015; Olander et al., 2010), which is independent of motor-co-ordination in clapping tasks (Olander, 2010). Furthermore, the ability to co-ordinate hand-motor movements in tapping to a beat is linked to processing speed of consonant sounds in adults with normal hearing (Tierney and Kraus, 2013) and to the ability to integrate motor and auditory feedback (Phillips-Silver et al., 2009, 2010; Pfordresher, 2006; Thaut, 2008). Falk et al. (2015) proposed that slower processing and reduced integration of the auditory-motor signal is common to impaired speech and performance in motor tasks in people who stutter. It is possible that the participants in the present study may have a neurological deficit in motor timing that is common to their speech production and accuracy in rhythmic motor tasks, as argued in other speech-impaired populations (Corriveau and Goswami, 2009; Cumming et al., 2015; Olander et al., 2010; Peter and Stoel-Gammon, 2008). If this is the case, the group data and case studies indicate that in both domains, this is exacerbated or mediated by hearing acuity, task demand, motor ability and previous learning.

The elevated degree of voice perturbation at the onset of sustained vowels demonstrates a difficulty in initiating phonation (Figure 8.1). In TD populations with phonatory difficulties this can indicate dysphonia (Braunschweig et al., 2008; Cavalli and Hirson, 1999; Fourcin and Abberton, 2008). The data suggest the use of additional muscle tension, in line with evidence given by Pryce (1994). Excessive muscular tension, as observed in Kerry and Andrew, can affect the sub-glottal pressure and stability of phonation (Braunschweig et al., 2008; Titze, 1994, 2008; Baken and Orlikoff, 2000). The degree of perturbation increased in speech tasks, which reflects the increased demands on the vocal system (Fourcin and Abberton, 2008). This is also typical in populations with dysphonia, whose voice takes longer to stabilise at onset (Schaeffler et al., 2015). The group demonstrated difficulties in producing appropriate changes in intensity in words (Table 8.2), which may result from reduced respiratory control (da Silva, 2010; Heselwood et al.,1995; Salguerinho et al., 2015). The study cannot exclude the presence of vocal fold pathology, but the data indicate that the group had difficulties in achieving the optimum phonatory pressure for speech required at initiation (Titze, 2008) with consequences for voice production and prosody. As the hearing acuity of Andrew, Kerry and Rachel is not known, it is also possible that a degree of HI contributes to their

voice production difficulties (Coelho et al., 2015; Deqhan and Scherer, 2011). However, in speech tasks, the degree of perturbation also increased with cognitive load (Figure 8.2).

Cognitive load and divided attention

Reduced STM may account for many of the participants' difficulties in all imitation tasks (Table 8.3). Limited memory capacity in DS affects the amount of information retained in tasks and can result in errors of serial recall (Jarrold et al., 2004; Purser and Jarrold, 2005). Reduced capacity may result in poor mental representation of temporal templates at slow tempi in beat entrainment tasks (Grahn and Schuit, 2012; Repp and Doggett, 2007), resulting in the observed instabilities in motor performance (Table 8.3: 12, 13). Errors in serial recall were made in interval-production tasks by Rachel and Kerry, and in rhythm imitation tasks for all but Robert: his production was limited to two claps. A reduced STM effectively adds to the cognitive load of any task that draws on auditory STM (Sweller, 2011), such as imitation tasks. Furthermore, an increase in pressure on working memory — for example, as the need to sustain or divide attention increases — results in the deterioration of performance (Beilock and Ramirez, 2011).

The results of this study suggest that reduced STM may make the speech and voice production systems of people with DS particularly vulnerable to tasks requiring working memory, in line with TD populations (Bosshardt, 2006; Johnstone and Scherer, 1999; Yap, 2012; Quatieri et al., 2015). For all participants in this group, tasks that required little attention or conscious processing were produced with fewer voice perturbations and fewer indications of stress-response (Table 8.3: 15, 17). By contrast, for Andrew, Kerry and Rachel, tasks that required overt attention resulted in more severe breakdowns in one or more aspects of production (e.g. Andrew, Figure 4.19; Kerry, Figure 5.16; Rachel, Figure 6.6). The combination of articulatory difficulties and memory difficulties prevent even simple speech processes from becoming 'automatic' in people with DS (Kumin and Adams, 2000; Silverman, 2007). As a result, every speech production task requires conscious effort and attention, and draws more extensively on working memory in production (Silverman, 2007).

Tasks that require dual attention place a greater strain on memory, which overloads processing resources (Mattys and Wiget, 2011; Koelewijn, Shinn-Cunningham, Zekveld and Kramer, 2014). This can result in poorer speech production (Jou and Harris, 1992; Tumber et al., 2014) and speech perception (Mattys and Wiget, 2011). Dual attention also affects accuracy in song production (Dalla Bella et al., 2012; Gordon et al., 2010; Racette and Peretz, 2007). The verbal and vocal responses from participants in this study suggest that as task-related demands increased beyond their STM capacity in speech and in song, they limited their attention to either segmental production or to suprasegmental production (Table 8.3). For example, an improvement was noted in HNR for Kerry and Rachel when they 'switched' from processing segmental features

in CNREP and in Syllables tasks to prosodic features (Table 5.10, Table 5.11, Table 5.12, Table 5.13, Table 6.3, Table 6.13, Table 6.14, Table 6.16). In contrast, Andrew's production in all tasks was more accurate at the segmental level, and less accurate in terms of prosody (Table 8.1; Chapter 4) or melody (Table 8.4); whereas Robert's speech was produced with sustained phonation but with a higher number of speech errors (Table 8.1, Chapter 7). Furthermore, all participants reproduced the 'global' features of new songs before they learned the details (Table 8.4). Although this is a characteristic of TD populations (Davidson et al., 1981; Welch, 1998; Welch 2016), a preference in attending to global features has been indicated in people with DS (Porter and Coltheart, 2006). People with DS are impaired in their ability to divide attention (Carney et al., 2013; Costanzo et al., 2013) and to process multiple sources of information (Lanfranchi et al., 2012). This study suggests that the individuals in this study might switch to a global style of processing when the amount of 'local' segmental detail increases. Furthermore, people who stutter have a reduced ability to divide attention (Bosshardt, 2006). Therefore, limiting the focus of attention may support speech and voice output by reducing task demands and stress, which can increase dysfluency.

Summary

Deficits in perception, phonation and articulation were common to both musical and speech domains. However, some observed behaviours, such as poor motor timing in speech and in motor-music tasks may be affected by one of several causes or a combination of these. Furthermore, the difficulties were affected in each domain by task demand, and may be ameliorated by prior learning. The data partially support the hypothesis 3: *given the shared physiology between singing and speech and correlation between certain speech and musical abilities in terms of perception and production, difficulties that affect the speech domain will correspond to difficulties in the musical domain.* However, the results are based on comparison of observed behaviours and the underlying causes for behaviours in each domain may be different.

8.2.3 Song learning

When learning new songs, TD children of primary age reproduce the words, initially, and then add increasing rhythmic and melodic detail to the production (Hargreaves, 1986; Reifinger, 2006; Welch et al., 1994, 1996, 1998). This group showed similar trends in learning except that they had greater difficulty in learning to produce the words (Table 8.2: 11; Table 8.4: 5). This affected their musical production (Table 8.2: 11) and Rachel and Robert seemed to have no concept of the meaning of the words (Table 8.2: 13). Children with DS are better able to learn the referent of new words than the phonological form (Jarrold et al., 2009) and they recall familiar words more easily than novel words (Smith and Jarrold, 2014). Therefore, teaching the meaning of songs may support their memory for words in songs.

All participants were able to reproduce the relative duration of the words in the first phrase of *Magic Book* (Table 8.4: 10). However, with the exception of Kerry, their rhythmic production of phrases remained affected by impaired recall of the words. In TD children, accuracy in rhythmic aspects of singing develops before accuracy in intervals (Trainor and Hannon, 2013; Gooding and Standley, 2011; Welch, 2016). The group did not develop in the accuracy of their melodic intervals. The song predominantly used the interval of the minor third (*so-me*) which is believed to be one of the first that TD children reproduce (Hargreaves, 1986; Welch, 2016), although Houlahan and Tacka (2008) argue that it can be more challenging than smaller intervals. Andrew, Kerry and Rachel were all capable of producing the interval to within half a semitone, but they were inconsistent in their production of it. Pitch matching and accuracy develops at about 5-6 years, but a sense of tonality typically develops at about 7 years (Hargreaves, 1986; Gooding and Standley, 2011). Kerry and Rachel both produced stable tonal centres, which indicates that productive difficulties may have affected their production, rather than developmental difficulties. Although Robert did not improve in pitching skills, he reproduced the melodic shape with a wider vocal range when re-assessed. Visual and motor tasks assist younger children (3-5 years) to reproduce pitch (Miyamoto, 2007, in Gooding and Standley, 2011) and the same effect was noted for Robert (Appendix 3, p. 348 [13], 349 [22]). The improvement may reflect increased perception of the melody as a result of the visual and gestural teaching style. With continued visual support, he may develop further.

Summary

The progress demonstrated in *Magic Book* highlights that most of the group may be capable of developing singing and musical skills beyond their current abilities. In addition, the data indicate that the different abilities observed within the group in music tasks are the result of previous experience, in line with the TD population. The data support hypothesis 3: a*dults with DS and SLD are capable of learning songs and of developing their melodic-rhythmic skills in singing. Individuals with DS can develop in both awareness and production of musical features, when these skills are explicitly taught.*

8.2.4 Transfer between domains

For novel lyrics, TD adults recall fewer words when asked to sing lyrics than when asked to speak them, possibly in response to speech imposing fewer processing demands (Racette and Peretz, 2007). With the exception of Andrew, the group produced fewer words spontaneously when speaking than when singing. The slower production in singing (Table 8.5: 6) may have given them more processing time (Racette and Peretz, 2007), or the novelty of the task may have affected their performance. The group was able to transfer knowledge learned in song to speech. With the exception of Andrew, the group altered the duration of vowels and the scope and direction of pitch

changes so that their spoken production was distinct from their singing (Table 8.5: 2-5). Despite the challenge of the task in separating the words from the melody (Racette and Peretz, 2007), aspects of vocal production were improved, relative to other speech tasks: for example, Kerry and Robert produced lengthier phrases than in other tasks (see Chapter 5, Chapter 7); and with the exception of Rachel, they produced fewer interharmonics when speaking *Happy Birthday* that when singing it (Table 8.4; and see Chapters 4-7). In addition, the task enabled Kerry to speak fluently; probably as a result of the strongly metrical production, that can assist motor-timing (see Chapter 5: Section 5.7).

Summary

The group was able to transfer pre-existing learning. There are indications that aspects of voice and speech production were superior in this task than in speech imitation tasks. The study supports hypothesis 4: *If the group is able to transfer learning between domains, and if the appropriate functional, emotional and cognitive conditions are met, then the development of musical skills may be employed to develop aspects of speech and vocal production.*

8.3 Implications for the application of findings

Section 8.3.1 considers how the findings from the study can inform our understanding of the abilities and of people with DS in musical and speech domains. It addresses the fourth aim of the study: *for people with DS, what is the potential for rhythmic and song-based music tuition to develop aspects of auditory processing, voice and speech production?* Section 8.3.2 makes suggestions for policy regarding the provision of music for people with DS. Section 8.3.3 recommends a programme of activities that might be used to support speech processing and production in people with DS.

8.3.1 General implications

The findings of this study indicate that the same types of physiological, neurological, cognitive and behavioural factors affected the participants' production in speech tasks and in music tasks. However, when learning to sing a song, most of the group developed in one or more aspect of speech production. An improvement was also observed in their performance of musical tasks when STM was supported visually, and an improvement in voice quality and vocal range was evident when the group produced gestalt phrases from long term memory. The production of phrases was typically fluent in all singing tasks, and in the transfer of song to speech. The findings suggest that the group is able to learn when teaching is adapted to their difficulties in hearing, processing and memory. They also suggest that song-learning may be a suitable means to support the learning of words and phrases for the participants.

The four participants were described as having severe learning disabilities. The majority of people with DS have moderate learning disabilities (Chapman and Hesketh, 2000; Dierssen, 2012), and the group's performance on key cognitive tasks (BPVS, Digit span) place them towards the lower spectrum of ability within the wider DS group. At the age of 25 years, their cognitive abilities may have peaked (Capone, 2004; Iacono et al., 2010; Vicari, 2006; Wuang and Su; 2011) or may be approaching their peak (Couzens et al., 2011a). However, in the wider DS population, cognitive development and adaptive behaviour may continue to develop into adulthood (Couzens et al., 2011; Dressler, Perelli, Feucht and Bargagna, 2010), even as IQ declines (Dierssen, 2012; Vicari, 2006). Despite their age and learning difficulties, the participants in the study demonstrated an ability to learn songs and to improve aspects of their production, in line with other studies involving DS adults (Ringenbach et al., 2006).

The study suggests that neither age nor ability should limit the opportunities for individuals with DS to take part in, and learn from, musical activities. However, early experiences are critical in shaping the DS brain (Capone, 2004). Furthermore, early musical experiences are likely to have a greater impact than late-occurring experiences (Habib and Besson, 2009), especially for auditory processing and motor processing (White et al., 2013). The early experiences of the participants in the study group are not known, but younger children with DS may have greater opportunities for learning as a result of recent improvements to health, educational and social provision (e.g. Cuckle, 1999; Lanfranchi and Carretti, 2012). It may be particularly important to support auditory processing in young children with DS as a result of the high incidence of glue ear (Buckley and Sacks, 2001; Maris et al., 2014) which can have long-term consequences for language learning (Buckley, 2003; Pettinato, 2009) and classroom-based learning (Bennets and Flynn, 2002). Therefore, if music training can potentially lead to neurological or behavioural benefits as is reported in TD children (e.g. Besson et al., 2011; François et al., 2013; White et al., 2013), a programme of music that supports auditory processing from the early years onwards would be most beneficial for supporting developing language skills in the DS population.

8.3.2 Implications for policy

Although further study is required to test the outcomes of music provision on speech development for this group, it is recommended that music provision is offered for individuals with DS from infancy through to adulthood. In TD musicians, the magnitude of neurological, cognitive and behavioural changes that benefit speech processing is linked to the age of onset of training and to the duration and intensity of practice (Habib and Besson, 2009; Chandrasekaran et al., 2015). Although changes can take place after a 'sensitive' period (White et al., 2013), change is easier if training begins by the age of 6 or 7 years (Schlaug et al., 1995; Strait et al., 2013). The typical developmental trajectory of those with DS indicates that earlier provision may be needed for this

group. Infants with DS show differences in the quality of their responses to infant-directed communication (d'Etiole et al., 2015 Glenn et al., 1982), and mother-infant interaction may be atypical (Mahoney and Robenalt, 1986; Slonims and McConachie, 2006). Infants with DS vocalise less, and produce fewer vocal imitations (Fidler, 2005). Additionally, bouts of otitis media in the first year of life may be linked to later deficits in auditory processing (Fidler, 2005). Infants with DS are therefore at risk of impoverished auditory stimulation: potentially, the 'rich' stimulation of music could compensate for this, supporting early auditory development.

Given that access to targeted SaLT may be limited and that the onus is on schools to support speech, language and communication needs (Dockrell, Lindsay, Roulstone, and Law, 2014), school-based music provision might provide essential opportunities for supporting speech and language for people with DS. Indeed, the requirements of the current EYFS (DfE, 2012) and Primary (DfE, 2013) curricula enable ready links between musical activities and development of listening and speaking, understanding of language and diction. However, music provision in the majority of state-funded schools is poor in terms of opportunity and quality of teaching (Ofsted, 2012). It is questionable whether primary teachers have the knowledge or confidence to teach music (DfE, 2011; Henley, 2011; Garrett, 2014; Welch and Henley, 2014), and whether teachers in general are able to adapt music provision for children with special educational needs (Jellison, 2013; Ofsted, 2012). Given a recognised lack of training for teachers in primary schools, current UK policy recommends collaboration between schools and specialist providers, through local Music Hubs (Henley, 2011; Ofsted, 2013). However, visiting specialists do not necessarily have the teaching skills required to differentiate for whole classes, nor the knowledge of pupils that might enable progression (Ofsted, 2012; Ofsted, 2013). This puts children with different learning needs at a disadvantage in terms of how well they can engage in lessons, and how effectively they can learn. The development of musical skills is a pre-requisite for transfer to the linguistic domain (Patel, 2011, 2014). If children with DS are to benefit from music provision in terms of speech perception or production, an approach needs to be adopted within schools or other music providers that enables progression, and that is readily tailored to the needs of pupils with SEND.

In contrast to the current knowledge-based approach to formal music education (DfE, 2013; 2014), the present study recommends a focus on hands-on music-making for pupils with DS in order to support musicality. Such an approach, if delivered effectively, can be used as a basis for developing academic learning (such as composition, reading of music notation), thereby also meeting the curriculum requirements. Approaches that focus on the integration of skills, rather than on mastery of a given instrument are recommended. The use of a Kodaly approach was successful in engaging the participants in the present study, and the same approach was highlighted as an example of good practice in schools (Ofsted, 2012). Furthermore, the Kodaly approach and other

developmental approaches (e.g. those of Orff and Suzuki) enable the development of musicality though playing. These enable easy differentiation in ensemble practice, making them appropriate for including pupils with DS in groups of mixed ability.

For pupils with DS, an approach that incorporates and focuses on singing is recommended. Singing should focus on the development of the voice as an instrument, on the use of songs for supporting musicality, and on developing accuracy. Given the high incidence of HI in children with DS and its impact on phonation, provision should be made for supporting vocal development, especially during puberty. In particular, visual feedback should be used to enable knowledge of results (Welch, 1998; Welch, Howard and Rush, 1989), especially in those with HI (Welch et al., 2015). Although the present study relied on gestural and simple visual forms of feedback, the availability and accessibility of mobile technology make this a useful tool for supporting accuracy of pitching (e.g. Lin, Anderson, Hamzeen and Lui, 2014).

Although a song-based approach was used in this study, such approaches could be augmented or integrated with other musical programmes that do not rely upon singing. This may be particularly valuable for younger children with DS, or those with DS and additional or complex learning needs. For example, participation in rhythmic musical activities can support and underpin spoken perception and production (Fujii and Wan, 2014; Hausen et al., 2013; Jackson et al., 1997; Schlaug et al., 2008; Stahl et al., 2013). Furthermore, non-vocal musical skills may transfer to linguistic learning (Patel, 2014). An approach that may be beneficial for people with DS is the *Sounds of Intent* framework. This enables the tracking of musical development across a range of skills, and links clearly to early stages of communication (www.soundsofintent.org). This approach also enables participation in musical activities through the use of assistive technology, making it accessible to groups with a wide range of needs. Furthermore, the programme is easily adopted by teaching staff (Ockelford et al., 2011). For older pupils with DS, or those of higher ability, such programmes could be integrated into a community-based approach to learning. Such informal teaching approaches allow easy differentiation within ensemble-playing, and can be tailored to the musical preferences of the group, enhancing engagement within classrooms (Evans, Bauchamp and John, 2015; Hallam, Creech and McQueen, 2016). Such approaches could still draw on developmental approaches to learning in order to monitor progress, and ensuring that the quality of teaching and learning remains central.

Finally, there needs to be opportunity to practice musical skills throughout the lifetime. Schools and service-providers should aim to integrate high-quality musical provision into the daily schedule. The easiest, and potentially most beneficial form of musical practice is singing. Regular opportunities for singing will support any musical skills taught in more formal sessions during

schooling. This will enable ongoing practice of a range of cognitive and speech skills, supporting any SaLT provision. Furthermore, regular engagement in group-based singing provides additional benefits for health and wellbeing (Clift, 2012; Welch et al., 2014), including those with intellectual disability (Dingle, Branden, Ballantyne and Baker, 2013).

In summary, the following policy-based recommendations are made in order for individuals with DS to benefit from potential transfer effects:

- musical provision should begin during early infancy in order to support early listening and auditory processing skills;
- there should be a focus on developmental music provision throughout schooling;
- teaching approaches should be based upon the development of musicality, rather than on knowledge or mastery of a single instrument;
- Where possible, singing should be used as a means of teaching musicality, and as a means of supporting vocal and linguistic skills;
- visual aids and mobile technology should be used to support the development of singing skills;
- opportunities for non-song based musical development should also be included in programmes, with the use of assistive technology, as appropriate; and
- regular opportunities to practice skills or engage in community music-making should be incorporated into daily schedules throughout the lifetime.

8.3.3 Recommended activities

The findings in the study suggest that the speech and musical abilities of the participants were commonly affected by perceptual and productive factors (Section 8.2.2) that are also common to people with DS. Based on these findings, the following music-based activities are recommended to support people with DS, regardless of age or cognitive ability. The activities are suggested to support the DS population in their ability to perceive temporal changes in music and speech rhythm and to support the processing of speech sounds and words; to support fluent speech production and accurate word production; and to support sustained, efficient vocal production.

Whole-body rhythmic activities

Whole body activities to the beat of music that engage the vestibular system can support and influence the perception of rhythm (Phillips-Silver and Trainor, 2007, 2008). Movement activities involving the vestibular system may therefore be critical in developing a sense of beat and rhythm, and such activities will be especially important for those at risk of HI.

Rhythmic clapping and tapping activities

Accuracy in beat entrainment requires co-ordination between the auditory-motor system and auditory feedback (Deutsch, 2013; Pfordresher, 2006), and precise motor-timing (Sowiński and

Dalla-Bella, 2013). Tapping tasks are strongly linked with faster processing of auditory signals Besson et al., 2007, 2011; Kraus and Chandrasekaran, 2010; Magné et al., 2006) and with enhanced phonological processing (Huss et al., 2011; Overy, 2000; Tierney and Kraus, 2013). Rhythmic tapping activities also engage the sensorimotor network that controls orofacial and articulatory movements which are necessary for speech (Fujii and Wan, 2014; Schlaug et al., 2008; Wan and Schlaug, 2013; Wan et al., 2010). It is argued that in MIT, the sensory experience of rhythm (e.g. by a therapist tapping a rhythm on the left hand) may prime the brain for speech production (Wan and Schlaug, 2013; Schlaug et al., 2008). Fuji and Wan (2014) argue that synchronisation to a pulse is sufficient to engage the neural networks for both speech perception and for speech production. Furthermore, emerging research indicates that the motor systems that control speech are linked to those that control motor movements (Chu and Barlow, 2016). Tapping tasks may therefore support the rapid processing of signals that underpins speech perception, and support production.

Behavioural studies indicate that rhythmic activity may support both perception and production of words. Imitating the clapped syllabic pattern of words has been shown to facilitate perception and production of multi-syllable words in people with moderate learning disabilities (Jackson et al., 1997). Paced rhythmic production supports fluency of speech (Alm, 2004; Thaut, 2008): practice to a metronome may therefore alleviate some of the anxiety-responses associated with stuttering, and allow accurate practice of production. Additionally, the ability to tap to a beat is linked to singing ability (Dalla Bella et al., 2015). Potentially, if an individual is to benefit from singing activities in terms of speech, they will also need to develop their beat-tapping skills.

Singing

As neurological changes are typically specific to the type of musical training (Habib and Besson, 2009), singing training may be a suitable vehicle for supporting processing of speech sounds and words. Singing may facilitate the segmentation of new words into syllables (Kolinsky et al., 2009; Schlaug et al., 2008), and reduce cognitive load by ‘chunking’ and grouping sounds (Rinta and Welch, 2008; Schlaug et al., 2008). Wan et al. (2010) and Thaut (2008) argue that singing may enhance perception of speech sounds for people with developmental speech difficulties, and contribute to enhanced speech production as a result of physiological overlap and greater interhemispheric connection.

There is evidence that singing may be beneficial for people who stutter (Alm, 2004). Although underlying causes of stuttering are uncertain and may differ between individuals (Kent, 2000), developmental stuttering is associated with deficits in limb-motor timing (Olander et al., 2010). Given that developmental stuttering is prevalent in people with DS (Howell, 2010; Kent and Vorperian, 2013; van Borsel and Vandermeulen, 2008) people with DS may benefit from the

potential of singing to support motor timing. Although singing is produced at a slower rate than speech, it requires greater control of motor-planning and precision than speech (Jeffries et al., 2003). Different brain regions control singing than those used for speech (Jeffries et al., 2003; Riecker, Ackerman, Wildgruber, Dogil and Grodd, 2000). However, evidence from neuroimaging studies by Jeffries et al. (2003), shows that the singing of words activates regions associated with control of motor timing for oral articulation and laryngeal control, which may confer benefits for people who stutter. If words can be produced fluently in the context of singing, then it allows individuals to develop and practice motor plans for speech. This may ultimately reduce the planning load for speech production.

The act of singing may support the physiological production of speech and voice (Rinta and Welch, 2008; Wan et al., 2010). The nature of sustained phonation in singing (Schlaug et al., 2008) will allow practice of efficient vocal production and prolonged respiration. This has implications for the control of voice in producing voice contrasts in speech and for prosody. Rinta and Welch (2008) argue that singing may be particularly beneficial in the remediation of speech and voice disorders in pre-pubertal children. Puberty may also be a critical time in the voice production of DS individuals (Albertini et al., 2010). Early training in soft onset and relaxation may reduce the effects of hypotonia in the vocal muscles, and help prevent or reduce symptoms of habitual MTD. Given the high incidence of HI in children with DS (Maris et al., 2014) and its potential impact on phonation (Coelha et al., 2015; Das et al., 2013), it is of particular importance to support vocal production in children with DS, whose voices will still be developing.

Teaching considerations

Sensory feedback is essential in learning to pitch (Berkowska and Dalla Bella, 2009b; Keough and Jones, 2009; Pfordresher, 2006; Pfordresher and Dalla Bella, 2011; Dalla Bella et al., 2015) and in motor-timing tasks (Palmer, 2013; Phillips-Silver et al., 2010; Thaut, 2008). People with DS will benefit from visual or tactile feedback during learning to support difficulties in auditory feedback. This may support those with reduced hearing acuity in particular.

A repertoire of music-based activities is recommended, as the potential benefits conferred from one activity may complement or develop those of another. For example, accuracy in pitching involves the same neural timing networks as those required for accurate beat-timing (Dalla Bella et al., 2015). If training in either beat or singing are to support the motor-timing mechanisms, it would be beneficial to incorporate both skills.

Group learning may benefit people with DS in a number of ways. Social learning is especially beneficial for rhythmic development, especially when the participant models an adult (Deutsch, 2013; Miendlarzewska and Trost, 2013). Social learning allows the practice of social

communication skills such as listening and turn-taking (Gerry et al., 2012; Groß et al., 2010; Honing et al., 2015). It also supports a range of skills that underpin learning, such as social-emotional development (Phillips-Silver, 2009; Kirschner and Tomasello, 2010) and a range of executive functions (Miendlarzewska and Trost, 2013).

Special consideration must be given to supporting STM and cognitive processing in order to support processing, vocal production, learning and behaviour. Songs will need to be taught and developed in suitably short sections. As singing requires the ability to attend to linguistic and musical features simultaneously (Gordon et al., 2010; Racette and Peretz, 2007), tasks should be divided into separate features to reduce cognitive load. Simple visual notation, gesture or Makaton would further support the learning process and reduce demands on auditory STM. A reduction in cognitive load may reduce the degree of voice perturbation that is associated with increased stress (Schellenberg and Weiss, 2013; Yap, 2012). In combination with a system of visual feedback for supporting voice and gentle onset exercises (e.g. the use of a kazoo), this may help prevent the early development of habitual MTD (Ogawa, 2013; Lee and Son, 2005). Finally, a reduction in processing demands may reduce behaviours associated with task overload, including task avoidance and poor attention. This type of learning condition will allow for precision in training (P), emotional connection (E) and attention (A). In singing and speech, there are shared neural and physiological resources that demonstrate an overlap between the domains (O) (Christiner and Reiterer, 2013; Patel, 2014). As outlined in Patel's OPERA hypothesis (2011), these are the conditions for learning that are required before benefits may transfer from the musical domain to speech. The group demonstrates an ability to transfer learning from song to speech, despite the cognitive effort that is required (Racette and Peretz, 2007; Wiesniewski et al., 2013) and their severe learning disabilities. All else that is required to fulfil the criteria is repetition (R).

8.4 Critical evaluation of the research and future direction of research

The study's aims were to establish the levels of cognitive development and abilities in speech, rhythmic-motor and singing tasks of a group of young adults with DS and SLD; to examine to what extent they could take part in and learn from a six-week programme of song-based musical activities; and to determine whether singing and music activities might be an appropriate approach to support their speech perception or production. A battery of perceptual and productive assessments were used, most of which were suitable for the participant's levels of ability. Recordings were made of each participant's singing and speech that allowed detailed analysis of their voice production, phonetic production and prosodic production in tasks of increasing cognitive load. The programme of teaching was suited to the group's abilities and interests and enabled

them to progress towards learning aims (Appendix 3, p. 340). All participants made progress in individual sessions and over the course of the teaching programme (Appendix 3 pp. 342-355). A thorough and detailed analysis of the rhythmic-motor abilities of the group and of their vocal production in speech and singing tasks has identified the degree of similarity between their productions in speech and music. It has also identified the key physiological, cognitive and behavioural difficulties that affected their production in each domain.

The current study was extensive in scope and design, and produced a rich body of data relevant to the participants' speech and musical abilities. Not all the data were suitable for detailed analysis, as a result of environmental noise. However, the assessments and ongoing field recordings produced sufficient high-quality data to answer the questions of this study. In the researcher's opinion, the study achieved its aims. However, there are a number of strengths and weakness associated with this study that may be addressed in future studies. These are discussed below.

8.4.1 Positive aspects of the research project

Case study approach

The case study approach allowed detailed analysis of the speech, voice and music data. This level of analysis would not have been achievable in a large-scale study. This would have been at the detriment of the voice analysis, in particular, as no automatic readings of voice perturbation were possible for aperiodic samples of voice. The case-study approach also allowed inclusion of participants with HI and behavioural difficulties, who may typically have been excluded. Although this makes it difficult to isolate causal factors, their inclusion highlights the complexities that affect learning within the wider DS group who have the same level of cognitive impairment (i.e. SLD). The data highlight the heterogeneity of the group, even within this sub-category of the Syndrome.

Teaching sessions

Although the teaching programme was two weeks shorter than intended, it was suited to the group's abilities and enabled participants to meet their learning aims (Appendix 3, p. 340). The use of visual teaching methods during sessions was especially beneficial for Robert, who has a moderate-severe HI (Chapter 7). The use of gesture and physical aids was also a useful gauge of conceptual understanding of pitch. The success and limitations of individual teaching sessions and activities for all four participants are described in more detail in Appendix 3 (pp. 340-355) and in each case study (see Sections 4.1.5, 4.3; 5.1.5, 5.3; 6.1.5, 6.3; 7.1.5, 7.3).

Breadth of research

The study drew on a large body of research in order to examine both musical and speech domains. The study has contributed to our understanding of musical abilities, phonatory difficulties, and motor-timing abilities in people with DS; and to our understanding of song-learning ability in people

with DS and SLD, including those with HI. It has added to the literature concerning the overlap between motor-timing deficits and speech abilities in the DS population and in other populations with speech and rhythm deficits. To the author's knowledge, it is the only study to have examined the connection between musical, speech and cognitive abilities in the DS population, and the first study to examine the impact of cognitive load on their voice production and behaviour in the group. The implications of implementing a programme of music for people with DS, regardless of age or cognitive ability, have been outlined above.

Depth of analysis of phonation

The analysis of the data has required the careful selection and preparation of over 200 individual samples of voice production per participant in order to reject contaminated samples and prepare samples for automatic voice analysis. The analysis of the group's voice production required the implementation of two programmes of voice analysis (Praat: Boersma and Weenink, 2011; and Sonneta: Mint Leaf Software, 2015) in a bid to capture numerically the changes in voice quality that were audible within tasks and between tasks. Visual examination of speech samples was a useful addition to the numerical data, as it revealed the prevalence of interharmonics and diplophonia. Inter-rater reliability measures indicated that the methods used for automatic perturbation measures and for manual measurements of interharmonics were reliable to a statistically signifiant level ($p<0.001$: Chapter 3, Table 3.2). Although measures were less reliable for Andrew (Table 3.2), this was accounted for by a difference in the reading of the duration of interharmonics in one vowel, to the order of 12 ms. In part, this may be attributable to the complexity of the waveform in his atypical productions. Additionally, the poor time-resolution of the narrow-band spectrograms and the gradual reduction in the intensity of interharmonics make it difficult to ascertain a clear and consistent cut-off point (see Figure 3.12). A reliable automatic measure would be needed to overcome these limitations.

In their small-scale study of voice production in DS adolescents, Lee et al. (2009) manually adjusted sudden pitch changes in Praat. They also excluded onset and offset of vowels: although an increase in perturbation is typical at onset and offset (Fourcin and Abberton, 2008; Fourcin, 2009). Lee et al. (2009) concluded that the perceived voice qualities of their participants did not reflect laryngeal parameters. This study challenges that view. The participants in the present study showed an elevated degree of perturbation, which is typical of people with dysphonia (Fourcin and Abberton, 2008; Schaeffler et al., 2015). Given the high degree of perturbation at the onset of vowels in all participants in this study, and the reduced presence of interharmonics mid-vowel, the perception of voice quality may well depend upon the incidence, intensity or duration of interharmonics at onset. Future studies should use whole vowel samples in DS and control groups to test this. For larger-scale studies, automatic acoustic measures would be needed that can capture the presence of additional interharmonics (such as the Subharmonics-to-Harmonics ratio,

Sun, 2002) and diplophonia (such as the algorithms developed by Aichinger, 2013, 2015). This would enable examination of a larger sample size and statistical analysis.

8.4.2 Limitations of the study

Assessment environment and delivery

The assessment and teaching environment was not conducive to listening, sustained attention, or the recording of voice samples to a standard that allowed reliable comparison with other studies. Excessive noise can contribute to perceptual difficulties and overload phonological working memory (Schellenberg and Weiss, 2012); a quieter teaching environment may have elicited different results in perceptual tasks, motor tasks and in learning. A clinical setting with a controlled noise level may be required for voice recordings in order to draw clear comparisons with other studies. However, the setting was naturalistic, and the study has greater ecological validity as a result.

Suitability and reliability of assessments

Not all tasks were suitable to the abilities of the group. In perceptual tasks, the group was unable to reliably respond to tasks requiring judgements of ‘same’ or ‘different’, and gave unreliable responses to the prosody perception tasks. In some tasks, participants demonstrated task-avoidance behaviours or inattention. It is also not certain whether or to what extent HI affected the performance on some tasks. Despite these limitations, performance measures were able to indicate perceptual abilities in music tasks.

The study would be enhanced by the inclusion of additional indicators of cognitive abilities to supplement DAP, and by further physiological measures. Up-to-date hearing assessments would be required to confirm the extent of HI on perception and production. Although auditory processing difficulties are reported in DS (Groen et al., 2008; Keller-Bell and Fox, 2007; Marcell et al.,1990; Marcell, 1995; Pekkonen et al., 2007), a direct measure of this would help to clarify the nature of any difficulties, and to distinguish the effects from cognitive difficulties. Respiratory capacity measures would also confirm the extent of respiratory difficulties, and changes in pressure when speaking. Electroglottographic measures would confirm the extent of difficulties in phonation, and measures of muscle tension would further clarify how physiological responses relate to the spectrographic data. This additional information would allow firmer conclusions to be drawn between physiological abilities and cognition and perception.

Much of the existing data on beat tapping skills relies upon the finger-tapping tasks, with a few exceptions (Patel, 2008). The participants in this study used claves, which will require a different range of motor movements. This may exaggerate or reduce differences between finger-tapping tasks, in comparison to other studies.

Manual calculations of pitch change and voice

As a result of the voice difficulties, it was not always possible to objectively measure pitch changes. As a result, pitches were approximated (see Chapter 3, Section 3.5.4) in order to measure the perceived changes in melody and song. Additionally, this resulted in a different method being used for Andrew and Kerry than for Rachel and Robert. Alternative methods of calculation would need to be developed for future studies: such as a Subharmonic-to-Harmonic ratio (Sun, 2002).

The timing between beats and to the beat was calculated manually. Although the results were reliable at a statistically significant level ($p<0.001$, Table 3.2), the use of an automatic measurement would allow for the collection and analysis of a larger body of data. For example, tasks could be repeated to track consistency and learning, and to compare performance in different modes (e.g. finger tapping, tapping with a stick). This would enable more accurate comparison between DS subjects, and more detailed examination of the types of factors that may affect their performance.

8.4.3 Recommendations for further research

The study has suggested a strong association in people with DS between motoric beat timing abilities and the ability to perceive, remember and act upon auditory information. Further studies should investigate these links further, to isolate the role of each parameter, and to investigate the effects of training and experience on both motoric ability and perceptual response. Beat-timing skills in this group appear to be linked to reading abilities, as in other populations: this too could be further explored.

The data have revealed a clear divide in voice quality within the group. Future studies could examine further the prevalence of diplophonia and the link between the number and intensity of interharmonics and perceptual correlates. It has also indicated that cognitive load is the primary trigger for increased perturbations, possibly as an emotional response to perceived task effort. Further investigation would be needed to confirm this link, using a wider range of tasks. Studies of voice quality may also examine whether changes in perceived voice quality may act as a marker of anxiety or stress in a population that does not necessarily verbalise this.

The group results suggest that divided attention affected production in speech and singing tasks. The extent to which people with DS can learn when task demands are reduced could be further investigated, in order to measure their learning potential. Although the group showed an ability to transfer previous learning from speech to song, a lengthier teaching programme would be required

in order to examine the extent to which new learning may transfer when optimum learning conditions are met.

8.4.4 Conclusion

To the author's knowledge, this is the first study to have compared the abilities of people with DS in speech and musical tasks, and the first to examine their abilities to learn songs and to transfer learning from singing to the speech domain. To the author's knowledge, it is only the second study to examine diplophonia in DS (see Beckman et al., 1983) and the first to examine the impact of cognitive load on their voice production. It has added to our knowledge of their musical abilities and contributed to the research investigating motor-timing in DS participants with speech and language difficulties. It is clear from this study that the DS group demonstrate an ability to learn new songs and the potential to transfer learning from song to speech. It is also evident that reduced auditory STM and cognitive load affect a wide range of behaviours, with consequences for their participation, emotional response, rate of learning, speech production and voice quality. It remains to be determined whether musical training can reduce the demands in tasks requiring auditory memory or working memory in the DS population. However, any reduction in the effort that is required to perceive, learn and reproduce speech may relieve the immediate load. This will make the imitation of speech less stressful and, potentially, reduce the associated tension that affects voice production.

The study recommends that structured musical activities should be a part of the lives of people with DS, especially for infants and children. Ideally, a curriculum such as one based on the Kodaly approach (Brindle, 2005; Houlahan and Tacka, 2008) would be most suitable. Such pedagogy integrates physical learning with simple visual notation and games that allow for people of a wide range of learning abilities to be actively included. The current government emphasises the potential of music to transform lives (DfE, 2011) and the revised National Curriculum (DfE, 2013; DfE, 2014) encourages opportunities for primary and secondary pupils to take part in music. However, pedagogical issues, such as lack of teacher knowledge or confidence in teaching music (DfE, 2011; Henley, 2011; Garrett, 2014; Welch and Henley, 2014), and lack of local music provision (Creech, Saunders and Welch, 2015) will undoubtedly limit how well musical provision can meet the needs of people with DS, or with other learning disabilities. Nevertheless, any opportunity that allows people with DS to take part in musical activities should be exploited. Singing and making music as part of a group can foster social inclusion (Welch et al., 2014) and the act of singing is believed to confer a range of social, emotional and health benefits including enhanced self-confidence and wellbeing, respiratory function and improved immune functioning (Clift, 2012; Hallam, 2010; Welch et al., 2014). Music and singing offer a wealth of benefits to people with DS, including benefits to those who might never fully develop language or verbal communication. An

emphasis should be placed on delivering high-quality music education that is suited to the learning needs and abilities of people with DS. If such learning began in the early years, the consequences for the DS population will reach far beyond that of simply singing or banging a drum.

Appendices

Appendix 1: Ethics documentation.....300

Appendix 2: Sample tests materials for music production and perception.....……..…….323

Appendix 3: Teaching sessions, evaluations and teaching materials….....340

Appendix 4: Summary of research for MSc (LACIC): *An investigation into the effects of singing tuition on aspects of intelligibility in young people with speech impairments and learning difficulties*.....359

Appendix 1
Ethics documentation

Ethics consent for pilot study 301

Request for amendment to ethics consent.....302

Consent to amendment.....303

Letters and information sheets for participants 298

Part 1 of 4: Letter to the participant.....304

Part 3 of 4: Participant's consent form 307

Part 4 of 4: Pictorial information sheet 308

Letter and Information sheets to parents/carers 310

- Letter 1.....310
- Letter 2.....311

Information sheets for parents/guardians.....312

Ethics consent for pilot study

The University Of Sheffield.

Department of Human Communication Sciences
31 Claremont Crescent
Sheffield S10 2TA UK

Head of Department
Professor Joy Stackhouse, PhD., F.R.C.S.L.T., C.Psychol., A.F.B.Ps.S

Telephone: +44 (0) 114 222 2418/ 2402/ 2405
International: +44 (0) 114 222 2418
Fax: +44 (0) 114 273 0547
Email: hcs-support@sheffield.ac.uk
http://www.shef.ac.uk/hcs

19th May 2011

Dear Tracy

Title: A pilot study to test the methods and feasibility of a project investigating the effect of singing tuition on speech

Thank you for your submission to the HCS Research Ethics Committee. The committee has reviewed your submission and supporting documents and grants you approval to commence the research.

We hope your project proceeds smoothly

Yours sincerely

Prof R Varley
Chair of HCS Ethics Committee

Request for amendment to ethics consent

Dear Professor Varley,

Thank you for your letter dated 19th May 2011 confirming Ethics Approval for my Pilot Study.

Due to circumstances beyond my control, the school that had expressed an interest in participating is no longer able to, which means that I must recruit participants from other organisations. This process will take time, and may also require me to recruit slightly older participants.

Therefore, I am writing to ask if it is possible to approve some changes to my application to reflect the changed circumstances, specifically:

1. To extend the dates of the study on my application, from 1/6/2011-31/7/2011 to 1/6/2011 – 31/12/2011?
2. To increase the upper age range of potential recruits to 25? Whilst it may still be possible to recruit 11-17 year-olds for the Pilot, it is more likely that I will successfully recruit 17-25 year-olds as there are several care homes catering for this age range within my locality.
3. To include 'care organisations' and 'youth clubs' as a potential source of participants, alongside schools and colleges?

In all other respects, the Pilot study remains the same.

Thank you for your assistance with this.

Yours sincerely,

Tracy Jeffery

Consent to amendment

From: Rosemary A Varley (email address removed)
Subject: **Re: FW: letter for Professor Varley**
Date: 30 June 2011 14:39:10 GMT+01:00
To: Tracy Jeffery (email address removed)

Dear Tracy

thank you for your letter regarding changes to your research plan. As you will be recruiting older (vs. younger) participants and there will be no changes to the research protocol, I can confirm that there is no need for further ethics review. I also confirm that the change in the dates of the study is also unproblematic.

I would be grateful if you can send revised documents to Kathryn for storage in your file. This would include a revised application form (recording new study period, new upper age limit, new recruitment locations), and if you make any changes in recruitment letters, information sheets and consent forms - copies of these are also lodged in your file.

Best wishes

Rosemary

Part 1 of 4: Letter to the participant

To: the participant

Hello. My name is Tracy. I am a teacher and a student. I am writing to ask if you would like to take part in a research study. The study is so that I can find out whether or not you enjoy taking part in singing lessons with me, and how easily and quickly you can complete some tests.

People who take part in the study will be asked to take part in some tests and some singing lessons. The information sheets explain more, and your parents and carers can tell you more about it. Please talk to them about it: they will answer your questions.

When you have decided if you would like to take part or not, please tell your parents/carers: they will show you what to do next,

Thank you,

Tracy Jeffery,

PhD student, University of Sheffield

Part 2 of 4: Information about the Study

What is the study about?

The study is about speaking and singing. I want to know if singing lessons help people to talk more easily.

Why are you doing this research?

There are 3 reasons I am doing the research project:

1. I would like to find out whether my singing lessons are easy to take part in and fun for people of different ages.
2. I want to know how easy or difficult it is for people your age to complete some tests, and how long it takes
3. I want to know if the way you talk or sing changes when you have finished the singing lessons.

Why are you asking me?

I am asking you because your carers/parents believe you enjoy singing and music, and because you have the same types of difficulties with speech and language as some other people.

Do I have to do this?

No. You can choose on your own; but your parents/carers may help you decide, and can tell you more about the study.

GUARDIAN: I have checked with the participant and they understand that participation is voluntary __ (initial)

Is this bad or dangerous for me?

No. If you are unhappy about taking part in the tests or singing lessons, you can stop. A familiar person will also be present to make sure that you feel OK.

Do I get anything for being in the research?

No.

Is everybody going to know about this?

Your parents and researchers will be able to look at the tests and listen to the recordings of you singing, making music and speaking, but they will not give these, or play the recordings to anybody else. When Tracy writes about what you did, she will not use your real name or personal details: nobody will be able to tell who took part.

Can I choose not to be in the research? Can I change my mind?

You can say 'no' now if you do not want to take part. You can change your mind at any time, if there is any part of the study you do not want to take part in.

Who can I talk to or ask questions to?

You can talk to your parents and carers about this. They will help you.

What is going to happen to me?

If you say 'yes', then this is what will happen:

1. You will take part in some tests. Each test will last approximately 15-30 minutes. She will ask to see you for 3 x 1 hour sessions before you start singing lessons, and for 3 x 1 hour sessions after you have finished singing lessons.
2. You will take part in Singing Lessons with some other young people from your school or organisation: there will be no more than 6 people in the group. You will take part in singing lessons once a week for one hour and a quarter; you will be given a break in the lesson.
3. Tracy will record your voice sometimes when you are speaking and singing.
4. Tracy will write about how easy or difficult you found the assessments and singing lessons. But when she writes, she will keep your name and personal details secret so that nobody will be able to tell that you took part.

GUARDIAN: I have checked with the participant and they understand the procedures _____(initial))

Will you tell me the results?

I will be able to tell you your results if you want them, and your parents and carers.

If you choose to be part of this research I will also give you a copy of this paper to keep for yourself. You can ask your parents to look after it if you want.

What do I do now?

- Read or listen to the information sheets.
- Think about whether you want to take part in the study.
- Talk to your parents/carers about it. Your parents/carers will ask you if you want to take part.
- Decide Yes or No.
- Complete the Consent form, or ask your parents/carers.
- Return the form to me.

Part 3 of 4: Participant's Consent Form

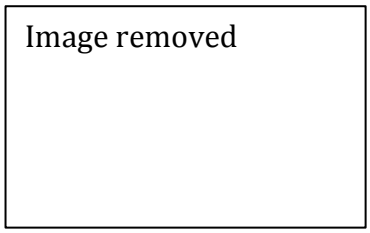
Participant's Consent Form.

Research Project: A study to test the methods and feasibility of a project investigating the effect of singing tuition on speech

Lead Researcher: Tracy Jeffery

Participant Identification number for this project .....

Do you want to take part? Yes/No.



If yes,

Please initial the box and sign.

1. I agree to take part in the Assessments
2. I agree to take part in the Singing lessons
3. I agree to be recorded
4. I know that I can choose to be in the research study or choose not to be in the research study. I know that I can stop whenever I want.
5. I have read this information (or had the information read to me) and I understand it.
6. I have had my questions answered and know that I can ask questions later if I have them.
7. I understand any changes to this will be discussed with me.
8. I agree to take part in the research.

1. Print name of participant …….

Signature of participant: ……..……. Date:…….

2. Witnessed by …….. …… (PRINT NAME)

Signature of witness……
Date……

Part 4 of 4: Pictorial Information Sheet

nb. Images have been removed for copyright reasons

1. Hello. My name is Tracy. I am doing a study:

The Study is about speaking and singing

2. If YOU want to take part in the study, YOU will:

Take part in some tests and singing lessons

3. In the study, Tracy will:

Record you speaking and singing, give you singing lessons, and write about you.

4. Other people will: read about how long it took to complete assessments and whether you enjoyed and learned from the singing lessons

5. Your name and details will be disguised so nobody will know who took part and your details will be kept confidential.

NOW YOU CHOOSE:

6. Do you want to know more? Do you want help deciding?

Talk to your family or carers. Learn more about it.

7. When you are ready, choose ‘yes’ or ‘no’

Choose YES or NO.

Tell your family or carers.

Thank you.

Letter and information sheets for parents

Letter 1

Dear Parent/Carer,

I am a music facilitator, and a PhD student with the University of Sheffield. I am researching the effects of singing tuition on speech and communication in children and young adults with speech, language and learning difficulties. I am writing to ask whether your son/daughter might be interested in participating in a Pilot study. This would involve your child taking part in a number of speech, cognitive and musical assessments lasting approximately 2 hours, followed by 6 weeks of singing tuition.

If you are interested in learning more, please return the reply slip to your school. I will then forward further information about the study that should answer most of your questions and provide you with the opportunity to ask me questions before you and your child make a decision.

Yours sincerely,

Tracy Jeffery, MSc. (LACIC), PGCE

Music facilitator/PhD student, University of Sheffield

……

Re. PILOT study for research study into singing and speech

Please return this slip to your child's school by (DATE)

Name of child ……

☐ I DO NOT wish for my son/daughter to take part in the study

---------OR----------

☐ Please send me further information about the study.

Signed……. (parent/carer) Date……

Letter 2

Dear Parent/carer

Thank you for your interest in this research study. I enclose the following documents:

1. Child's pack: Introductory Letter ; Information sheets for your child; Assent form
2. Guardian's pack: Information sheets (9 pages); Consent Form (1 page). PLEASE NOTE: It is your child's decision whether he or she would like to take part in the study, but as legal guardian, you have the right to give or withhold consent.

Please take your time to read through the information, and to ask any questions. You may contact me with any queries that you or your child may have (see contact details below).

If your child or you decide NOT to take part, please initial the Assent form (Section B) and return it to the school. Please note that no personal details will be passed on to the researcher.

If your child wishes to take part, please complete the Assent form (Section A), signed by your child, and initialled by you where shown, and the Guardian's Consent form. I will then advise you of the likely timeline of the study.

When you and your child have reached an informed decision, please return the relevant forms to your son/daughter's school by (DATE).

Thank you for taking the time to read this, and I look forward to hearing from you.

Yours sincerely,

Tracy Jeffery, MSc. (LACIC), PGCE

Music facilitator/PhD Research Student, University of Sheffield

Contact details: (email address removed); telephone (mobile number removed)

Information sheets for Parents/Guardians.

This document provides information on the following questions relating to the study. If you have any further questions, or queries, please contact Tracy on (email address removed).

Information sheets for Parents/Guardians. 312
ABOUT THE RESEARCHERS and the INVITATION 314
Why this school? Why my child? 314
What details do the researchers already know about me and my son/daughter? 314
What experience or qualifications does the lead researcher have to ensure my child will receive the best education and care should I consent? 314
What other precautions are in place to protect my child? 314
Does my child get anything for being in the research? 314
ABOUT THE RESEARCH STUDY 315
What does the Pilot study aim to achieve? 315
Why conduct a Pilot study at all? 315
What is the background to the Pilot and Proposed studies? 315
What is the format of the Pilot study? 315
What are the benefits to my son/daughter taking part in the study? 315
What are the disadvantages or risks of my son/daughter taking part in the study? 315
ABOUT THE ASSESSMENTS 317
What is the purpose of the assessments? 317
Can you give me an estimate of how long the assessments last? . 317
ABOUT THE SINGING TUITION 317
What will my child do in singing sessions? 317
How long do they last? 318
What are the aims of the sessions? 318
What will be the impact of the singing lessons for my child? 318
ABOUT THE DATA, CONFIDENTIALITY and ETHICS 318
What data will be collected and how will it be used? 318
Confidentiality: what measures are taken to ensure that my child's identity is protected during and after the study? 318
Ethics Approval 319
May I or my child have access to the data or results of the assessments? 319
Who else will have access to the data? 319
How will the data be stored? 319

If I give permission, how might this data be used in future research?319
ABOUT CONSENT..... 319
Consent and Assent: Why are there 2 forms?..... 319
May I withdraw my child from the study, or may my child refuse to take part? 320
How do I withdraw consent?..... 320
WHAT DO I DO NOW? 320
What happens next?..... 320

ABOUT THE RESEARCHERS and the INVITATION

Who is conducting the research?

The lead researcher is Tracy Jeffery, who is conducting the research in part fulfillment of a PhD, being offered by the Human Communication Sciences Dept. at the University of Sheffield. The research is supervised by Dr. Sandra Whiteside, at the University of Sheffield.

Why this school? Why my child?
The Pilot and the planned research are aimed at teaching children with speech and learning difficulties. The final study will aim to measure whether singing can be used to develop speaking skills. When approached, the school expressed an interest in learning more about the study and its aims. After initial consultation, your child was invited because he or she has ongoing difficulties in aspects of speech and learning, and has shown an interest in singing.

What details do the researchers already know about me and my son/daughter?
None. The researchers do not know who has received this letter. The school will only pass on your details IF you consent to take part. Any details that you do supply will remain confidential during and after the study (see Ethics section).

What experience or qualifications does the lead researcher have to ensure my child will receive the best education and care should I consent?
Tracy has been teaching and working with people with learning difficulties for over 5 years. She holds a teaching qualification (post-16) and an MSc in *Language and Communication Impairments in Children*. For over 5 years, she has been teaching music as a means to develop communication in young people with learning disabilities, and has worked as a music facilitator in primary and secondary schools.

What other precautions are in place to protect my child?
Tracy, and any other staff that might be involved in assessing or teaching your child will be fully CRB checked; a member of the school's staff will also be present if you, the school or the researcher require it.

Does my child get anything for being in the research?
No. However, participation is, and remains, voluntary; participants may withdraw at any time.

ABOUT THE RESEARCH STUDY

What does the Pilot study aim to achieve?
The pilot study aims to find out:

a) how long it takes the participants to complete the tests in the assessments for the main study.

b) whether all the tests within the assessment battery are necessary and appropriate to the participants' abilities.

c) whether any changes in singing/musical skills are observed after 6 weeks of teaching.

Why conduct a Pilot study at all?
The Pilot study is a smaller scale of the proposed study, and is designed to check whether the proposed study is feasible and appropriately designed. This will help the researchers decide whether to proceed with the proposed study, and whether any changes need to be made to the proposed study.

What is the background to the Pilot and Proposed studies?
Research has shown that singing and musical activities can help develop communication and speech skills, but there has not been much research into whether this helps people with speech and learning difficulties. For her MSc. study, Tracy researched the effects of singing on speech intelligibility in young people with speech and learning difficulties. She found evidence that singing lessons seemed to help young people with speech and voice difficulties to use their voices more efficiently. The proposed study will be designed to find out more about whether singing and musical development can be used to modify both speech and voice.

What is the format of the Pilot study?
The study is in 2 main parts:

1. Participants will be assessed on speech, cognitive and musical skills;
2. Participants will receive singing lessons over a 6-week period, once a week

What are the benefits to my son/daughter taking part in the study?
There ma**y** be no benefits from the assessments or from the singing tuition. It is hoped that they will enjoy and benefit from the experience, and that some might develop their singing and musical skills in the process.

What are the disadvantages or risks of my son/daughter taking part in the study?
The main disadvantage will be the duration of the assessments and the possibility of stress or anxiety that arises because of this. However, all steps will be taken by the researcher to ensure that your child is at ease, that assessments are kept to the

minimum duration, that breaks are given as needed, and that your child may withdraw, or will be withdrawn, if stressed.

May I withdraw my child from the study, or may my child refuse to take part?

Yes, you may withdraw at any time. You do not need to give a reason. Your child may refuse to take part in any part of the study.

How do I withdraw consent?

Contact your school; they will inform the Tracy.

What do I do if I have a complaint, or if my son/daughter has a complaint?

If you have any complaints or concerns regarding your child's participation in the singing lessons or assessments, in the first instance, please notify the school or Tracy so that these can be immediately addressed in order to ensure your child's well-being. If you have any further concerns or complaints regarding the study, please contact either the project co-ordinator:

Dr. Sandra P. Whiteside, Reader, Department of Human Communication Sciences, University of Sheffield, 31 Claremont Crescent, Sheffield S10 2TA.

Email: (email address removed) Tel: (telephone number removed)

OR

Head of Department: Professor Shelagh Brumfitt, Department of Human Communication Sciences, University of Sheffield, 31 Claremont Crescent, Sheffield S10 2TA.

Tel: (telephone number removed)

OTHERWISE

You can use the normal University complaints procedure and contact the following person:

Registrar and Secretary, Registrar and Secretary's Office, University of Sheffield, Firth Court, Western Bank, Sheffield, S10 2TN (telephone number removed)

ABOUT THE ASSESSMENTS

What is the purpose of the assessments?

The assessments are designed to help the researchers and understand how your child uses and understands language and speech, and to learn about their musical and singing skills. Some of the tests give scores that are comparable to IQ scores: this allows researchers to understand how your child performs in these tests relative to other children, and also shows if they are better at verbal tests (which rely on the use of language), or at non-verbal tests (which rely on the use of pictures, objects or symbols).

What types of assessments are used and what will my child be asked to do?

There are 3 main types of assessment: 1) verbal 2) nonverbal and 3) musical.

Verbal assessments are used to find out how well your child understands words, phrases and instructions, and how many words they can remember easily. Your child will be asked to describe pictures, to point to pictures, and to repeat sounds, words and phrases. These assessments will also be used to find out about your child's speech and speaking voice.

Non-verbal assessments are used to learn about how your child can complete tasks that do not use words. Your child will be asked to complete puzzles, and to draw a picture.

In musical tests, your child will be asked to listen to and compare short pieces of music and to say if they sound the same or different; and to copy clapped patterns, songs, and short extracts of songs. They will also be asked to sing their own favourite song.

Can you give me an estimate of how long the assessments last?

Assessments will probably take 2 hours in total. This can be broken down into sections with regular breaks as your child needs. If it is apparent that an assessment is too difficult, stressful or other wise inappropriate for your child, then the researcher will stop, give your child a break and move on to another test if appropriate and if your child agrees.

ABOUT THE SINGING TUITION

What will my child do in singing sessions?

Singing lessons will follow a routine: warm-up routines designed to relax key muscles and to stimulate deep, controlled breathing; vocal warm ups, vocal exercises, group singing and solo singing. Group activities will be used to develop basic music skills in a fun, informal setting, and to encourage participation and confidence. Mostly,

participants will be asked to copy, but some may also feel comfortable with demonstrating or leading activities once they are familiar with them. Participants will be encouraged to try all activities; but as participation will remain voluntary, they may choose to abstain.

How long do they last?
Singing lessons will be group activities of 4-6 participants and will run during the school day. Sessions will last for 75 minutes, with allowance for a 10 minute break after 40 minutes.

The sessions will run for one school term, beginning the week after the initial assessments.

What are the aims of the sessions?
The primary aim is to observe whether the participants find the activities easy to follow and enjoyable, and that the activities allow the participants opportunities to develop their singing skills.

What will be the impact of the singing lessons for my child?
Your child will be able to learn new songs and techniques for developing his/her voice. The focus for your child will always be on having fun by participating in the activities themselves, rather than on an end point. It is not possible to guarantee that he or she will improve in any aspect of singing, as this will depend on a combination of factors.

ABOUT THE DATA, CONFIDENTIALITY and ETHICS

What data will be collected and how will it be used?
Some examples of your child's speech and singing will be recorded as audio files: the recordings will be made to ensure that they are of suitable quality, that results are consistent, and to ensure that the recording technique is not intrusive. Once this is known, the recordings will be wiped.

Some of your child's responses during assessments will be written down as ticks against a checklist, or as words. All verbal assessments will be recorded to ensure accuracy.

Confidentiality: what measures are taken to ensure that my child's identity is protected during and after the study?
Information regarding the results of assessments and singing tuition will be treated confidentially by the researchers, and by any other staff (e.g. teachers) involved in the study. This information will not be passed on to third parties.

When the results of this study are published, no personal details will be given that would enable a reader to identify your child, you, or the school.

Ethics Approval
The research study has been approved by the Department of Human Communication Sciences Ethics Committee, University of Sheffield, to ensure that no participants are put at risk as a result of the study. Your school has given approval for this study.

May I or my child have access to the data or results of the assessments?
You may request to see the assessments data during the study, and may ask for a brief report regarding your child's singing and musical skills, after the final singing session. To do this, please email Tracy on: (email address removed), or contact her via your child's school.

Who else will have access to the data?
The researchers will have access to the original recordings and test results to check the results are of the required quality.

If you and your child consent, these results may be given to your school if requested. You and your child have may ask to see the data at any time, up until the end of the study (this is estimated to be July 2014).

How will the data be stored?
Recorded data will be stored anonymously on digital card, and as an AIFF/MP3 file on a password-protected laptop. Your child will be given a number and then a pseudonym to protect his/her identity, and recordings will be coded using this pseudonym. Anonymised paper assessments will be stored securely at the Principal researcher's home, and anonymised recordings will be stored securely on digital cards or audio discs at the lead researcher's home: all data will be locked in a filing cabinet. At the end of the study (estimated to be July 2014), the recordings will be wiped, and paper assessments will be shredded, unless permission is given to retain the audio for future use. In this case the audio data will be retained on the original recording medium and as computer files for 10 years from the date of your child's final singing lesson and/or assessment. At the end of the agreed period, the paper and digital materials will be destroyed by shredding and erasing, respectively.

If I give permission, how might this data be used in future research?
The data collected may potentially be of use in other studies that seek to find out more about the speech, learning and musical abilities of children with speech and learning difficulties. If you agree to your child's data being retained for future use, it will still be treated as confidential information by the researchers involved, and no personal details will be passed on or published.

ABOUT CONSENT

Consent and Assent: Why are there 2 forms?
There is a Consent form for you, and an Assent form for your child. The Assent form is for your child to complete. He or she has the right to decide whether or not to participate, but may not be fully aware of the consequences and therefore your consent is also required. As legal guardian, you may withhold consent even if your

child wishes to participate – it is your consent that determines whether or not your child will take part in the study**.**

May I withdraw my child from the study, or may my child refuse to take part?
Participation is voluntary and your child may withdraw at any time. You do not need to give a reason. Your child may refuse to take part in any part of the study.

How do I withdraw consent?
Contact your school; they will inform the researcher.

WHAT DO I DO NOW?

If you are satisfied that all your questions are answered, and have discussed this with your child then please:

- Complete the Child's Assent form and initial it where shown;
- Complete the Guardian's Consent form; and
- Return both documents to your child's school.

What happens next?
Tracy will inform you of when the study will begin.

Guardian's Consent Form

Title of Research Project: A pilot study to test the methods and feasibility of a project investigating the effect of singing tuition on speech

Name of Lead Researcher: Tracy Jeffery

Participant Identification Number for this project: **Please initial box**

1. I confirm that I have read and understand the information sheet explaining the above research project, and I have had the opportunity to ask questions about the project. ☐

2. I understand that my child's participation is voluntary and that he/she is free to withdraw at any time without giving any reason and without there being any negative consequences. In addition, should he/she not wish to answer any particular question or questions, s/he is free to decline.
   To withdraw consent, contact Tracy on (email address removed) ☐

3. I understand that my child's responses will be kept strictly confidential.
   I give permission for members of the research team to have access to my
   child's responses. I understand that my child's name will not be linked with
   the research materials, and that neither my child nor I will not be identified
   or identifiable in the report or reports that result from the research. ☐

4. I agree for the data collected from my child to be retained for use in this study, until approximately July 2014.

5. I agree for the data collected from my child to be used in future research, within 10 years from the date of my child's final singing lesson and/or assessment. I understand that the data will be safely destroyed after this time.

6. I agree to my child taking part in the above research project. ☐

_____	_____	_____
Name of Parental Guardian	Date	Signature
_____	_____	_____
Lead Researcher	Date	Signature

To be signed and dated in presence of the participant

Copies: *Once this has been signed by all parties the participant should receive a copy of the signed and dated participant consent form, the letter/pre-written script/information sheet and any other written information provided to the participants. A copy of the signed and dated consent form should be placed in the project's main record (e.g. a site file), which must be kept in a secure location.*

Appendix 2

Sample test materials for music production and perception

Sample test materials for music production and perception 323
Music production skills checklist 324
Pulse.....324
Pitching skills – all on ‘La’.....325
Singing songs.....325
Pitch production – visual support 326
Rhythm production and syllable segmentation – visual support 327
Test 13: Minimal pairs perception test 328
Slides 1-5: consonants.....322
Slides 7-11: vowels and dipthongs.....323
Test 14: Simple prosody perception test.....324
Test 26: Pitch and interval perception test.....332
Slides 1-6: pitch direction.....326
Test 27: Rhythm perception test.....333
Slides 1-4: training on identification of same/different using visual patterns.....333
Slides 5-7: same/different identification of rhythm patterns.....334
Slides 8-13: same/different pitch perception tests.....335
Test 28: 'Find the sound' interval and rhythm perception test..... 330
Beat Entrainment Task 1.....331
Rhythm Entrainment Task 2.....332

Music production skills checklist

The checklist was used to record gross motor movement skills and ability to perform each task independently (with aural stimuli) and with visual support (see samples in this Appendix).

Pulse:

	AURAL/Independently	VISUAL assistance
1. tap feet to pulse		
2. alternate feet to pulse		
3. walk to pulse		
4. both arms to pulse		
5. alternate arms to pulse		
6. clap to pulse		
7. shoulders – lift both		
8. shoulders alternate		
9. create movements to pulse		

Rhythm:	AURAL	AURAL_VISUAL
1. copy rhythm		
2. entrain to rhythm		
3. copy rhythm with visual support		
4. play from visual support		
5. play word rhythm – monkey		
6. say words to pulse		
7. clap words		
8. say and clap		
9. drum		

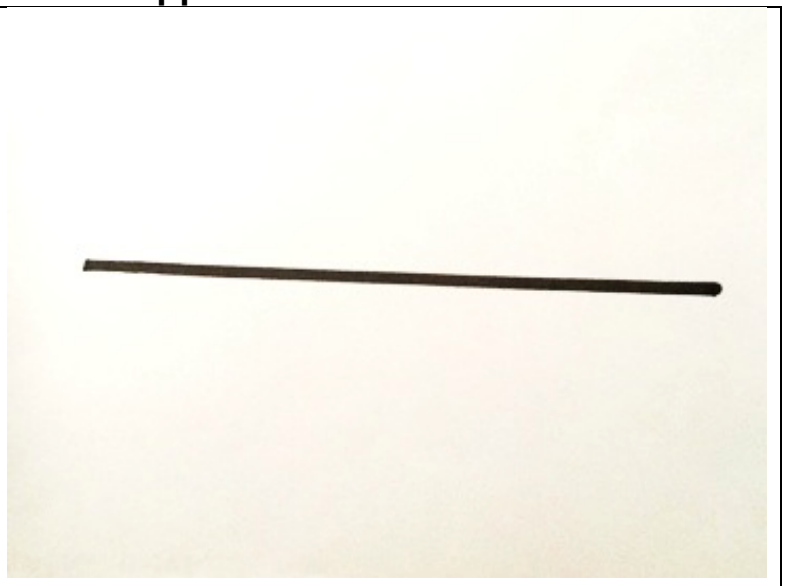
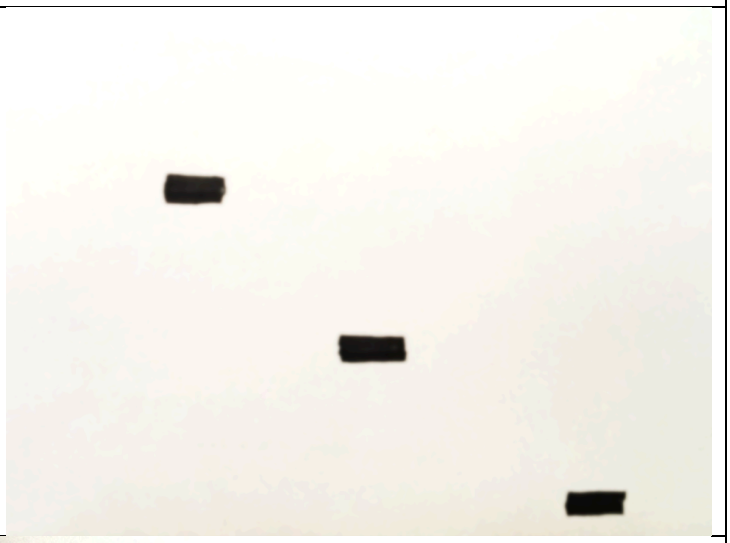
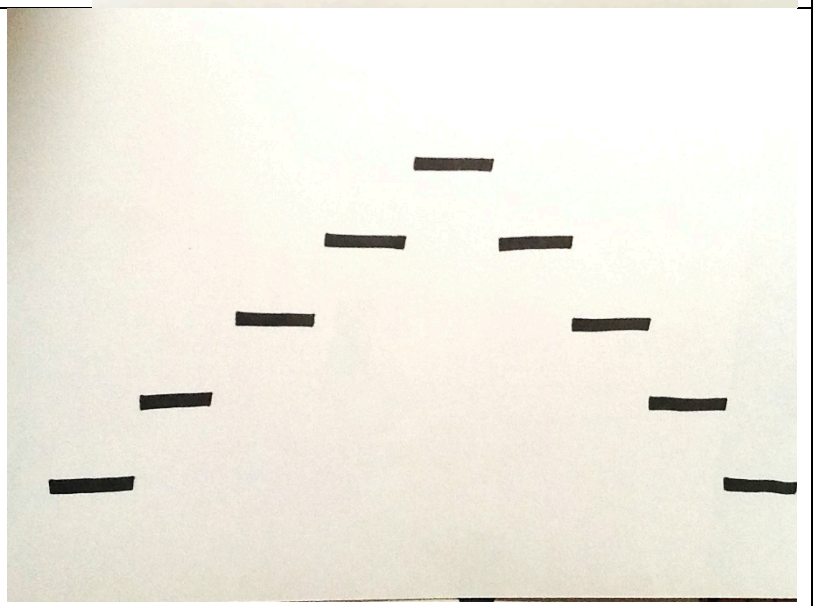
Pitching Skills – all on 'La'

	AURAL	AURAL_VISUAL
1. match note – middle C		
2. match note – G above/below middle C		
3. interval – so-me		
4. interval – me-do		
5. interval – so-me-do		
6. interval – so-do		
7. scale: 12345		
8. scale: 123454321		
9. interval – me-so		
10. interval – so-la-so		

Singing Songs

1. sing Happy Birthday	
2. repeat a new phrase – *aye aye aye* – words, melodic shape, rhythm	
3. without, then with Makaton	
4. perform song –	

Pitch production – visual support

Straight line representing a sustained note. This was modelled by the researcher who placed a finger at the left of the line and ran it along the line whilst singing 'la' on a sustained G3 note.	
Descending So-Me-Do. This was modelled by the researcher who placed a finger on each mark whilst singing the interval pattern of so-me-do on 'la'.	
Ascending and descending 5-note scale. This was modelled by the researcher who placed a finger on each mark in turn, whilst singing each note of the scale do-re-me-fa-so-fa-me-re-do on 'la'.	

Rhythm production and syllable segmentation – visual support

Image of Cat

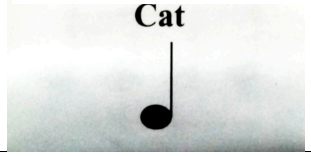


Image of Monkey

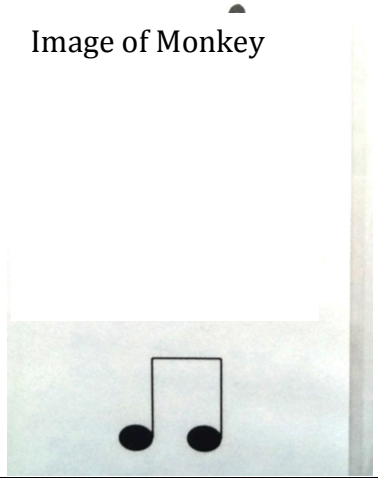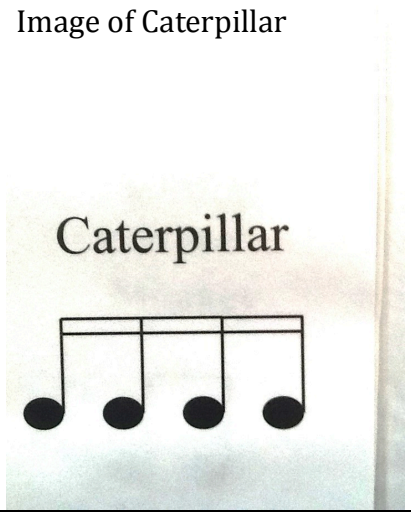


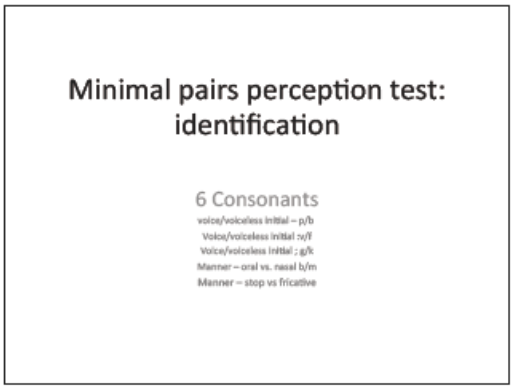
Image of Caterpillar

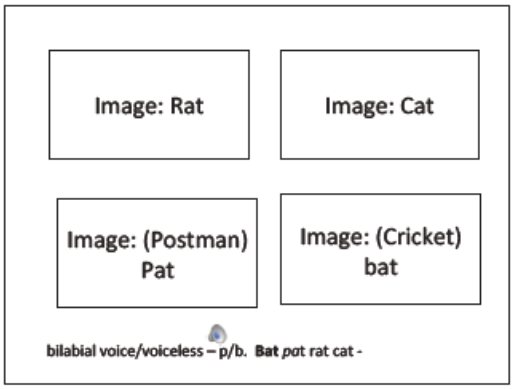


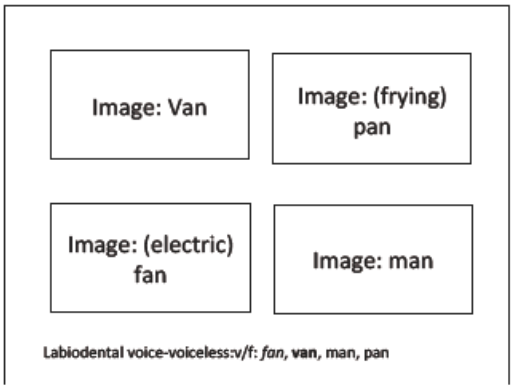
Images were used to support the participants in reproducing the spoken rhythmic sequence: 'Monkey, caterpillar, monkey, cat'. The musical notation was used as a prompt for clapping the syllables if participants were unable to do this aurally, unaided.

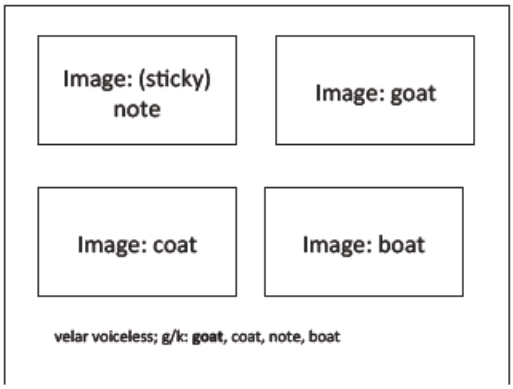
Test 13: Minimal pairs perception test

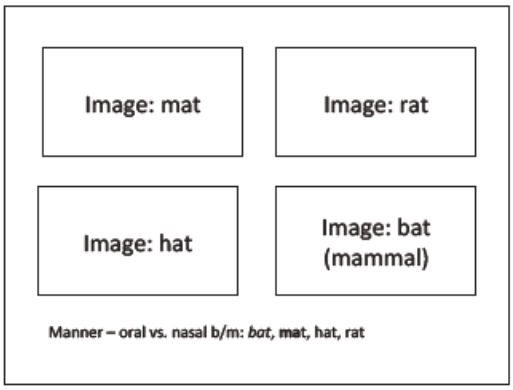
1-5: consonants

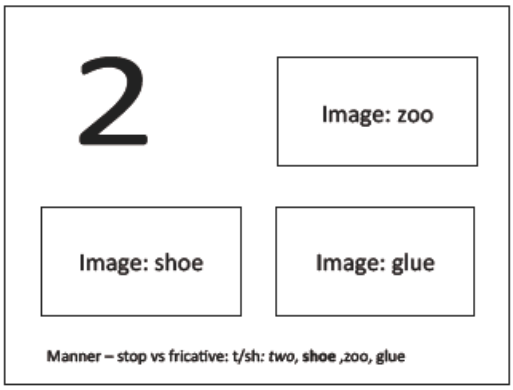






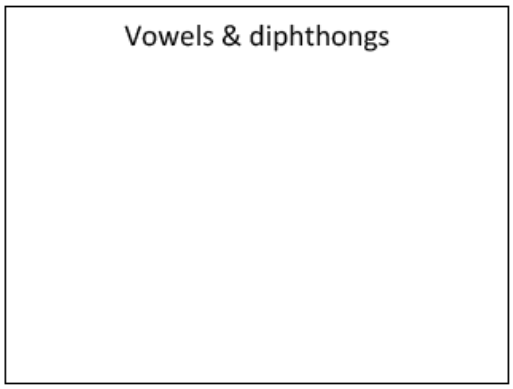


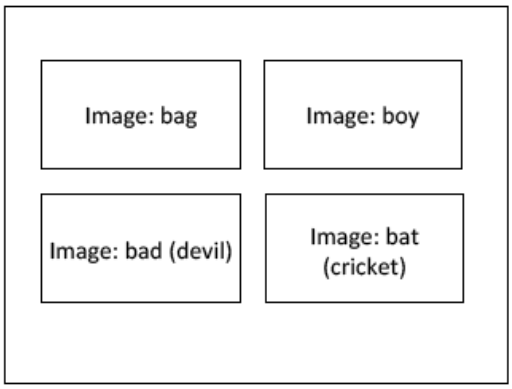


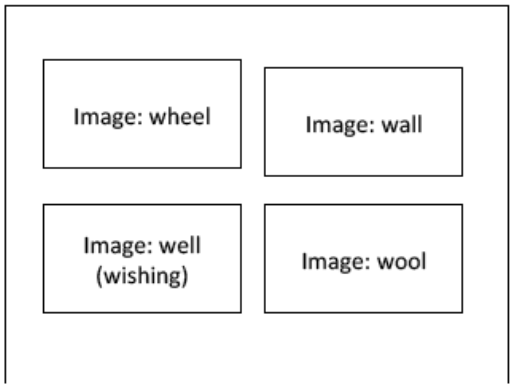


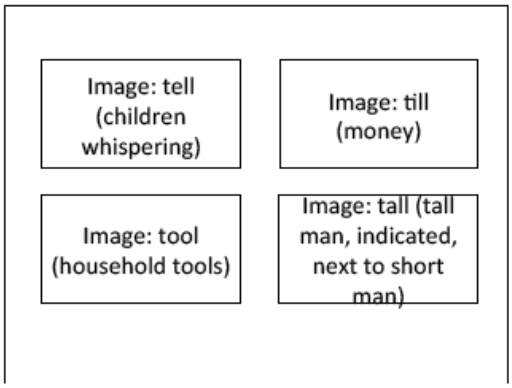
Test 13: Minimal Pairs Perception Test

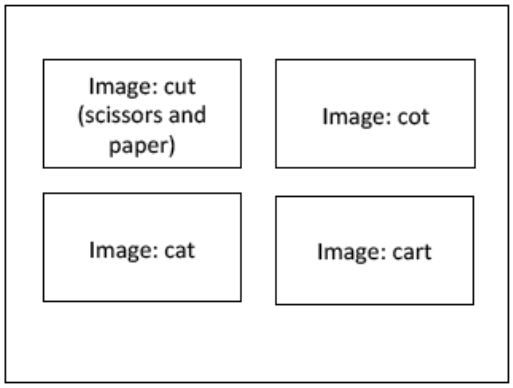
6-11: vowels and diphthongs

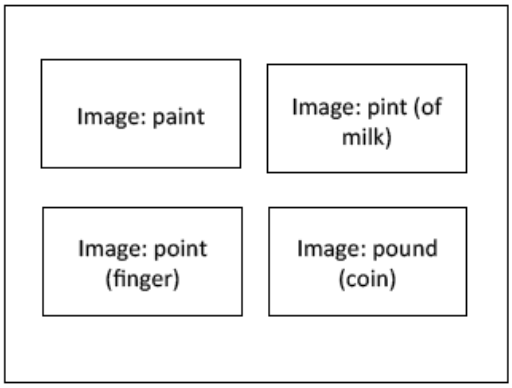






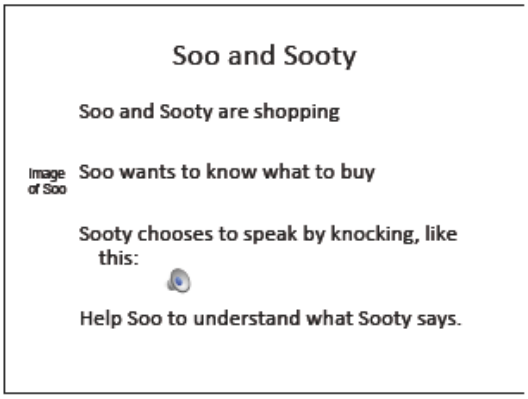


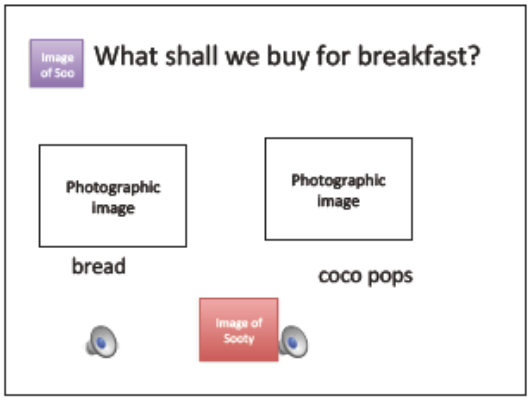


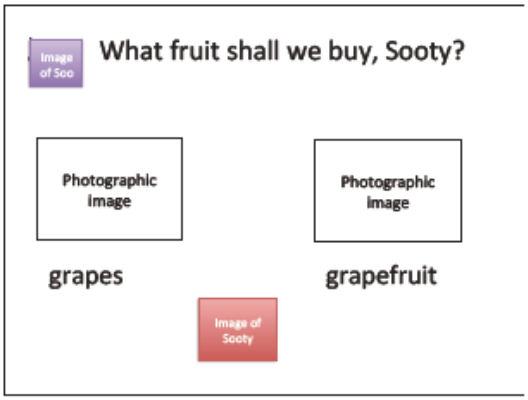


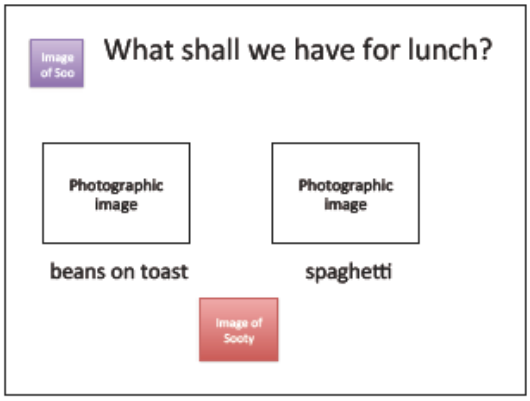
Test 14: Simple prosody perception test

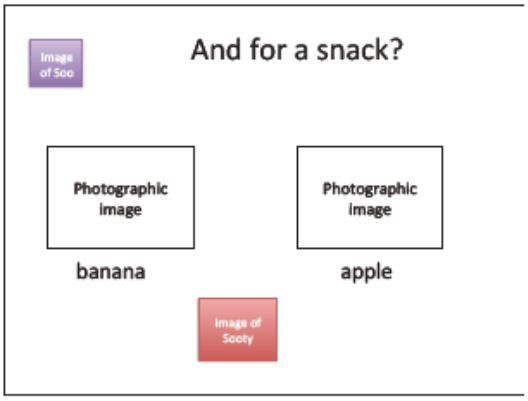












Prosody perception test

Word rhythm and melody

Soo and Sooty

Soo and Sooty are shopping

Image of Soo — Soo wants to know what to buy

Sooty chooses to speak by knocking, like this:

Help Soo to understand what Sooty says.

What shall we buy for breakfast?

Image of Soo

Photographic image — bread

Photographic image — coco pops

Image of Sooty

What fruit shall we buy, Sooty?

Image of Soo

Photographic image — grapes

Photographic image — grapefruit

Image of Sooty

What shall we have for lunch?

Image of Soo

Photographic image — beans on toast

Photographic image — spaghetti

Image of Sooty

And for a snack?

Image of Soo

Photographic image — banana

Photographic image — apple

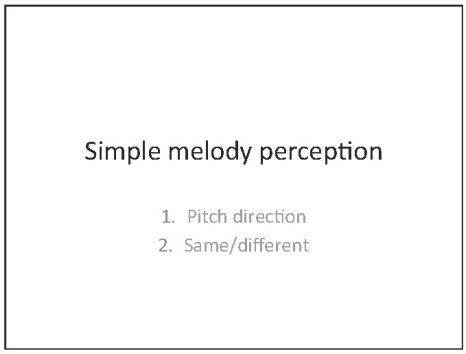
Image of Sooty

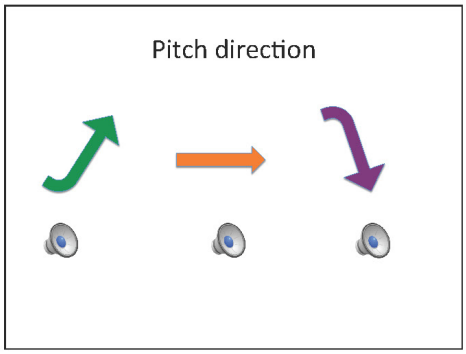
Test 26. Pitch and interval perception test

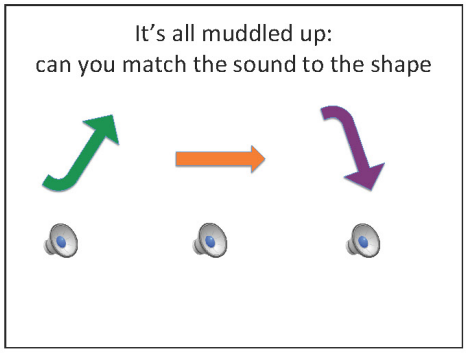
Pitch direction:

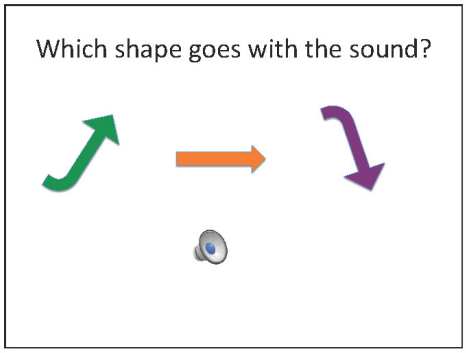
Slides 2-3 are training slides: participants listen to ascending, straight and descending tones and identify their representative line/colour.

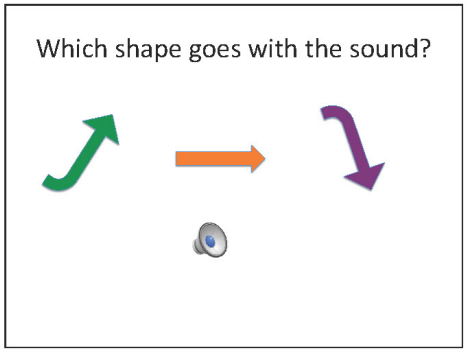
Slides 4-6 require participants to listen to a sound and identify the corresponding line/colour.

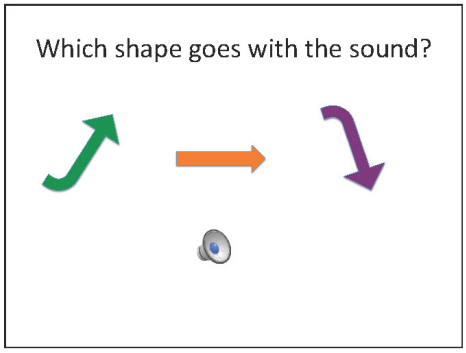








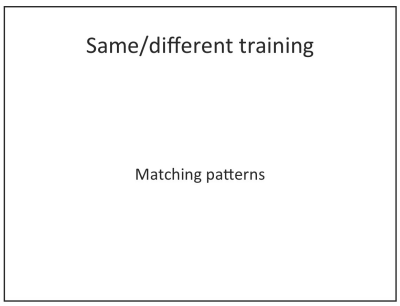




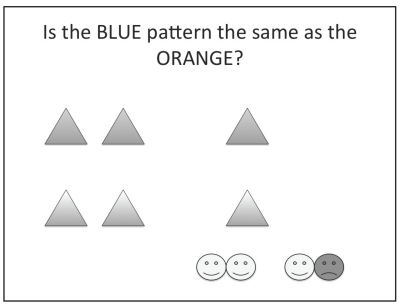
Test 27: Rhythm Perception test

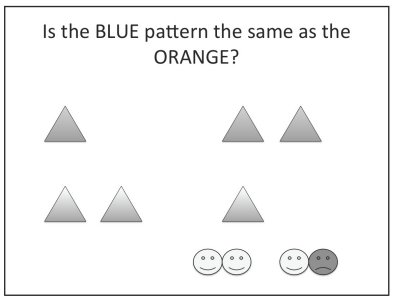
Training

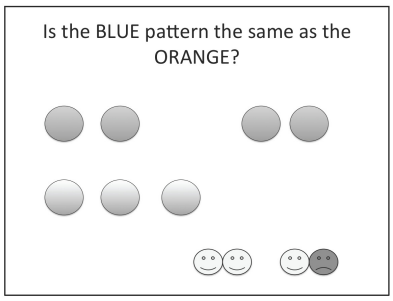
Slides 7-10 are training slides for choosing same/different images, based on visual patterns. The top line of images are coloured blue, and the images on the lower line are coloured orange. Instructions were read to the participants, and they were asked to choose the corresponding emoticons, positioned in the lower right corner of each slide, representing ‘same’ (two yello smiley faces) and ‘different’ (a yellow smiley face, and a blue sad face).



This can be with real objects eg. coloured counters, beads, blocks – the idea is to focus on the pattern being same or different, not the colour

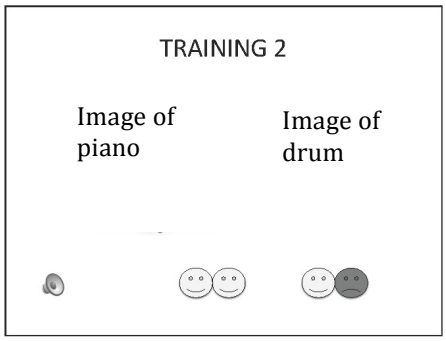


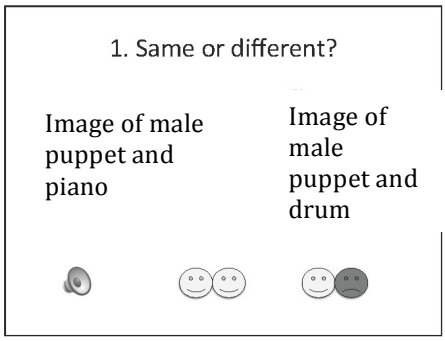


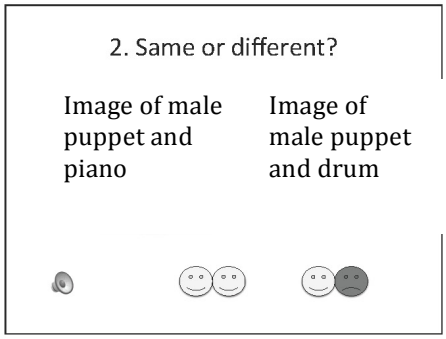


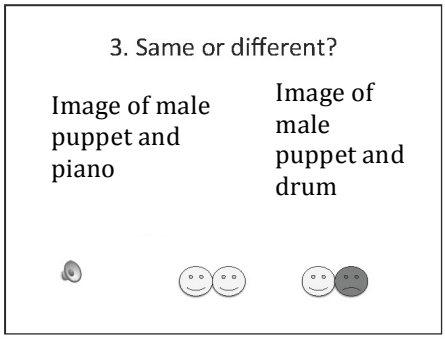
Test 27: Rhythm Perception test

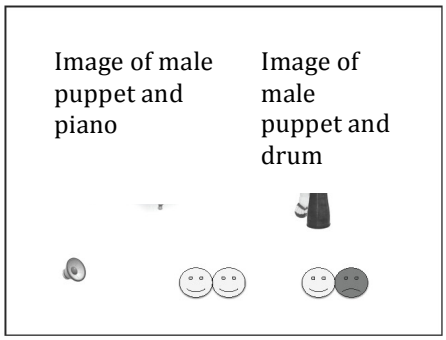
Slides 5-7: Same/different identification of rhythm patterns

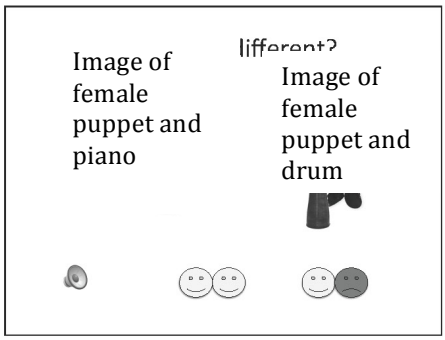






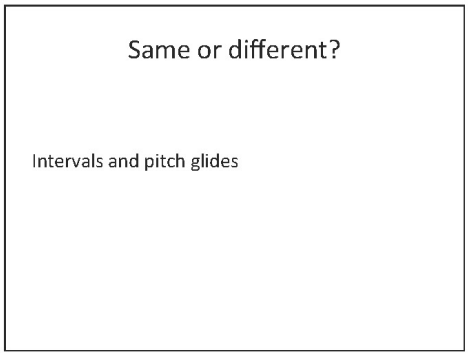


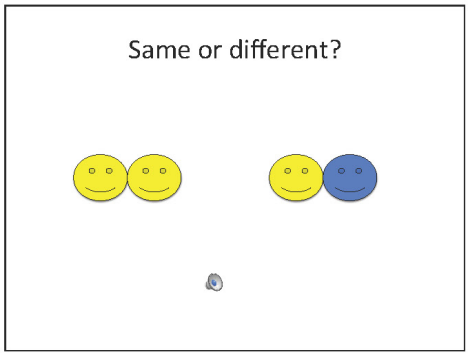


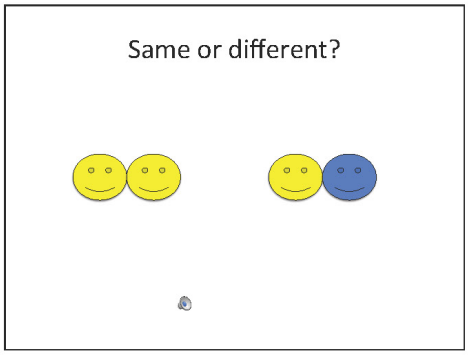


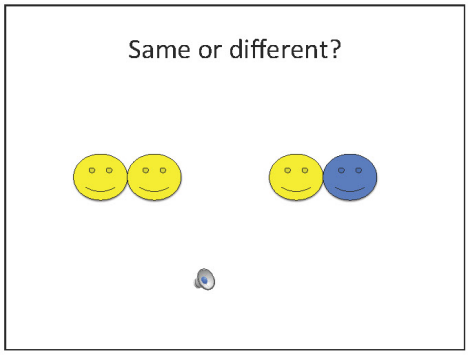
Test 27. Same/Different pitch perception tests

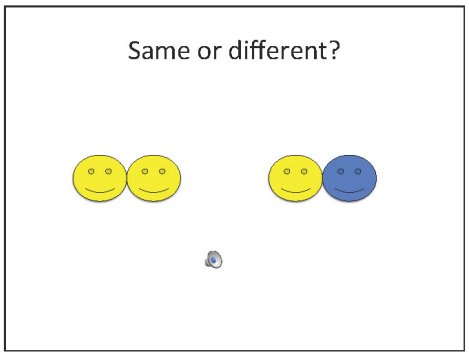
Pitch and interval perception test: the sounds used in order are: Descending 5ths (s); Octave (d); ascending minor 3rds (s); Pitch bend down (s); descending and ascending minor 3rd (d); unison (s).

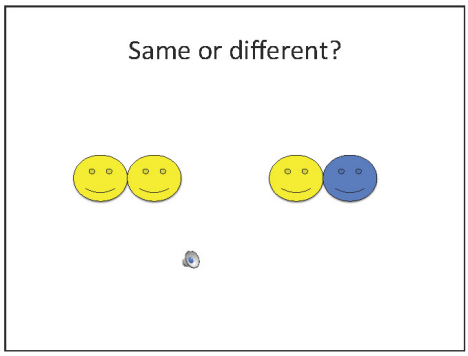




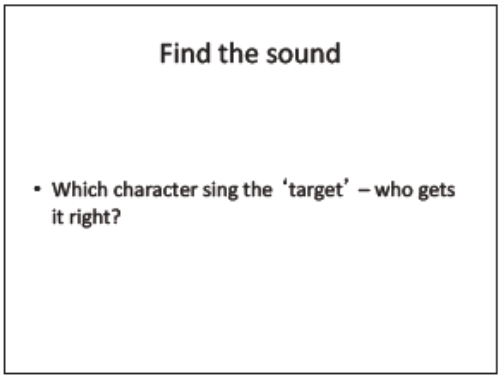


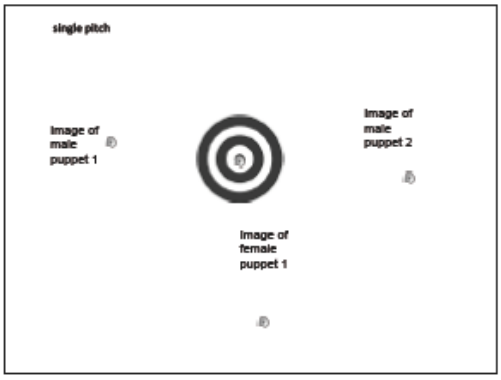


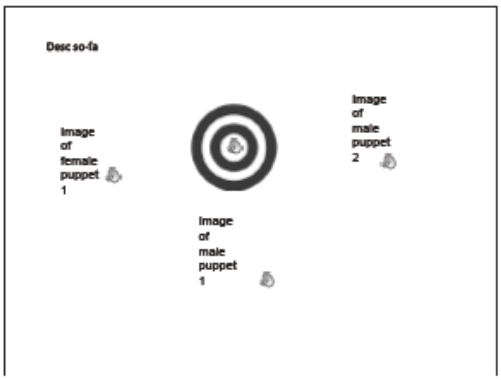


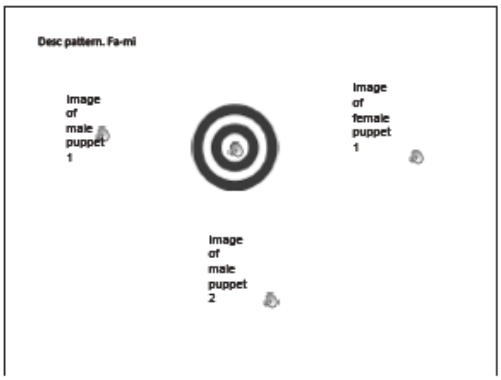


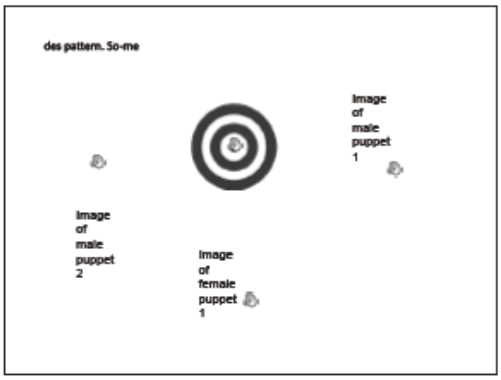
Test 28: 'Find the sound' interval and rhythm perception test.

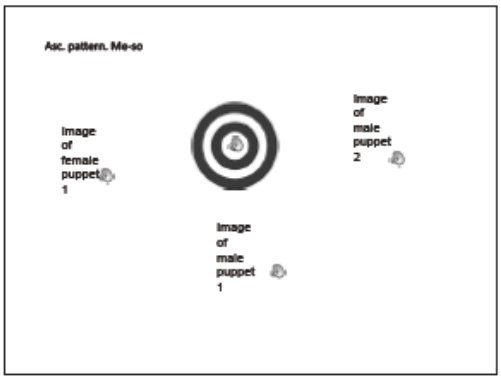


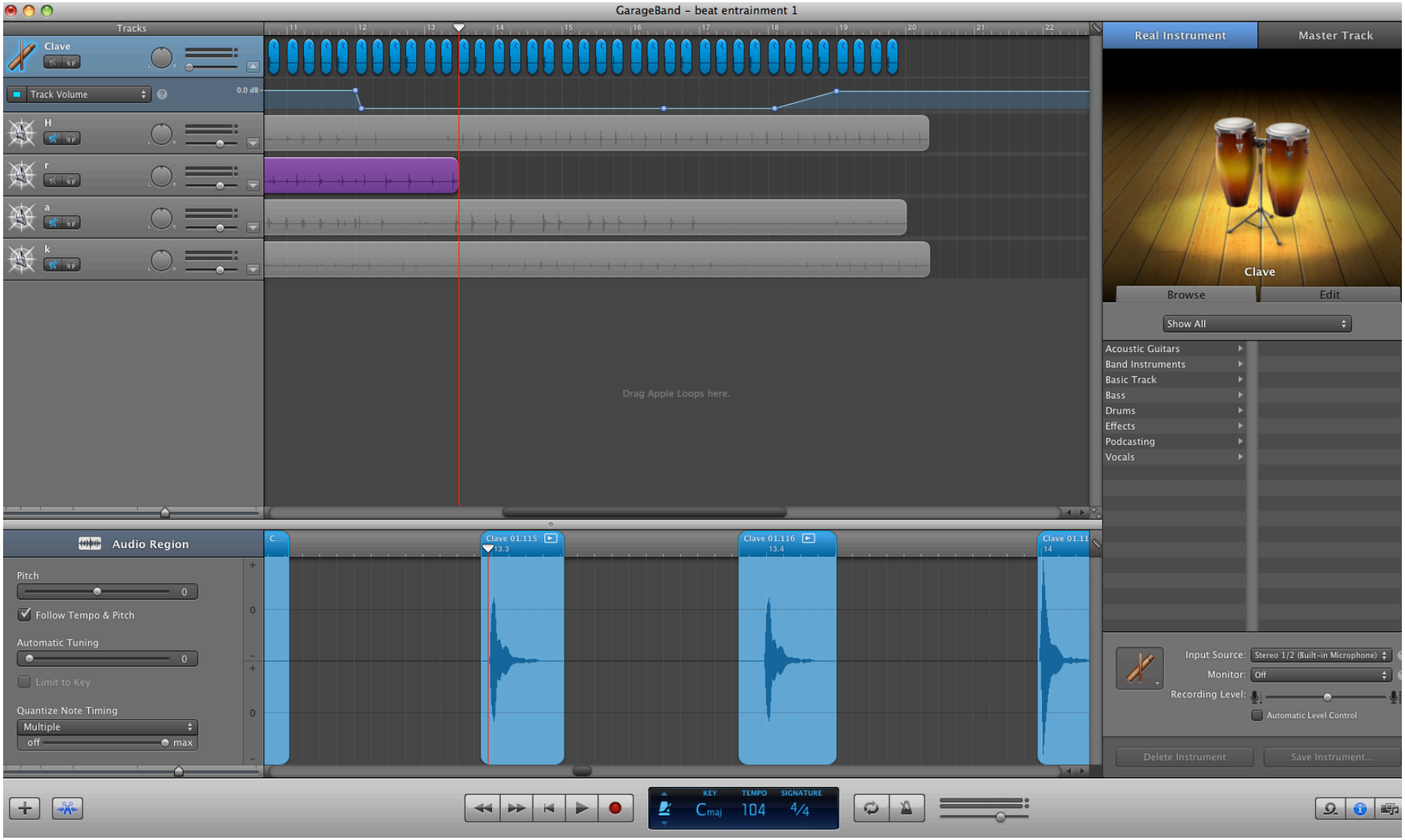






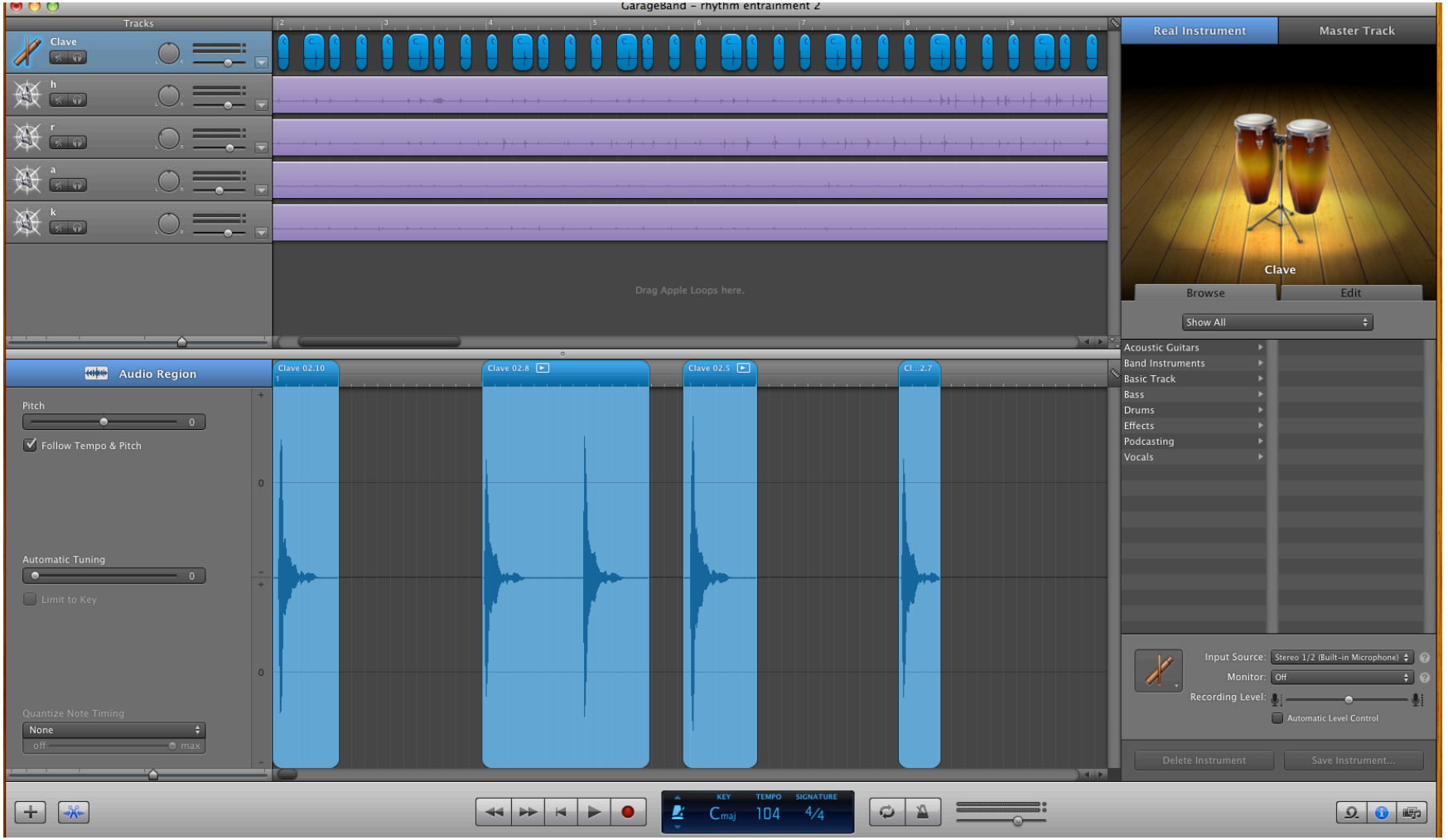






Beat entrainment task 1

Screenshot of the recording for the beat entrainment task at 104 bpm. The ‘clave’ track and detail shows the stimulus. The tracks below show the recordings made of participants.



Rhythm entrainment task 2.
Screenshot of the recording for the rhythm entrainment task, *ta tee-tee ta ta*, at 104 bpm. The 'clave' track and detail shows the stimulus. The tracks below this show the recordings made of participants.

Appendix 3

Teaching sessions, evaluation and sample teaching materials

Six-week teaching plan 340
Learning aims for all students..... 340
Teaching repertoire and expected progression..... 340
Evaluation of the plan, repertoire, expected aims and progression..... 341
Session plans and notes..... 342
Session 1: 30th October 2011 342
Plan..... 342
Notes/Evaluation..... 342
Session 2: 14th Nov 2011 344
Plan..... 345
Notes/Evaluation..... 346
Session 3: 21st November 2011..... 347
Plan..... 347
Notes/Evaluation..... 347
Session 4: 28th November 2011..... 352
Plan..... 352
Notes/Evaluation..... 353
Session 5: 5th December 2011..... 355
Repertoire: sample teaching songs..... 356
Hello song..... 356
Rocky Mountain..... 357
Magic Book..... 358

Six-week teaching plan

Teaching aims: The sessions will aim to teach the link between the aural experience and visual representation of pitch direction and melodic shape. Students will be encouraged to use inner hearing to sing melodic patterns from visual shapes. Throughout sessions they will also practise clapping word rhythms in songs, and in the final session they will use inner hearing to recall a phrase and clap its rhythm.

Learning aims: for all students

- Extend singing range, as appropriate
- Develop accuracy in pitching single notes and intervals
- Develop resonance/phonation
- Increase breath awareness and management
- Practise inner audiation for melody and rhythm
- Practise syllable segmentation, rhythmic singing and ensemble singing
- Show the pulse and rhythm of songs
- Practise volume control

Teaching repertoire and expected progression

Rocky Mountain: use the song to develop attention to word rhythm and ability to clap and say/sing the words rhythmically. Practise: pulse week 1; clap words, week 2; read rhythm, week 3 and week 4; play rhythm week 5; sing week 6

Hello, How Are You?: use the song to develop so-me tuning and awareness of the interval using hand gesture: monitor pitching, show melody,

Do Re Mi: practise imitating and sustaining single notes and upward scale; learn a phrase, with actions to help

We Will Rock You: draw melodic line; learn and articulate words to selected phrases: 'buddy ...', articulating /b/; 'we will, we will...', articulating /w/ on the beat

Monkey Man: develop accuracy in word rhythm and melodic shape of chorus refrain 'aye aye aye'

Magic Book: learn words to the song, develop accuracy in pitching of *so-me la;* practise keeping pulse, tapping rhythm

Individual performance: progression expected in vocal; range; articulation of practised sounds; pitching of individual notes in **Do Re Mi** and of the intervals *so-me*, and *so-me-la*

Audiation: perform the **Rocky Mountain** rhythm whilst thinking the song; perform the rhythm to **Monkey Man**; sign **Monkey Man** whilst thinking words

Evaluation of the overall plan, repertoire and expected aims and progression.

The 6-week plan was reduced to just four weeks of teaching, due to unforeseen circumstances. This has consequences for the development students could have made in terms of musical skills – especially those related to rhythm and syllable segmentation which were to be taught in the final two weeks.

The teaching aims were achieved in most sessions but Session 3 was disrupted due to technical failure (the computer programme cut out before the teaching was complete) and difficulties in managing the behaviour of Andrew; Session 4 was also disrupted due to behavioural problems, and some planned activities were substituted with improvised activities to encourage participation. Session 5 was cancelled by the setting after participants had completed the warm-up to allow them to take part in an extra-curricular activity. The teaching for Session 6 was also cancelled by the setting, but re-assessments of pitch and the assessment of song-learning was allowed to take place in lieu of this session.

The students all made progress in their learning of songs: all learned elements of *Magic Book* and *Monkey Man*. The students made progress in using the kazoos for phonation and in sustaining phonation. All made progress in linking the visual form of notation with aural.

The teaching repertoire was appropriate to the students' interests and abilities with the exception of the bean-bag game, which the students found too difficult. The use of the magic trick worked well to encourage solo performance and peer. It helped draw attention to which elements needed further practice.

The session length and format was appropriate; flexibility in the plan was necessary, though, to allow for disruptions, behavioural issues and emotional issues.

Note: Comments on the achievement of individual goals are highlighted in the following notes with underscored text. Please be aware that these field notes were informal records of teaching sessions, based on notes taken at the time, and are not full records of teaching or learning.

Line numbers are given for ease of reference in the main body of the thesis.

Session plans and notes

Session 1: 30th October 2011

Aim: to introduce teaching repertoire and the session format

Plan

Warmups:

- *Hello, how are you?* – demonstrate the song and show melodic shape; request 'answer' of 'very well, thank you'
- Physical: move in time to ABBA
- Breathing and voice: blow pinwheels, then use kazoos to 'sing' to *Bad Romance* by Lady Gaga; practise yawn-sigh
- Pitching with visual support: sing/use kazoos to sing lines: up, down, straight and siren.
- Pitching to chime bars: do, re, mi – each play a chime bar at one note, then all sing the pitch
- Pitching to song/articulation/rhythm: *Doo Wah Diddy*
- Beat/pacing: drum and sing *Doo Wah Diddy*

New songs:

- *Rocky Mountain*: pass the ball to the beat; one person to drum to beat; listen/sing along in preparation for learning next week
- *Monkey Man*: demonstrate whole song then say the words and clap the rhythm; demonstrate again whilst one participant at a time plays beat; demonstrate chorus of 'aye aye aye' then encourage call and response of chorus; whole group to sing when cued; demonstrate verse for teaching next week

BREAK

Solos:

- Perform choice and listen to feedback.
- *Magic Book:* demonstrate song and game.

Notes/Evaluation

Warmups:

Kazoos: Kerry struggled but managed to make a sound after 5 minutes of trying.

Blowing:

- Andrew maintained 3 seconds, with good lip shape and breath control; tense shoulders on inhalation.

- Rachel: good lip shape, blew continuously for 4 seconds
- Kerry: found it difficult; maintained outbreath for one second
- Robert: blew for 4 seconds

Visuals during warmups worked well: Andrew and Rachel each demonstrated how to sing a pattern from a visual.

***Rocky Mountain*:** there was no time to do this.

***Monkey Man*:** Andrew and Rachel, then Kerry and Robert accompanied the song in pairs, on drums.

Solos:

Kerry: We Will Rock You – good word rhythm and volume throughout.

Andrew: *Please, Please* by McFly. Most of the song was performed in a loud voice, close to shouting; there was rhythmic, but not melodic, variation and words were unclear and produced in monosyllables; voice breaks were apparent towards the end.

Rachel: A Whole New Word: good melodic shape; words unclear

Robert: *Spirit in the Sky:* sung throughout in a quiet, high-pitched voice; some melodic variation but all sung on vowels with no clear attempt at words.

Magic Book:

Andrew: recalled most of the words and melodic shape; showed good recall and improved on all aspects after singing one line at a time.

Rachel: forgot the phrase 'what's inside the' consistently, even after practice; reasonable rhythm and so-me interval

Kerry: good recall of melodic shape and words

Robert: unable to recall words or melody on own, but with assistance by prompting of the initial sound, he sang key words without melodic shape

Session 2: 14th November 2011

Aim: use visual mode to reinforce aural by introducing the use of line drawings to represent pitch direction and interval; associate visual representation of melodic interval with sound; teach the *hello* song

Objectives: by the end of the session, learners will associate an image of a melodic shape with changes in pitch direction; some will associate relative change of visual image with interval

- Develop skills by individual objectives
- Develop skills in singing up and down, using visual feedback
- Relate visual to aural representation of *so-me* and *so-me-la* interval in *hello* song: sing interval with visual rep and kinaesthetic rep
- Phonation: practise sustained phonation using kazoos- pitch glides and sustained notes
- Pitch: sustain vowels on range of pitches
- Learn hand gestures for *Do Re Mi*
- Melodic shape: show melodic shape of chorus of *We Will Rock You* and *Rocky Mountain*
- Rhythm/ pulse : keep a steady pulse when playing; clap words of *Rocky Mountain* with clear changes in duration

Individual targets:

Kerry: develop accuracy in pitching of familiar melodic intervals; produce melodic and rhythmic elements of *We Will Rock You*; articulate key consonant sounds in *We Will Rock You*; sustain phonation for 5 seconds with constant quality

- Practice phonation by using a kazoo; obtain a continuous sound for 3 seconds at any pitch – achieved on siren, up and down.
- Imitate *so-me* model in *hello* and *magic book* – very accurate pitching after listening to a demo, with visual support
- Articulate /b/ sounds in verse to *We Will Rock You* – said words to the verse

Robert: develop vocal range, explore lower voice; sing some words in *We Will Rock You* with vowel contrast and some phonetic detail; begin to show melodic contour in chorus of *We Will Rock You*

- Sustain phonation for 3 sec using kazoo at any pitch – achieved on siren, up and down
- Explore lower range of voice using kazoo with visual assistance -not yet
- Articulate /w/ and vowels in *We Will Rock You* – achieved, with visual support

Rachel: develop accuracy in performing melodic intervals s-m and s-m-l with visual and gestural assistance; sing songs with increased clarity in contrasting vowels and consonants in *A Whole New World*; develop inner audition for melodic pattern and pitch; inner audition; listen and think a single pitch before producing

- Imitate melodic interval in *Hello* and *Magic Book* and use *so-me* gestures - achieved
- Practice consonant articulation at different tempo – not yet
- Sustain phonation for 3 seconds at constant, given pitch - achieved

Andrew: produce sustained phonation with no change in quality for 3 seconds at 3 different pitches; develop accuracy in pitching melodic patterns; sing the chorus of *We Will Rock You* with appropriate melodic contour; develop inner audition for pitch and pattern

- Listen and think before copying single pitches – some improvement; still rushed in
- Sustain phonation on kazoo at 3 pitches for 3 seconds – achieved
- Imitate so-me patterns with visual and gestural references – achieved during *hello song*
- Practice articulation of chorus of *We Will Rock You* at different tempos – achieved with visuals and gestures to support
- Show melodic contour of *We Will Rock You* – achieved: drew in the air using Makaton signs for initial consonant

Plan

Warm ups:

- Movement to music by ABBA
- Breathing: blow pinwheels to Lady GaGa's *Bad Romance* - changed to *It's Raining Men*
- Voice/kazoos: initiate phonation; perform sirens from visual info; sustain phonation for 3 seconds at 3 pitches - *all used kazoos on YMCA chorus*
- *Hello, how are you?:* listen, copy, follow melodic shape individually RECORD on ZOOM to chart learning and pitching
- Pitching with chime bars- pentatonic: Listen and sing each note, and copy *do re mi* Kodaly hand signs; sing *Do Re Mi*
- Articulation/pulse/tempo: *We Will Rock You*. Draw the melodic shape of phrase *we will rock you* - on *la*- in the air with hand, then add words. Sing verse 1 and chorus to drum beat. *Demonstrated, and used the Makaton signs for 'w', 'r' and 'y' to cue words whilst drawing melodic shape*

New songs

- *Rocky Mountain* (for later inner audition test): play game and keep pulse
- *Monkey Man:* recall words/ signs and rhythm; practice parts a line at a time- listen and repeat; watch signs and copy; put it together

BREAK 15 mins, 2:15-2:30

- **Solo songs** and feedback RECORD Students giving feedback SPONTANEOUS SPEECH (Zoom)
- *Magic Book* - demonstrated song and actions. Individuals sing one at a time RECORD to chart development of song learning

Notes:

- *Andrew*: phonation sustained on kazoo, but breaks apparent when singing/speaking
- *Kerry*: good VQ when singing at higher pitch; excellent melodic shape
- *Rachel*: sings at pitch higher than speaking; good VQ
- *Robert*: speaks and sings in high voice, sustained vowels are notably quavery

Session 3: 21st November 2011

Aims: use visual image to develop accuracy in singing a single note and intervals; use voice to produce a visual line drawing; reinforce link between auditory and visual concepts;

Practice syllable segmentation; perform songs with group pulse and rhythm at 3 tempos

Plan:

Warm-ups/melody:

- Use iPad to practise accuracy of interval
- Repeat words to *Rocky Mountain* and clap word rhythm to pulse
- Draw melody of chorus /phrase of *We Will Rock You*
- Articulate and clap words to verse 1 of we will rock you at different tempos
- Perform *Monkey Man* with actions; learn verse
- Pinwheels, kazoos, up down, sirens, lines - repeat on vowels
- *Hello, how are you?:* sing and draw in air
- *Do re mi*: sing each pitch and draw in air
- *We Will Rock You:* draw in air
- Use iPad/ sing and see to practise singing straight lines, up, down, and interval

Songs/Rhythm:

- *Rocky Mountain:* play game with beanbags
  - Listen and repeat a line
  - Listen and clap a line
  - Sing and clap to beat

- *We Will Rock You*: repeat above stages
- *Monkey Man*- recall words/ actions and sing along to guitar
- Learn verse 1
- Solo
- *Magic Book*

Notes on the session recording

Breathing and kazoos:

Warmup: dance to Glee's *Don't Stop Believing*

Breathing:

With pinwheels

- *Rachel*: exhaled for 3 counts – vigorous but sustained

- *Andrew*: exhaled for 2 counts; achieved 3 counts after instruction to relax shoulders
- *Robert*: exhaled for 5 counts
- *Kerry*: exhaled for 3 counts

Kazoos:

- *Andrew*: n/k
- *Rachel*: sustained for 4 counts
- *Robert*:-sustained a long straight line and sirens with kazoo circa 5/6 seconds
- *Kerry*: n/k

'Ah' sirens and down:

- *Robert*: produced 'oo' in upward and downward shape in response to instruction; and for the downward shape; on the visual for 'down'- he went up then down
- *Rachel*: imitated up and down in high voice; went up in response to visual
- *Andrew*: produced straight tone in response to visual instruction for direction- produced low tone in response to visual for downward direction: some movement but very limited pitch range;
- *Kerry*: short up and down

(5 mins after physical warmup)

***Do Re Mi* with chime bars:**

Pitch matching – Do, re, mi

- *Rachel*: matched pitch well; all others were arguing and the performance was disrupted by Andrew
- *Do*: All practiced matching *Do* except Andrew who sang *re*, then *ti*
- *Me*: Rachel played the note; and matched well, Kerry matched, Robert attempted a match – not quite on pitch,
- *So*: Kerry played the note – Robert matched, Rachel sang at a lower pitch, Kerry sang close to target but lower; Andrew, sang lower

Interval matching using *Hello how are you?*

Question to group: 'can you recall the song?' No

However, all sang the song with gestural prompts for words and interval; Andrew was able to produce the response independently

Solos: responses to the sung question *hello, how are you?*:

- *Andrew* responded with: 'I'm very well *thank* you' – he spoke with some rhythmic and melodic expression
- *Robert* produced the words 'very' and 'thank you' with visual prompts to recall words; no melodic variation, but rhythmic approximation
- *Rachel* was unable to recall the words independently, but she sang the words with phonological prompts
- *Kerry* produced the words with appropriate rhythm and melodic shape at a pitch that was close to the demonstration, without visual or verbal cues

Pop songs and melodic shape: singing phrases with visual support

Don't Stop Me Now (using the first phrase in the song's chorus):

Demonstrated the pattern with a visual line-drawing and gesture. The group were asked to draw the pattern in the air as I demonstrated; then to draw the shape as they sang it

- All sang twice observing visual representation of sounds and melodic shape
- *Kerry* sang clearly with melodic shape, but with some hesitation on 'now'
- *Andrew* sang all words and produced the appropriate downward gesture with his hand
- *Robert* sang the approximate melodic shape when supported with a visual prompt

Sing and See demo. The activity aimed to use a computer to demonstrate a real-time image of singing pitch ('Sing and See' programme) in order to reinforce the idea of visual and gestural images of pitch. It was explained to the group as being able to 'paint your voice'. However, the computer failed before the group was able to fully explore its use. *Kerry* sang a straight line and siren into Sing and See; *Rachel* sang a siren into Sing and See; *Andrew* sang a line with a slight siren, and *Robert* produced upward movement.

The same principle was then applied to the songs that each participant had chosen to sing as solos, later in the session. A section of the chorus was chosen and the researcher demonstrated the melody with gesture.

***Uptown Girl*.** In response to a demonstration of the melodic shape and words of the first line to the chorus:

- *Kerry* imitated the section accurately; *Rachel* imitated the words but sang on one note; *Andrew* recalled the words and imitated part of the melodic shape but used two voice qualities (falsetto style/gruff style)

A Whole New World. In response to a demonstration of the melodic shape and words of the first line to the chorus:

- *Rachel* sang words and melodic shape; then she produced the words with greater attention to clarity, but less musical accuracy
- *Andrew* produced some appropriate distinction in pitch: he sang 'new' on a higher pitch and 'world' on a lower pitch.
- *Robert* sang an upward melodic glide, without distinguishable words
- *Kerry* struggled to get words out initially. However she sang the end of melody and continued into the next line

I'm a Believer. In response to a demonstration of the melodic shape and words from the chorus*:*

- *Robert* sang the word 'face'. After a second demonstration of the words, and shape, then tune, and with Makaton signs to prompt words, he produced 'face' and 'believer'

We Will Rock You. I recapped what we had learned about drawing key sounds and melodic shape in air; then focussed on articulation. I drew the melodic shape of the chorus in the air using a /w/ to cue the sounds. The group sang to this cue; then we applied it to the chorus

- *Robert* produced some differentiation between vowels /i/ and /u/
- *Rachel* placed some emphasis on known words
- *Kerry* produced clear sounds at the start of words in the chorus
- *Robert* produced initial consonant sounds and a clear 'you'

***Rocky Mountain* game:**

I demonstrated the tune as the group tapped the beat on their knees –

- *Kerry* sang and *Andrew* sang along;

I demonstrated the tune again and tapped the beat; then demonstrated again and modelled passing the beanbags around the circle in time to the beat. We played the game at different speeds, passing the bean bags around the circle as I sang the words.

- *Andrew* and *Robert* joined in periodically at the start (Andrew retained the ball and refused to pass it to Robert)
- *Andrew* sang the song, with prompts and accompanied all others, singing key words and sounds
- *Kerry* sang with prompts
- *Rachel* sang with prompts, and produced the melodic shape
- *Robert* watched and listened

Monkey Man:

- *Andrew* recalled the words ‘aye aye’ from a gestural prompt; he sang solo when given gestures to prompt recall
- *Kerry* recalled the key words
- All joined in Chorus

Session 4: 28th November 2011

Aims: sing and show pulse: combine an action to the pulse with word rhythms to the songs *Rocky Mountain* and *Magic Book*

Teaching objectives:

- Singing: use visual aid to fine tune pitching and explore range,
- Musicianship: use inner audition to clap word rhythm; relate word rhythm to *tee tee* and *ta* syllables;
- Voice: practise articulation of words in given pop songs and apply when singing the melody
- Produce resonant vowel across 3 notes
- Perform *Monkey Man* with Makaton
- Perform *Hello* and *Magic Book* with Makaton, and from memory

Individual objectives:

- *Andrew*: achieve sustained phonation without using diaphragm – keep stomach relaxed
- *Rachel*: articulate sounds of words in *Monkey Man*, *hello* song, *Magic Book*
- *Kerry*: articulate words in familiar songs; increase accuracy in intervals and pitching
- *Robert*: explore lower range; articulate key words; demonstrate rhythm in songs

Plan

Warmups

- Movement: to McFly
- Breathing: practise deep inhalation, and 'ah' downward sigh
- Phonation: produce 'yummmm', based on favourite foods
- Range: sirens, from notation – on 'oo', and 'humm'
- Pitching single notes:
- Pitching intervals: *Hello, how are you?* – use visual notation; all sing answer; sing greeting to each other (use notation)
- *Do Re Mi* – pitch single notes; sing a line at a time SOLO; sing whole song together
- Articulation, ECHOING words: *We Will Rock You* at different tempos to the drum; pass the phrase around the circle and clap PULSE (copy me; go around the circle and each copies it in time to increasing beat); each lead a phrase

SONGS

- *We Will Rock You* : show or copy melodic shape
- Sing together: *We Will Rock You*

- Words and rhythm: play *Rocky Mountain*: tap rhythm on body (melody); clap and say words
- Memory and musicality: *Monkey Man* – sing with visual prompts; all perform actions; one perform at a time to guitar
- *Magic Book*: practise memory and interval pitching: perform as group and solo.

Notes/Evaluation:

No new teaching: *Andrew* was undo-operative and quite hostile to *Robert* (again), and as he is the loudest, this affected all. The plan was mostly implemented, except *We Will Rock You* etc. didn't work: *Andrew* and *Rachel* refused to sing, so I asked them to play accompaniment, and then they sang. I modified this: all played accompaniment using boomwhackers; then they played the chords to the pulse, as one sang. *Rocky Mountain* also didn't work: the beanbag game was too hard, as it was too difficult for *Robert* and *Kerry* to pass beanbags and *Andrew* kept the bags to himself. We sang, played once, then moved on. *Andrew* played the drum and improvised a rhythm on it and all improvised a song. Then each sang *Monkey Man*, supported with actions to cue them. Then as a group, all sang the verse from memory, with actions. We finished with *Magic book.*

BREAK

Solos

Recording notes:

Warmups:

- Breathe in and then produce descending 'ah' glide – Robert imitated well; then Kerry echoed
- Robert sang sirens – all except Andrew (who refused) sang from visual scores of lines
- 'Yumm' siren sounds were produced to suggestions of food

Hello, how are you? Individuals answered to my 'call'; then individuals sang the call to another

- *Robert* sang the question; *Rachel* sang the response
- *Kerry* sang the question to *Andrew*, who responded

Do Re Mi with boomwhackers

- *Do*: the low note was imitated and sustained by all; variable production
- *Re*: *Robert* played the note then sang it, rising in pitch; good pitch matching from all others
- *Mi*: *Andrew* sang with a gruff vocal quality; *Robert* sang higher than pitch

- *Fa*: *Andrew* sang with a higher register
- *La*: *Kerry* didn't match pitch

We Will Rock You

All sang in time as I conducted the rhythm with a boomwhacker – excellent timing from all

- *Robert* produced some consonants and approximate vowel with some anticipation of key words

Monkey Man

Andrew remembered the start of the song independently and sang to actions. He recalled key words and some changes in pitch

- *Kerry* remembered key words and melodic intervals
- *Robert* sang some melodic intervals and key words clearly
- *Rachel* sang words after imitation and intervals with appropriate shape
- All sang the verse we learned last week: *Andrew* recalled key words, *Kerry* recalled key words.
- *Andrew* correctly sang the interval change in the opening line.

*Robert***:** Parts of key words from *Greased lightning* were articulated clearly. Other words were produced on a single sound until *oo oo oo* – the words were anticipated, and produced with clear rhythm and melodic descent.

Magic Book: Robert sang after a demo, with appropriate word rhythm and melodic shape, and articulated the word 'abracadabra'.

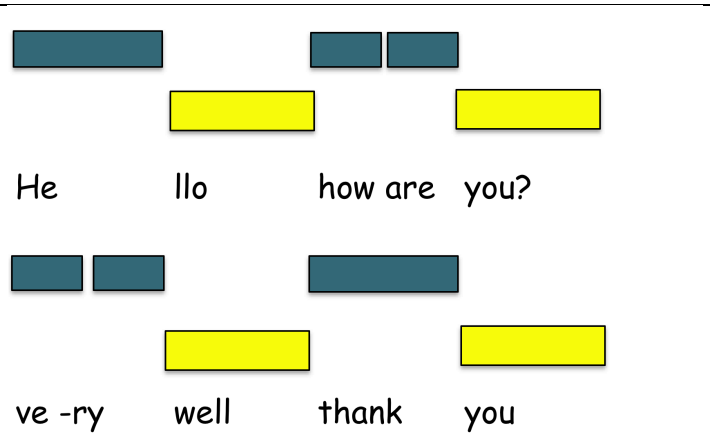
Session 5: 5th December 2011

Aims: sing and clap word rhythm; practise inner audiation of rhythm

No notes available. The session began, but participants were withdrawn in the first 10 minutes in order to take part in a different activity.

Repertoire: sample teaching songs

Visual scores were produced by the researcher. The songs and activities are taken or adapted from Brindle (2005).

Hello song: hello, how are you? The visual represents the melodic and rhythmic pattern. The song is sung with the following interval, from any starting pitch: *So-me, so-so-me* *So-so me, so-me* Adapted from Brindle, 2005	

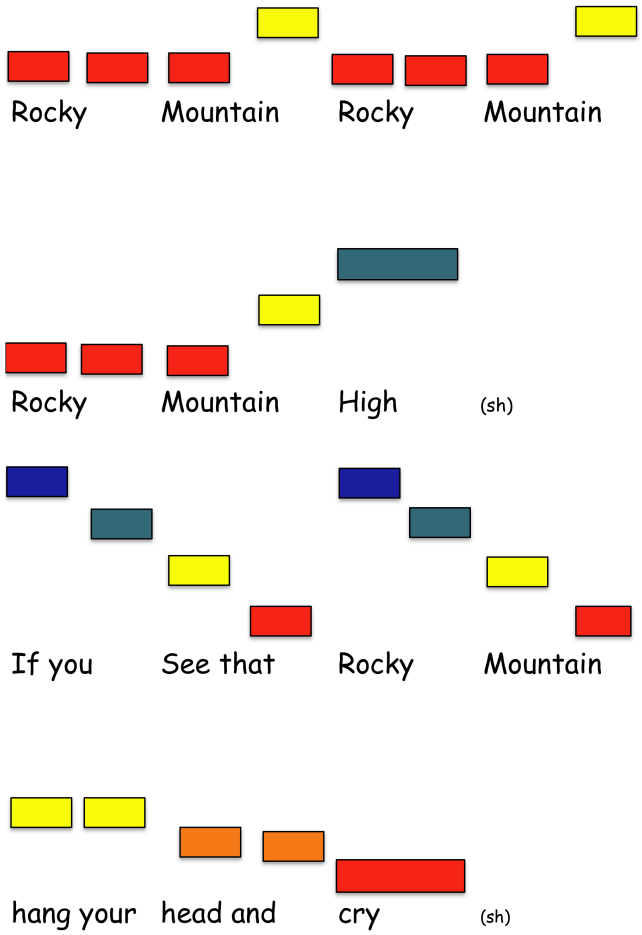
Rocky Mountain

The graphic shows the melodic shape and rhythm.

For rhythmic development, participants were first asked to tap the pulse on their knees, which comes at the start of 'rocky' and 'mountain'; once that is mastered, the game is played by passing beanbags in time to the beat.

The song is sung on:
do-do-do-me
do-do-do me
do-do -do- me- so
la-so-me-do
la-so-me do
me-me-re-re-do

Adapted from Brindle, 2005



Magic Book

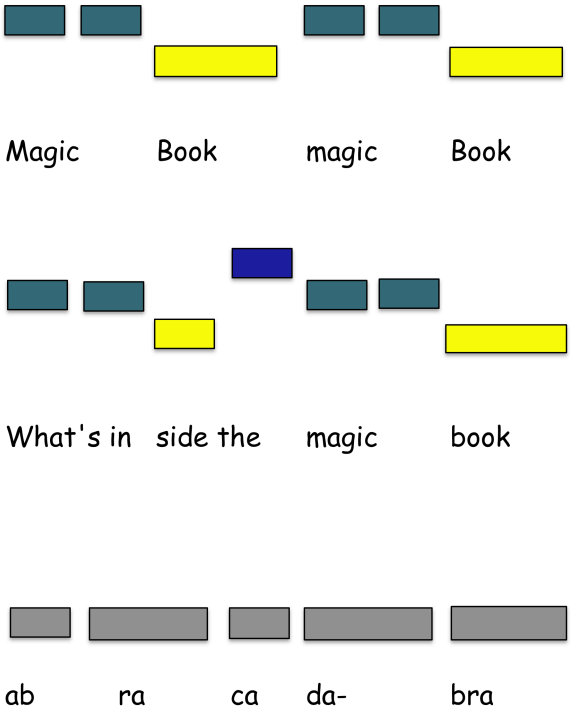
The graphic shows the melodic shape and rhythmic pattern. The final 'abracadabra' is spoken.

Game: with a magic wand, participants tap the beat over the magic book, whilst singing the song. At abracadabra, they speak and tap the book with the rhythm of the words. This is learned by demonstration and imitation, rather than teaching of the concepts.

According to the quality of the singing, the teacher may show the pictures in black and white, which indicates that improvement is needed; or in colour, which shows a good performance in specific aspects.

The melody is:
So-so-me
So-so-me
So-so-me-la
So-so-me

Adapted from Brindle, 2005



Appendix 4
Summary of MSc.

The following pages summarise the dissertation submitted in partial fulfilment of my Msc. in Language and Communication Impairment in Children. An abridged form is given of the aims, methods, key results and findings. Group results are presented, and a case study is given of one participant - a young man with Down Syndrome - for whom there were notable changes in voice quality as perceived by listeners, and as measured acoustically.

An investigation into the effects of singing tuition on aspects of intelligibility in young people with speech impairments and learning difficulties.

(Tracy Jeffery, 2007)

Abstract

Poor speech intelligibility can have significant effects on the lives of people with speech impairments. The intelligibility of speech has been linked to prosody and voice quality, aspects of speech that are also found in singing. This study investigates whether singing would be a suitable medium for effecting changes in the speech of four young adults with learning difficulties, three of whom have Down syndrome, and whether any changes observed would affect overall intelligibility.

Speech samples taken before and after intervention are analysed for changes using perceptual measures and spectral and numerical data generated by Praat. Significant changes ($p<0.01$) are found for two of the subjects in terms of improved HNR, jitter and shimmer. The results are considered in light of data collected from a control group, and the implications of the results for this population are discussed, and suggestions are made for future research.

Overview and Aims

This study aimed to find out whether singing can be used as a means to improve specific segmental and suprasegmental aspects of speech in learning-impaired adults. A group of young adults with speech difficulties and learning difficulties, including individuals with Down Syndrome (DS), took part in seven sessions of a planned eight-week programme of singing tuition, once a week over a period of ten weeks. Samples of connected speech were collected from a test group and from a control group in the week before the singing intervention, and in the week after the intervention. The data were used to learn whether singing tuition - comprising of tuition in posture, breathing, pitching, rhythm, resonance and articulation - affected aspects of speech that are related to intelligibility. To test this, the speech samples were analysed to establish changes in the following parameters:

1. mean pitch, pitch range and standard deviation of pitch;
2. intensity of voice;
3. rhythm in words or phrases;
4. voice quality;
5. articulation of phonemes; and
6. Intelligibility of words and phrases.

Participants

Twelve subjects were selected and invited to participate in the study; all participants were aged between 18 and 22 and attended a residential FE college for students with special educational needs. Of the twelve invited, ten agreed to take part. A test group of five young people with speech difficulties and learning difficulties were offered singing lessons. A second group of five participants, matched for age, gender and level of learning ability acted as controls: these were offered singing sessions after the intervention period. Two of the test group and four of the control group had pre-existing diagnoses of severe learning disability: the remaining participants had been diagnosed with moderate learning disability. All had been identified by the college speech and language therapist (SaLT) as having poor verbal intelligibility and received one hour of speech and language therapy a week. Seven of the students had diagnoses of DS, three of whom took part in the singing intervention and both speech assessments. Of the ten participants, eight were present for both pre- and post-assessment (four in each group): only the data from these are included in this study.

Methods

Speech samples

Samples of connected speech were collected from participants before and after the intervention period. For the test group, speech samples were collected the day before their first singing lesson, and in the week of their last lesson. Speech samples were elicited using a series of ten colour photographic pictures. Each subject was shown the picture which was annotated with a short, descriptive phrase, and was asked to describe the picture or to read the text. The samples were chosen to elicit a range of responses, from single word answers to more complex descriptive phrases. Each participant was then asked the question 'What did you do at the weekend?' in order to record their conversational speech.

Sample collection took place in the Speech and Language Therapy room, a detached outbuilding in the College grounds. Precautions were taken to minimise environmental noise, although complete silence was not possible. Recordings were made using a Sony ECM-MS907 stereo microphone (set at 90 degrees) and a SONY MZ-R70 Minidisc recorder. The recorder was placed approximately 15 cm from the speakers mouth, at an angle of about 45 degrees from the work surface.

Ten students provided speech samples pre-intervention; eight were recorded post-intervention. The samples from the eight subjects were used to determine whether there were any changes in speech, by comparing intelligibility scores, and by analysing samples acoustically.

Intervention

The test group was invited to attend weekly singing lessons, held in the evenings at the college. Each lesson lasted approximately an hour and a half, with a break of fifteen minutes. Eight lessons were planned over a ten-week period, allowing for a two-week Easter break between lessons 4 and 5. Due to illness, the researcher was able to deliver seven sessions only; this also meant that three weeks elapsed between lessons 4 and 5.

The lessons were themed, with a different focus and set of group targets each week. Some activities and were designed with reference to the types of speech difficulties experienced by participants in the group (e.g. exercises to stimulate deep breathing, soft onset of phonation etc). In addition, each

participant was given a set of individual targets to achieve each lesson, which were set depending on their performance in the previous lesson and with one or more of their targets pertaining to the focus of that week's lesson.

Each lesson consisted of a physical warm up; vocal warm-ups and exercises/drills; a group song; a break; additional warm-up/drills; individual singing with verbal feedback; and a final group song. Progress was charted through the setting of individual targets, which were set according to each participant's performance each week. The researcher evaluated each lesson to establish which exercises and methods were successful and which were not: not all the exercises selected, and trialled were suitable for the test group and were modified for subsequent sessions, or dropped.

Acoustic measures

All pre- and post-intervention speech samples were analysed in Praat. Visual spectrograms were used to identify segmental and suprasegmental features, and phonetic transcriptions were created. Spectrograms were also used to identify contours in pitch, intensity and formants. Praat was used to obtain voice perturbation data for the voiced parts of each sample. Manual selection of vowels in words enabled the production of voice reports, which generated data for pitch and voice quality (jitter, shimmer, HNR). Whole phrases were analysed if the samples collected before and after intervention matched closely in terms of the words spoken: where the content of phrases differed between initial and post-intervention samples, identical words were selected for comparison.

Perceptual measures

Perceptual tests were also used to test for changes in intelligibility and voice quality. For each participant, four of the ten elicited picture-based samples were chosen at random by numbering each track and by pulling the number out of a box. These samples, and each sample of spontaneous speech, were compiled as a playlist using the shuffle function on Windows multimedia player to randomise presentation order. The prepared samples were played to a panel of four judges to derive intelligibility scores. The judges were volunteers, all of whom underwent a hearing test administered by a nurse using an audiometer. One of the judges had previous experience of working with the client group; the remainder were naïve listeners. The test was trialled with one judge and it was found that the more intelligible samples that had been derived from pictorial stimuli were influencing her interpretation of other samples. A further test was then conducted using selected samples from those available, but eliminating those samples with the highest intelligibility rating (as determined in the trial), and minimising repetition of samples. Therefore, a further test was conducted with the remaining three judges. The judges listened to the samples in random order, and their responses were analysed as follows.

1. The words recorded by judges were compared to the target words uttered by the subjects.
2. The syllables identified by the judges were compared to the syllables in the target words spoken by the subjects, with one mark allocated per syllable.
3. A mark was awarded if the gist of each phrase was identified by the judges.

The judges were also asked to rate samples of spontaneous speech in terms of voice quality using a modified GRBAS scale: the scale was modified to exclude parameters that the test subjects did not exhibit (eg. aesthenia), and the vocabulary was simplified in order for non-specialists to use the scale (eg. the term *grade* was replaced by *overall quality*).

Results: Group data

Acoustic measures of voice quality

The acoustic data collected from the voice samples of the test group were compared to data collected from the samples of the control group (Figures 1-3). Post-intervention, the mean local jitter across all samples decreased for three of the four subjects in the test group; it increased for three of the four subjects in the control group (Figure 1). The mean local shimmer decreased post-intervention for all the subjects in the test group; it decreased for three of the four subjects in the control group (Figure 2). The mean HNR increased post-intervention for all the subjects in the test group; it increased for

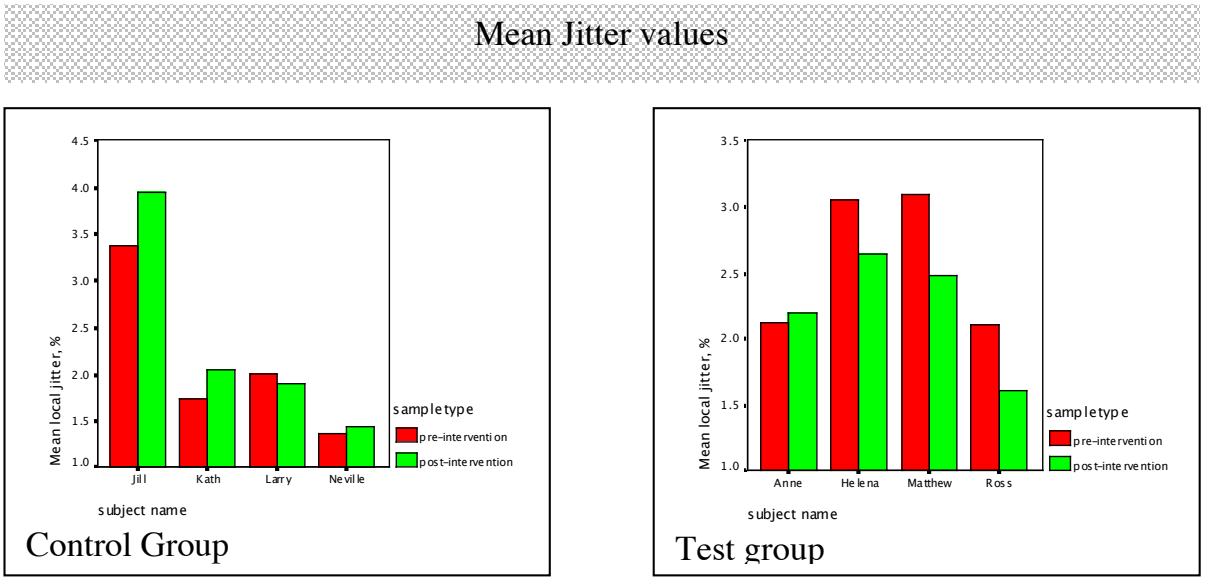


Figure 1: : mean jitter values before and after intervention a) control group b) test group

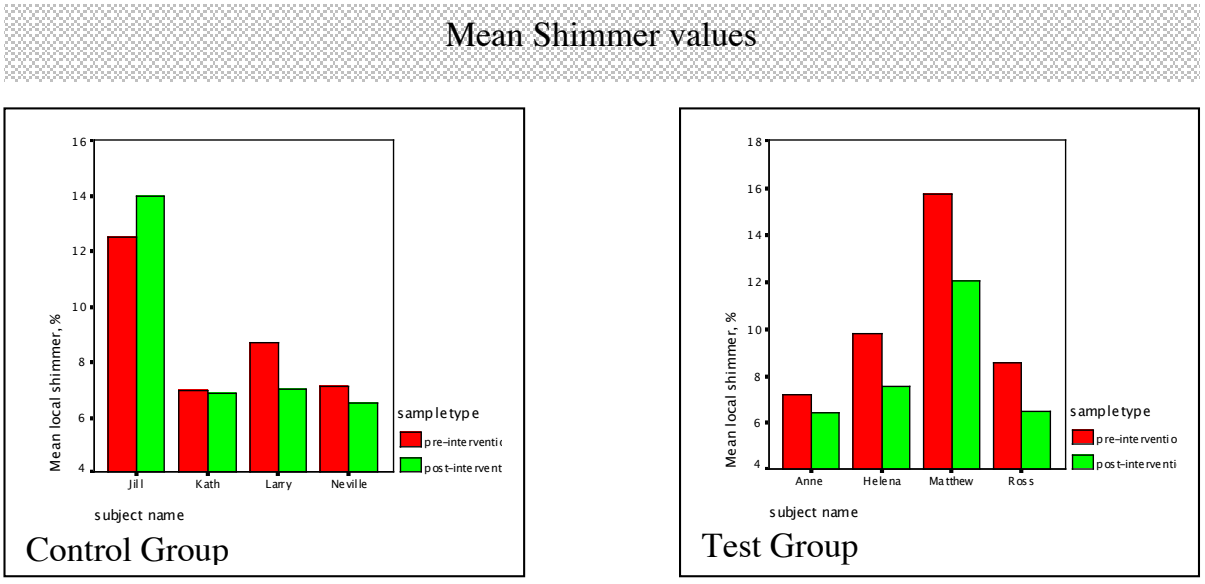


Figure 2: mean shimmer values before and after intervention a) control group b) test group

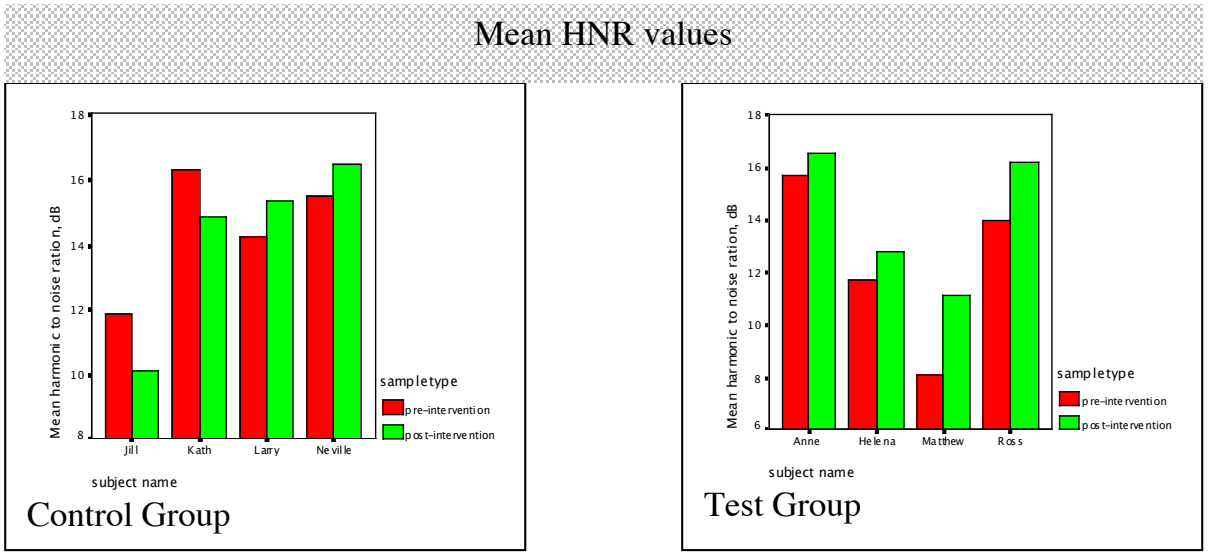


Figure 3: mean HNR values before and after intervention a) control group b) test group

two of the subjects in the control group, and decreased for two (Figure 3). The largest and most consistent changes are observed in the subject, Matthew, in the test group: all changes for Matthew are positive and indicate improved vocal quality post-intervention.

Measurements of voice quality and pitch, including mean, minimum, maximum and standard deviation were compared in all subjects and tested for significance using the Wilcoxon Signed Ranks Test. A change in median speaking pitch was significant for one of the control group and for one subject in the test group. Positive changes in HNR, jitter and shimmer were significant for two subjects in the test group (Matthew and Ross: Figures 1-3).

Perceptual measures of intelligibility and voice quality

According to scoring of the transcriptions made by three judges, the intelligibility of the test group improved post-intervention. However, the intelligibility of three of the subjects in the control group also improved.

The judges perceived an overall improvement in voice quality for three subjects from the test group (Helena, Matthew and Ross: Anne was omitted from this test as her voice quality was affected by a cold post-intervention). Although the judges were consistent in their judgements of changes in overall voice quality, they differed in their judgements of nasality, harshness and breathiness. The results for the test group were compared to the results for the control group and were tested for significance using Wilcoxon paired tests: the changes were not significant ($p<0.5$).

Matthew: a case study

Of the test group, the judges perceived a greater change in the voice of one male subject with DS, Matthew. In addition, the researcher had received an e-mail from his mother a week after the final singing session in which she stated that she believed that Matthew's voiced had become softer since taking part in singing. At the onset of the study, Matthew's singing voice and speaking voice were loud and harsh. Targets were aimed at producing a resonant tone with a soft onset, primarily through

humming exercises, and through breathing exercises. There was no progress for three sessions, due to the subject lifting and bracing his shoulders when asked to hum or breathe in. However, progress was made during the fifth session when humming was enabled following a new tactic, and soft onset was achieved following this. Matthew was subsequently able to sing with a little more variation of volume, and could sing with more rhythmic awareness when given a visual cue.

In addition to a perceived change in vocal quality, there were notable changes in acoustic voice perturbation measures for Matthew (see Figures 1-3). Data from voice reports for all speech samples showed an increase post-intervention in median pitch and HNR, and a decrease in jitter and shimmer. The data were analysed for significance using the Wilcoxon Signed Ranks test. Changes in mean pitch, median pitch, HNR, jitter and shimmer were significant ($p<0.01$). For those samples of speech that were the same in lexical content before and after intervention, data from the same vowel in each syllable of matched words revealed a post-intervention increase in HNR and reduction in shimmer and jitter values. Changes in jitter were significant ($p<0.01$) and changes in shimmer and HNR were highly significant ($p<0.001$).

Discussion

Acoustic measures

HNR had increased for all of the subjects in the test group, and for two of the subjects in the control group). HNR indicates how much extraneous noise there is in the voice signal, and as such, reflects how efficiently the vocal folds function: for non-pathological voices, the HNR is in the region of 15-25 dB. Studies have indicated that for pathological voices, HNR is lower and is a suitable indicator for indicating a pathological voice, and for measuring differences and changes in the vocal organ (Moura et al., 2007; Niedzielska, 2001). The improvement in HNR was significant for one subject (Matthew) in the test group, indicating an improvement in vocal fold functioning.

Studies have reported correlations between elevated measures of jitter and shimmer, and atypical voice quality (Hirano et al., 1988). Both values are reportedly higher for children with DS (Moura et al., 2007). Jitter values measure irregularities in frequency, reflecting how regularly the vocal folds are opening and closing; and shimmer measures irregularities in amplitude of the vocal signal, reflecting the coordination between vocal fold action and breath flow. Improvements in jitter were measured in three of the four subjects in the test group, and in one of the four subjects in the control group: of these, the change was significant in two of the test group indicting an improvement in vocal fold functioning following intervention. Improvements in shimmer were measured in all four subjects in the test group and three of the four subjects in the control group: of these, the change was significant for two subjects in the test group.

Perceptual measures

The results of the intelligibility tests are indeterminate, as improvements were seen in both groups. There are a number of confounds that make the tests unreliable. In particular, results were highly variable between judges, and it is likely that the ability of each judge to calibrate an individuals speech was a key factor (Kent, 1992). Secondly, the words in some of the spontaneous samples were difficult to comprehend in the absence of contextual cues, and given the pre-existing speech difficulties of the participants. Furthermore, speech production in people with DS is inconsistent (Dodd 1976; Stoel-Gammon, 2004), which means that single words may vary between productions

in segmental and suprasegmental features as a matter of course. The verbal responses of participants before and after intervention did not always match in segmental, suprasegmental or lexical content, making it difficult to establish or measure changes.

The judgements made by naïve listeners indicate that the voice quality of the test subjects had improved, relative both to samples collected before intervention and to results collected from subjects in the control group. Although these results were not significant, the results would indicate that the objective data obtained from Praat could be correlated with perceptual changes in the voices of some of the subjects, confirming the validity of this method of analysing changes in voice.

Matthew

The data from the perceptual tests for Matthew support both the acoustic data and the observation made by his mother that his voice had become softer following the intervention. Matthew presented at the outset with a voice that is typical to many with DS: gruff, with harsh onset (Montague and Hollien, 1973; Moran, 1986). Over the course of the singing lessons, he worked on eliminating tension in his breathing and on achieving a soft onset through humming exercises. In healthy voices, research has shown that such warm-up exercises can improve phonation the vocal folds open and close with greater efficiency, and with more economic use of breath (Amir et al., 2005; Elliot et al., 1995). It is therefore likely that the changes in Matthew's voice quality are due to an improvement in vocal fold functioning as a result of exercising the vocal folds.

For Matthew, the positive post-intervention changes in jitter, shimmer and HNR indicate a change in vocal fold functioning. Laver (1980) described the 'harsh' voice quality as characterised by irregularities of the waveform, additional noise in the signal, and perturbations of pitch, all caused by excessive laryngeal tension which causes the vocal folds to approximate harshly (Laver, 1980). It is therefore likely that the warm-up and singing exercises reduced laryngeal tension and promoted more efficient vocal fold action. Matthew's mean fundamental frequency pre-intervention was 160 Hz, with the minimum consistently below 90 Hz at the lower end of the pitch range for male modal voice, and in keeping with the descriptions of harsh voice as presented in Laver (1980). Post intervention, the mean fundamental frequency was closer to 190 Hz. Changes in pitch are made at the vocal fold level by adjusting both the length of the vocal folds and their tension or stiffness through adjustments to the surrounding laryngeal muscles, and may also be made through increasing the flow of air through the glottis (Miller, 1996; Husler and Rodd-Marling, 1976). An overall change in fundamental frequency during speech could therefore be attributed to relaxing and toning the muscles involved in changing pitch, and in toning and relaxing the muscles associated with breathing.

Matthew's voice changed in both mean and median pitch at a significant level; however, so did the median pitch of a member of the control group. The fundamental frequency of a speaker is related to their emotional state (Traunmüller and Eriksson, 1993) as well as to physiological attributes. It is therefore possible that the changes in median pitch reflect changes in the participants' emotional states as much as it does changes in vocal fold use. However, for Matthew, the change in pitch accompanied statistically-significant reduction in voice perturbation measures and a perceived improvement in voice quality. These data suggest a change in voice function. This is supported by evidence from visual analysis of the spectrographic data (not shown). Visual analysis showed that the majority of pitch contours in samples taken pre-intervention revealed a sudden dip in pitch as

intensity increases, indicating a difficulty in sustaining phonation as pressure increased. In contrast, the pitch contours taken from the post-intervention samples revealed fewer sudden dips, indicating better coordination between breath flow and phonation.

Conclusion

The aim of this study was to investigate whether singing tuition could affect the intelligibility of the speech of young adults with learning difficulties. The results indicate that there were no significant improvements in intelligibility for the test group as judged by scoring the transcripts of naïve listeners. However, the results suggest that singing tuition may provide benefits in terms of improved phonation for some of this population. Significant changes in voice parameters were found for two of the test group subjects, both males with DS, suggesting an improvement in phonation for these individuals. These results are in keeping with previous research into the effects of vocal warm-ups on phonation. Improvements in HNR, jitter, shimmer and mean fundamental frequency for Matthew suggest a change in vocal mode, possibly as a result of improved vocal fold functioning as a result of combined vocal and breathing exercises. Further research is required in order to pinpoint the exact nature and cause of the changes, and the endurance of any changes beyond the intervention period.

The outcome of the study was reliant on successful teaching. The teaching sessions were experimental in that there were no guidelines with respect to how best to teach this population, and the researcher had not previously taught singing to this client group. Whilst every effort was made to modify exercises to take account of the subject's preferred learning styles before the intervention commenced, some of the exercises were unsuccessful with the test group. For example, none of the participants in the test group were able to develop sufficient awareness or control of their breathing to affect a change at this level, and alternative methods had to be found to encourage deep inhalation without muscular tension: the success of methods used depended on the individual. Whilst all the subjects in the test group made progress in singing, the teaching could be improved in light of this and subsequent experience. Future studies could consider more extensive use of visual feedback, and whilst the group setting may confer advantages in terms of social skills and motivation, sessions could be tailored to include 1:1 sessions by using support workers to oversee group activities.

The results in this study suggest that there were significant changes in phonation following intervention for two subjects with Down's syndrome. Future research may consider isolating the effects of specific exercises involved, as the effects observed could be attributed to changes in muscle tone due to breathing exercises, relaxation exercises or humming exercises; alternatively they may have been due to the removal of mucus from vocal folds through use. The effect of improving vocal quality in this population is unknown but research suggests that improvements in vocal quality could have implications for intelligibility, and that this might impact positively on expressive language (Pryce, 2004). The exercises used are easily replicated and implemented; establishing which, if any, are important could lead to such exercises being incorporated into daily exercises, with possibly far-reaching effect.

References (MSc Summary)

Amir, O., Amir, N., and Michaeli, O. 2005, Evaluating the influence of warmup on singing voice quality using acoustic measures. *Journal of Voice*, 19 (2), 252-260

Dodd, B.,J., 1976. A comparison of the phonological systems of mental age matched normal, severely abnormal and Downs Syndrome Children. *British Journal of Disorder of Communication* 11, 27-42

Elliot, N., Sundberg, J., and Gramming, P., 1995. What happens during vocal warmup? *Journal of Voice*, 19 (1), 37-44

Hirano, M, Hibi, S, Yoshida, T, Hirade, Y, Kasuya, H, Kikuchi, Y., 1988. Acoustic analysis of pathological voice. *Acta otolaryngol, 105: 432-438*

Husler, F. and Rodd-Marling, Y., 1976. *Singing: The Physical Nature of the Vocal Organ*. London: Hutchinson and Co.

Kent, R.D., 1992. *Intelligibility in Speech Disorders: Theory, measurement and management*.USA: John Benjamins Publishing Company

Laver, J., 1980, *The Phonetic Description of Voice Quality*. UK: Cambridge University Press

Moran, J., 1986, Identification of Downs Syndrome Adults from Prolonged Vowel Samples. *Journal of Communication Disorders,* 19, 387-394

Montague, J.C., and Hollien, H., 1973, Perceived voice quality disorders in Downs Syndrome children. *Journal of Communication Disorders,* 6, 6-87

Moura, C.M, Cunha, L.M., Vilarinho, H., Cunha, M.J., Freitas, D., Palha, M., Pueschel, S.M., Pais-Clememt, M., 2007, Voice Parameters in Children with Down Syndrome. *Journal of Voice*, in press

Niedzielska, G., 2001 Acoustic analysis in the diagnosis of voice disorders in children. *International Journal of Pediatric Otorhinolaryngology*, 57, 189-193

Pryce, M., 2004. The voice of people with Down syndrome: An EMG biofeedback study. *Down Syndrome Research and Practice*, 2(3), 106-111.

Stoel-Gammon, C (2004) Down Syndrome phonology: Developmental patterns and intervention strategies *The Down Syndrome Educational Trust* retrieved 23/7/2006 from http://www.down-syndrome.net/library/periodicals/dsrp/07/3/93/DSRP-07-3-093-EN-GB.htm

Traunmüller, H., and Eriksson, A., 1993, The frequency range of the voice fundamental in the speech of male and female adults. Retrieved 20th April 2007 from: http://www.ling.su.se/staff/hartmut/f0_m&f.pdf.

References

Abbeduto, L., Pavetto, M., Kesin, E., Weissman, M., Karadottir, S., O'Brien, A., & Cawthon, S. (2001). The linguistic and cognitive profile of Down syndrome: Evidence from a comparison with fragile X syndrome. *Down Syndrome Research and Practice*, *7*(1), 9-15

Aichinger, P. (2015). *Diplophonic Voice - Definitions, models, and detection*. Unpublished PhD thesis, Graz University of Technology, Vienna, Austria. Retrieved from https://www.spsc.tugraz.at/sites/default/files/Thesis_finalVersion.pdf

Aichinger, P., Schneider-Stickler, B., Bigenzahn, W., Fuchs, A. K., Geiger, B., Hagmuller, M., & Kubin, G. (2013). Double pitch marks in diplophonic voice. *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings, 7437–7441*. http://doi.org/10.1109/ICASSP.2013.6639108

Albertini, G., Bonassi, S., Dall'Armi, V., Giachetti, I., Giaquinto, S., & Mignano, M. (2010). Spectral analysis of the voice in Down Syndrome. *Research in Developmental Disabilities*, 31(5), 995–1001. http://doi.org/10.1016/j.ridd.2010.04.024

Alcock, K. (2006). The development of oral motor control and language. *Down Syndrome Research and Practice*, 11(1), 1-8.

Alexander, M., Petri, H., Ding, Y., Wandel, C., Khwaja, O., & Foskett, N. (2015). Morbidity and medication in a large population of individuals with Down syndrome compared to the general population [Electronic version]. *Developmental Medicine & Child Neurology*, 58(3), 246-254. http://doi.org/10.1111/dmcn.12868

Alm, P. A. (2004). Stuttering and the basal ganglia circuits: A critical review of possible relations. *Journal of Communication Disorders*, *37*(4), 325–369. http://doi.org/10.1016/j.jcomdis.2004.03.001

Almeida, G. L., Corcos, D. M., & Hasan, Z. (2000). Horizontal-plane arm movements with direction reversals performed by normal individuals and individuals with Down syndrome. *Journal of neurophysiology*, 84(4), 1949-1960.

Anderson, J. D., & Wagovich, S. A. (2011). Relationships among linguistic processing speed, phonological working memory, and attention in children who stutter. *Journal of Fluency and Discord*, *35*(3), 216–234. http://doi.org/10.1016/j.jfludis.2010.04.003.Relationships

Anvari, S. H., Trainor, L. J., Woodside, J., & Levy, B. A. (2002). Relations among musical skills, phonological processing, and early reading ability in preschool children. *Journal of Experimental Child Psychology*, *83*(2), 111–130. http://doi.org/10.1016/S0022-0965(02)00124-8

Archibald, L. M., & Gathercole, S. E. (2006). Nonword repetition: A comparison of tests. *Journal of Speech, Language, and Hearing Research*, 49(5), 970-983

Aronson, A. E., & Bless, D. (2011). *Clinical voice disorders* (4th Ed.). New York: Thieme.

Baddeley, A. D., & Hitch, G. (1974). Working memory. *The psychology of learning and motivation*, 8, 47-89.

Baddeley, A., & Jarrold, C. (2007). Working memory and Down syndrome. *Journal of Intellectual Disability Research*, 51(12), 925-931.

Bailly, L., Henrich, N., & Pelorson, X. (2010). Vocal fold and ventricular fold vibration in period-doubling phonation: physiological description and aerodynamic modeling. *The Journal of the Acoustical Society of America*, 127(5), 3212–3222. http://doi.org/10.1121/1.3365220

Bajaj, A. (2007). Working memory involvement in stuttering: Exploring the evidence and research implications. *Journal of fluency disorders*, 32(3), 218-238.

Baken, R. J., & Orlikoff, R. F. (2000). *Clinical measurement of speech and voice* (2nd Ed.). San Diego, USA: Singular Publishing.

Baker, J.E. and Lane, R.D. (2008). Emotion processing deficits in emotional voice disorders. In Izdebski, K. (Ed.). Emotions in the Human Voice, Volume 3: Culture and Perception (Vol. 3, pp. 105-136). Abingdon: Plural Publishing.

Ball, M. J., Esling, J., & Dickson, C. (1995). The VoQS system for the transcription of voice quality. *Journal of the International Phonetic Association*, 25(02), 71-80.

Barker J. (1999). Singing and music as aids to language development and its relevance for children with Down syndrome. *Down Syndrome News and Update*. 1999;1(3);133-135.

Barnes, E. F., Roberts, J., Mirrett, P., Sideris, J., & Misenheimer, J. (2006). A comparison of oral structure and oral-motor function in young males with Fragile X syndrome and Down syndrome. *Journal of Speech, Language, and Hearing Research*, 49(4), 903-917.

Beck, J. M. (1988). *Organic variation and voice quality*. Doctoral dissertation, University of Edinburgh, Edinbugh, Scotland.

Beck, J.M. (2010). Organic Variation of the Vocal Apparatus. In Hardcastle, W.J., Laver, J. and Gibbon, F.E. (Eds.), *The handbook of phonetic science*s (2nd ed., pp 155-201) Wiley-Blackwell.

Beckman, D. A, Wold, D. C., & Montague, J. C. (1983). A noninvasive acoustic method using frequency perturbations and computer-generated vocal-tract shapes. *Journal of Speech and Hearing Research*, *26*(2), 304–314.

Beilock, S.L. and Ramirez, G (2011). On the interplay of emotion and cognitive control: implications for enhancing academic achievement. In Mestre, J.P. and Ross, B.H. (Eds.), The psychology of learning and motivation (vo.55, pp. 138-162). London: Academic Press

Bell-Berti, F., & Raphael, L. J. (1983). On identifying syllable boundaries. *Speech Communication*, 2(2), 145-147.

Bella, S. D., Berkowska, M., & Sowiński, J. (2011). Disorders of pitch production in tone deafness. *Frontiers in Psychology*, *2*(JUL), 1–11. http://doi.org/10.3389/fpsyg.2011.00164

Benda, C. E. (1965) Mongolism. In C. H. Carter (Ed.), *Medical Aspects of Mental Retardation*. Springfield: Illinois

Bennett SJ. (2009). Music may improve verbal memory – Implications for children with Down syndrome. *Down Syndrome Research and Practice*. 2009;12(3);173-174

Bennetts, L. K., & Flynn, M. C. (2002). Improving the classroom listening skills of children with Down syndrome by using sound-field amplification. *Down's Syndrome, Research and Practice*, *8*(1), 19–24. http://doi.org/10.3104/reports.124

Berkowska, M., & Dalla Bella, S. (2009). Reducing linguistic information enhances singing proficiency in occasional singers. *Annals of the New York Academy of Sciences*, 1169(August), 108–111. http://doi.org/10.1111/j.1749-6632.2009.04774.x

Berkowska, M., & Dalla Bella, S. (2009a). Acquired and congenital disorders of sung performance: A review. *Advances in Cognitive Psychology*, 5, 69-83.

Besson, M., Chobert, J., & Marie, C. (2011). Transfer of Training between Music and Speech: Common Processing, Attention, and Memory. *Frontiers in Psychology*, *2*(May), 94. http://doi.org/10.3389/fpsyg.2011.00094

Besson, M., Schön, D., Moreno, S., Santos, A., & Magne, C. (2007). Influence of musical expertise and musical training on pitch processing in music and language. *Restorative Neurology and Neuroscience*, 25(3-4), 399–410. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/17943015

Bhagyalakshmi, G., Renukarya, A., & Rajangam, S. (2007). Metric analysis of the hard palate in children with Down syndrome-a comparative study. *Down Syndrome Research and Practice*, 12(1), 55-59.

Bispham, J. (2006). Rhythm in music: What is it? who has it? and why? *Music Perception*, *24*(2), 125–134.

Blacketer-Simmonds, D. A. (1953). An investigation into the supposed differences existing between mongols and other mentally defective subjects with regard to certain psychological traits. *The British Journal of Psychiatry*, 99(417), 702-719.

Blamey, P. J., Sarant, J. Z., Paatsch, L. E., Barry, J. G., Bow, C. P., Wales, R. J. et al. (2001). Relationships among speech perception, production, language, hearing loss, and age in children with impaired hearing. *Journal of Speech, Language, and Hearing Research : JSLHR*, *44*(2), 264–285. http://doi.org/10.1044/1092-4388(2001/022)

Boersma, P., & Weenink, D. (2011). Praat: Doing phonetics by computer [Computer program]. Version 5.3. 03. 2011.

Bolfan-Stosic, N., & Simunjak, B. (2007). Effects of hearing loss on the voice in children. *Journal of otolaryngology*, 36(2), 120-123.

Borsel, J. V. (1996). Articulation in Down's syndrome adolescents and adults. *European Journal of Disorders of Communication*, 31(4), 415-444.

Bosshardt, H.-G. (2006). Cognitive processing load as a determinant of stuttering: summary of a research programme. *Clinical Linguistics & Phonetics*, *20*(5), 371–85. http://doi.org/10.1080/02699200500074321

Bowen, C. (2015). *Children's speech sound disorders* (2nd Ed.). Chichester: John Wiley & Sons.

Brandt, A., Gebrian, M., & Slevc, L. R. (2012). Music and early language acquisition. *Frontiers in Psychology*, *3*(SEP), 1–17. http://doi.org/10.3389/fpsyg.2012.00327

Braunschweig, T., Flaschka, J., Schelhorn-Neise, P., & Döllinger, M. (2008). High-speed video analysis of the phonation onset, with an application to the diagnosis of functional dysphonias. *Medical Engineering and Physics*, *30*(1), 59–66. http://doi.org/10.1016/j.medengphy.2006.12.007

Bray, M. (2008). Speech production in people with Down syndrome. *Down Syndrome Research and Practice*, 60–61. http://doi.org/10.3104/reviews.2075

Brindle, 2005. *How to be brilliant at teaching music at KS1*. [Course handbook].

Brown-sweeney, S., & Smith, B. (1997). The development of speech production abilities in children with Down syndrome. *Clinical Linguistics & Phonetics*, 11(5), 345–362. http://doi.org/10.3109/02699209708985200

Buchanan, T. W., Laures-Gore, J. S., & Duff, M. C. (2014). Acute stress reduces speech fluency. *Biological Psychology*, *97*(1), 60–66. http://doi.org/10.1016/j.biopsycho.2014.02.005

Buckley S, Bird G. (2002). Cognitive development and education: perspectives on Down syndrome from a twenty-year research programme. In: M. Cuskelly, A. Jobling, S. Buckley S (Eds.), *Down syndrome across the life span* (pp. 66–80). London England: Whurr Publishers.

Buckley, S., & Sacks, B. (2001). An overview of the development of infants with Down syndrome (5-11 years). DSE Enterprises.

Bunn, L., Roy, E. A., & Elliott, D. (2007). Speech perception and motor control in children with Down syndrome. *Child Neuropsychology : A Journal on Normal and Abnormal Development in Childhood and Adolescence*, *13*(3), 262–275. http://doi.org/10.1080/09297040600770738

Bunton, K., & Leddy, M. (2011). An evaluation of articulatory working space area in vowel production of adults with Down syndrome. *Clinical Linguistics & Phonetics*, *25*(4), 321–34. http://doi.org/10.3109/02699206.2010.535647

Callan, D. E., Tsytsarev, V., Hanakawa, T., Callan, A. M., Katsuhara, M., Fukuyama, H., & Turner, R. (2006). Song and speech: Brain regions involved with perception and covert production. *NeuroImage*, *31*(3), 1327–1342. http://doi.org/10.1016/j.neuroimage.2006.01.036

Cantor, G. N., & Girardeau, F. L. (1959). Rhythmic discrimination ability in mongoloid and normal children. *American journal of mental deficiency*, *63*(4), 621-625.

Capio, C. M., & Rotor, E. R. (2010). Fundamental movement skills among Filipino children with Down syndrome. *Journal of Exercise Science & Fitness*, 8(1), 17-24.

Capone, G. T. (2004). Down syndrome: Genetic insights and thoughts on early intervention. *Infants & Young Children*, 17(1), 45-58.

Carney, D. P. J., Brown, J. H., & Henry, L. A. (2013). Executive function in Williams and Down syndromes. *Research in Developmental Disabilities*, *34*(1), 46–55. http://doi.org/10.1016/j.ridd.2012.07.013

Carr, J. (2005). Stability and change in cognitive ability over the life span: a comparison of populations with and without Down's syndrome. *Journal of Intellectual Disability Research*, 49(12), 915-928.

Carroll, D. (1996). A study of the effectiveness of an adaptation of melodic intonation therapy in increasing the communicative speech of young children with Down syndrome. Unpublished dissertation, McGill University, Montreal, Canada.

Cason, N., & Schön, D. (2012). Rhythmic priming enhances the phonological processing of speech. *Neuropsychologia*, 50(11), 2652-2658.

Cason, N., Astésano, C., & Schön, D. (2015). Bridging music and speech rhythm: Rhythmic priming and audio–motor training affect speech perception. *Acta psychologic*a, 155, 43-50.

Cason, N., Hidalgo, C., Isoard, F., Roman, S., & Schön, D. (2015). Rhythmic priming enhances speech production abilities: Evidence from prelingually deaf children. *Neuropsychology*, 29(1), 102.

Cavalli, L., & Hirson, a. (1999). Diplophonia reappraised. *Journal of Voice*, 13(4), 542–556. http://doi.org/10.1016/S0892-1997(99)80009-5

Cawson, R. A., & Odell, E. W. (2008). *Cawson's essentials of oral pathology and oral medicine*. Elsevier Health Sciences.

Chandrasekaran, B. Xie, Z., Reetke, R. (2015). Music training and the neural processing of speech: a critical review of the literature. In Agwuele and Lotto (Eds.), *Essays in speech processes* (pp. 139-174). Sheffield: Equinox.

Chang, H. W., & Trehub, S. E. (1977). Infants' perception of temporal grouping in auditory patterns. *Child development*, 1666-1670.

Channell, M. M., Thurman, A. J., Kover, S. T., & Abbeduto, L. (2014). Patterns of change in nonverbal cognition in adolescents with Down syndrome. *Research in Developmental Disabilities*, *35*(11), 2933–2941. http://doi.org/10.1016/j.ridd.2014.07.014

Chapman, R. S., & Hesketh, L. J. (2001). Language, cognition, and short-term memory in individuals with Down syndrome. *Down's Syndrome, Research and Practice*, *7*(1), 1–7. http://doi.org/10.3104/reviews.108

Chen C-C (JJ), Ringenbach S.D.R., Biwer A., Riekena, A. A cerebral lateralization of the EEG during perceptual-motor integration in young adults with Down syndrome: A descriptive study. *Brazilian Journal of Motor Behavior*. 2015: 9(2): 1-7.

Childers, D. G., & Lee, C. K. (1991). Vocal quality factors: analysis, synthesis, and perception. *The Journal of the Acoustical Society of America*, 90(5), 2394–410. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1837797

Christiner, M., & Reiterer, S. M. (2013). Song and speech: Examining the link between singing talent and speech imitation ability. *Frontiers in Psychology*, *4*(NOV). http://doi.org/10.3389/fpsyg.2013.00874

Chu, S. Y., & Barlow, S. M. (2016). A Call for Biomechanics to Understand Hypotonia and Speech Movement Disorders in Down Syndrome. *Advances in communication disorder*. http://www.avidscience.com/wp-content/uploads/2016/04/ACD-16-01_April-16-2016.pdf

Chua, R., Weeks, D., & Elliott, D. (1996). A functional systems approach to understanding verbal-motor integration in individuals with Down syndrome. *Down Syndrome Research and Practice*, *4*(1), 25–36. Retrieved from http://www.down-syndrome.org/research-practice/4/1/25/

Cleland, J., Wood, S., Hardcastle, W., Wishart, J., & Timmins, C. (2010). The relationship between speech oromotor, language and cognitive abilities in children with Down's syndrome. *International Journal of Language & Communication Disorders*, *45*, 83–95. http://doi.org/10.1016/j.buildenv.2006.10.027

Clift, S. (2012). Singing, wellbeing, and health. In *Music, Health and Wellbeing (pp. 113-124)*. Oxford: Oxford University Press.

Coelho, A. C., Medved, D. M., & Brasolotto, A. G. (2015). Hearing loss and the child. In *Update on hearing loss: Hearing loss and the voice* (Vol. 228, pp. 28–30). Retrieved from Ana Cristina Coelho, Daniela Malta Medved and Alcione Ghedini

Comblain, A., & Rondal, J. (1996). Language in adults with Down syndrome. *Down Syndrome Research and Practice*, *4*(1), 3–14. http://doi.org/10.3104/reviews.58

Corrigall, K.A. & Schellenberg, G.E. (2016). Musical cognition in childhood. In McPherson, G. E. (Ed., pp 81-101). *The child as musician: A handbook of musical development*. Oxford: Oxford University Press.

Corriveau, K. H., & Goswami, U. (2009). Rhythmic motor entrainment in children with speech and language impairments: tapping to the beat. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *45*(1), 119–30. http://doi.org/10.1016/j.cortex.2007.09.008

Costanzo, F., Varuzza, C., Menghini, D., Addona, F., Gianesini, T., & Vicari, S. (2013). Executive functions in intellectual disabilities: A comparison between Williams syndrome and Down syndrome. *Research in Developmental Disabilities*, *34*(5), 1770–1780. http://doi.org/10.1016/j.ridd.2013.01.024

Couzens, D., & Cuskelly, M. (2014). Cognitive strengths and weaknesses for informing educational practice. *Educating learners with Down syndrome research, theory, and practice with children and adolescents*, 41-59.

Couzens, D., Cuskelly, M., & Haynes, M. (2011). Cognitive development and down syndrome: age-related change on the Stanford-Binet Test. *American journal on intellectual and developmental disabilities*, 116(3), 181-204.

Couzens, D., Haynes, M., & Cuskelly, M. (2012). Individual and environmental characteristics associated with cognitive development in Down syndrome: a longitudinal study. *Journal of Applied Research in Intellectual Disabilities : JARID*, *25*(5), 396–413. http://doi.org/10.1111/j.1468-3148.2011.00673.x

Coyne, N. M., Dwyer, M. L., Kennedy, M., & Petter, N. M. (2000). The effects of systematic implementation of music on behavior and performance of the Special Needs student. Unpublished dissertation, Saint Xavier University Chicago, Illinois.

Cuckle, P. Getting in and staying there: Children with Down syndrome in mainstream schools. *Down Syndrome Research and Practice*. 1999;6(2);95-99.

Cumming, R., Wilson, A., Leong, V., Colling, L. J., & Goswami, U. (2015). Awareness of rhythm patterns in speech and music in children with Specific Language Impairments. *Frontiers in Human Neuroscience*, *9*(December), 672. http://doi.org/10.3389/fnhum.2015.00672

Da Silva, V. Z. M., de França Barros, J., de Azevedo, M., de Godoy, J. R. P., Arena, R., & Cipriano, G. (2010). Bone mineral density and respiratory muscle strength in male individuals with mental retardation (with and without Down Syndrome). *Research in Developmental Disabilities*, *31*(6), 1585–1589. http://doi.org/10.1016/j.ridd.2010.05.003

Dalla Bella, S, & Peretz, I. (2003). Congenital amusia interferes with the ability to synchronize with music. *Annals of the New York Academy of Sciences*, 999(1), 166-169.

Dalla Bella, S., Berkowska, M., & Sowiński, J. (2015). Moving to the beat and singing are linked in humans. *Frontiers in Human Neuroscience*, 9(November 2015), 1–13. http://doi.org/10.3389/fnhum.2015.00663

Dalla Bella, S., Giguère, J. F., & Peretz, I. (2007). Singing proficiency in the general population. *The journal of the Acoustical Society of America*, 121(2), 1182-1189.

Dalla Bella, S., Tremblay-Champoux, A., Berkowska, M., & Peretz, I. (2012). Memory disorders and vocal performance. *Annals of the New York Academy of Sciences*, 1252(1), 338-344.

Das, B., Chatterjee, I., & Kumar, S. (2013). Laryngeal Aerodynamics in Children with Hearing Impairment versus Age and Height Matched Normal Hearing Peers. *ISRN Otolaryngology*, *2013*. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&AuthType=cookie,ip,shib&db=awn&AN=23970976&site=ehost-live

Davidow, Jason H; Bothe, Anne K; Yun, H. (2011). Systematic Studies of Modified Vocalization: speech production changes during a cariation of metronomic speech in persons who do and do not stutter. *Journal of Fluency and Discord*, *36*(2), 93–109. http://doi.org/10.1016/j.jfludis.2011.03.003.Systematic

Davidson, L. (1985). Tonal structures of children's early songs. *Music Perception*, *2*(3), 361–373. http://doi.org/10.2307/40285304

Davidson, L., & Scarlett, W. G. (1987). When is a song a song? The development of singing in early childhood. *Day Care and Early Education*, 14(3), 30-31.

Davidson, L., McKernon, P., & Gardner, H. (1981). The acquisition of song: A developmental approach. Documentary report of Ann Arbor Symposium: Applications of psychology to the teaching and learning of music (pp. 301-315). Reston, VA: Music Educators National Conference

Davis, A. S. (2008). Children with down syndrome: Implications for assessment and intervention in the school. *School psychology quarterly*, 23(2), 271.

de l'Etoile, S. K. (2015). Self-Regulation and Infant-Directed Singing in Infants with Down Syndrome. *Journal of Music Therapy*, *52*(2), 195–220. http://doi.org/10.1093/jmt/thv003

de Pijper, n.d. Semitone Conversions. http://users.utu.fi/jyrtuoma/speech/semitone.html. Accessed Nov 2015-January 2016

Dehqan, A., & Scherer, R. C. (2011). Objective voice analysis of boys with profound hearing loss. *Journal of Voice*, 25(2), e61–e65. http://doi.org/10.1016/j.jvoice.2010.08.006

Dejonckere, P. (1998). Effect of louder voicing on acoustical measurements in dysphonic patients. *Logopedics Phoniatrics Vocology*, 23(2), 79-84.

Dejonckere, P. H., & Lebacq, J. (1983). An analysis of the diplophonia phenomenon. *Speech Communication*, *2*, 47–56.

Delage, H., & Tuller, L. (2007). Language development and mild-to-moderate hearing loss: does language normalize with age? *Journal of Speech, Language, and Hearing Research : JSLHR*, *50*(5), 1300–1313. http://doi.org/10.1044/1092-4388(2007/091)

Demorest, S. M., & Pfordresher, P. Q. (2015). Singing accuracy development from K-adult. *Music Perception: An Interdisciplinary Journal*, 32(3), 293-302.

Department for Education. (2011). *The importance of Music. A National Plan for Music Education.* Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/180973/DFE-00086-2011.pdf.

Department for Education. (2013). *National curriculum in England: primary curriculum.* London: HMSO.

Department for Education. (2014). *National curriculum in England: framework for key stages 1 to 4*. London: HMSO

Deutsch, D (2013). The processing of pitch combinations. In Deutsch, D. (Ed.), *Psychology of music* (3rd Ed., pp. 249-313). London: Elsevier.

Deutsch, D., & Feroe, J. (1981). The internal representation of pitch sequences in tonal music. *Psychological review*, 88(6), 503.

Dierssen, M. (2012). Down syndrome: the brain in trisomic mode. Nature Reviews. *Neuroscience,* 13(12), 844–58. http://doi.org/10.1038/nrn3314

DiGiammarino, M. (1990). Functional music skills of persons with mental retardation. *Journal of Music Therapy*, 27(4), 209-220.

Dingle, G. A., Brander, C., Ballantyne, J., & Baker, F. A. (2013). 'To be heard': The social and mental health benefits of choir singing for disadvantaged adults. Psychology of Music, 41(4), 405-421.

Dockrell, J., Lindsay, G., Roulstone, S., & Law, J. (2014). Supporting children with speech, language and communication needs: an overview of the results of the Better Communication Research Programme. *International Journal of Language & Communication Disorders*, *49*(5), 543-557.

Dodd, B., Zhu, H., Crosbie, S., Holm. A. and Ozanne. A. (2002). *Diagnostic Evaluation of Articulation and Phonology.* London: Psychology Corporation.

Dowling, W. J. (1978). Scale and contour: Two components of a theory of memory for melodies. *Psychological review*, 85(4), 341.

Dowling, W. J. (1982). Melodic information processing and its development. In Deutsch, D (ed)., *The Psychology of Music* (pp. 413-430). New York: Academic Press

Drake, C., & Bertrand, D. (2012). The quest for universals in temporal processing in music. *The Cognitive Neuroscience of Music*, 17–27. http://doi.org/10.1093/acprof:oso/9780198525202.003.0002

Drake, C., & Gérard, C. (1989). A psychological pulse train: how young children use this cognitive framework to structure simple rhythms. *Psychological Research*, *51*(1), 16–22. http://doi.org/10.1007/BF00309271

Drake, C., Jones, M. R., & Baruch, C. (2000). The development of rhythmic attending in auditory sequences: attunement, referent period, focal attending. *Cognition*, 77(3), 251–88. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11018511

Dressler, A., Perelli, V., Feucht, M., & Bargagna, S. (2010). Adaptive behaviour in Down syndrome: a cross-sectional study from childhood to adulthood. *Wiener Klinische Wochenschrift*, 122(23-24), 673–80. http://doi.org/10.1007/s00508-010-1504-0

Duckworth, M., Allen, G., Hardcastle, W., & Ball, M. (1990). Extensions to the International Phonetic Alphabet for the transcription of atypical speech. *Clinical Linguistics & Phonetics*, 4(4), 273-280.

Dunn, L. M., Dunn, L. M., Whetton, C., & Burley, J. (1997). British Picture Vocabulary Scale (2nd ed.). Windsor, England: NFER-Nelson.

Dykens, E. (1996). The draw-a-person task in persons with mental retardation: What does it measure?. *Research in developmental disabilities*, 17(1), 1-13.

Edenfield, T. N., & Hughes, J. E. (1991). The relationship of a choral music curriculum to the development of singing ability in secondary students with Down syndrome. *Music Therapy Perspectives*, 9(1), 52-55.

Elliott, D., & Weeks, D. J. (1993). Cerebral specialization for speech perception and movement organization in adults with Down's syndrome. *Cortex*, 29(1), 103-113.

Elliott, D., Weeks, D. J., & Chua, R. (1994). Anomalous cerebral lateralization and Down syndrome. *Brain and Cognition*, 26(2), 191-195.

Elmer, S. S. (2011). Human singing: Towards a developmental theory. *Psychomusicology: Music, Mind and Brain*, 21(1-2), 13.

Evans, S. E., Beauchamp, G., & John, V. (2015). Learners' experience and perceptions of informal learning in Key Stage 3 music: a collective case study, exploring the implementation of Musical Futures in three secondary schools in Wales. *Music Education Research*, *17*(1), 1-16.

Falk, S., Müller, T., & Dalla Bella, S. (2014). Sensorimotor synchronization in stuttering children and adolescents. *Procedia-Social and Behavioral Sciences*, 126, 206-207.

Falk, S., Müller, T., & Dalla Bella, S. (2015). Non-verbal sensorimotor timing deficits in children and adolescents who stutter. *Frontiers in psychology*, 6.

Fawcus, (2009).The physiology of phonation. In Freeman, M. and Fawcus, M. (Eds.), Voice disorders and their management (3rd Ed., pp. 1-17). London: Whurr.

Findlay, E. (1971). *Rhythm and movement: applications of Dalcroze Eurhythmics*. Van Nuys, CA: Alfred Music.

Flaugnacco, E., Lopez, L., Terribili, C., Zoia, S., Buda, S., Tilli, S., ... & Schön, D. (2014). Rhythm perception and production predict reading abilities in developmental dyslexia. *Frontiers in Human Neuroscience*, 8(392), 10-3389.

Flipsen Jr., P. (1999). Prepausal accuracy in the speech of adults with Down syndrome. *Clinical Linguistics & Phonetics*, *13*(1), 55–65. http://doi.org/10.1080/026992099299248

Forgeard, M., Schlaug, G., Norton, A., Rosam, C., Iyengar, U., & Winner, E. (2008). The relation between music and phonological processing in normal-reading children and children with dyslexia. *Music Perception: An Interdisciplinary Journa*l, 25(4), 383-390.

Fourcin, A. (2009). Aspects of voice irregularity measurement in connected speech. *Folia Phoniatrica et Logopaedica*, *61*(3), 126–136. http://doi.org/10.1159/000219948

Fourcin, A., & Abberton, E. (2008). Hearing and phonetic criteria in voice measurement: clinical applications. *Logopedics, Phoniatrics, Vocology*, *33*(1), 35–48. http://doi.org/10.1080/14015430701251574

Fraile, R., & Godino-Llorente, J. I. (2014). Cepstral peak prominence: A comprehensive analysis. *Biomedical Signal Processing and Control, 14*, 42-54.

François, C., Chobert, J., Besson, M., & Schön, D. (2013). Music training for the development of speech segmentation. *Cerebral Cortex*, *23*(9), 2038–2043. http://doi.org/10.1093/cercor/bhs180

Fujii, S., & Wan, C. Y. (2014). The role of rhythm in speech and language rehabilitation: The SEP hypothesis. *Frontiers in Human Neuroscience*, 8, 777.

Fuks, L. (1999). Computer-aided musical analysis of extended vocal techniques for compositional applications. *VI Simpósio Brasileiro de Computaçäo Música, Rio de Janeiro*. Retrieved November, 4, 2007.

Garrett, B. (2014). *Exploring primary school teachers' motivation for music: an investigation into the impact of personal and social problems.* Doctoral dissertation, Lancaster University, Lancaster, England.

Gathercole, S. E. (2006). Nonword repetition and word learning: The nature of the relationship. *Applied Psycholinguistics*, *27*(04), 513–543. http://doi.org/10.1017/S0142716406060383

Gathercole, S. E., & Baddeley, A. D. (1993). Phonological working memory: A critical building block for reading development and vocabulary acquisition?. *European Journal of Psychology of Education*, 8(3), 259-272.

Gathercole, S. E., & Baddeley, A. D. (1996). Children's Test of Nonword Repetition. Psychological Corporation: London, UK.

Gathercole, S. E., Pickering, S. J., Ambridge, B., & Wearing, H. (2004). The structure of working memory from 4 to 15 years of age. *Developmental psychology*, 40(2), 177.

Gérard, C., & Auxiette, C. (1992). The processing of musical prosody by musical and nonmusical children. *Music Perception: An Interdisciplinary Journal*, 10(1), 93-125.

Gérard, C., & Drake, C. (1990). The inability of young children to reproduce intensity differences in musical rhythms. *Perception & Psychophysics*, *48*(1), 91–101. http://doi.org/10.3758/BF03205015

Giangreco, M. F. (2006). Foundational concepts and practices for educating students with severe disabilities. Instruction of students with severe disabilities, 5, 111-169.

Gibbon, F. E., McNeill, A. M., Wood, S. E., & Watson, J. M. (2003). Changes in linguapalatal contact patterns during therapy for velar fronting in a 10-year-old with Down's syndrome. *International Journal of Language & Communication Disorders*, 38(1), 47-64.

Gick, B., Wilson, I., & Derrick, D. (2013). *Articulatory phonetics*. Chichester, UK: Wiley-Blackwell.

Giddens, C. L., Barron, K. W., Byrd-Craven, J., Clark, K. F., & Winter, A. S. (2013). Vocal indices of stress: A review. *Journal of Voice*, *27*(3). http://doi.org/10.1016/j.jvoice.2012.12.010

Glenn, S. M., Cunningham, C. C., & Joyce, P. F. (1981). A study of auditory preferences in nonhandicapped infants and infants with Down's Syndrome. *Child Development*, 52(4), 1303-1307.

Glenn, S., & Cunningham, C. (2005). Performance of young people with Down syndrome on the Leiter-R and British picture vocabulary scales. *Journal of Intellectual Disability Research*, 49(4), 239-244.

Goodenough, F. L. (1963). Goodenough-Harris drawing test. New York: Harcourt Brace Jovanovich.

Gooding, L., & Standley, J. M. (2011). Musical development and learning characteristics of students: A compilation of key points from the research literature organized by age. *Update: Applications of Research in Music Education*, *30*(1), 32–45. http://doi.org/10.1177/8755123311418481

Gordon, E.E. (1979). Primary measures of music audiation (PMMA). Chicago: GIA

Gordon, R. L., Schön, D., Magne, C., Astésano, C., & Besson, M. (2010). Words and melody are intertwined in perception of sung words: EEG and behavioral evidence. *PLoS ONE*, 5(3). http://doi.org/10.1371/journal.pone.0009889

Goy, H., Fernandes, D. N., Pichora-Fuller, M. K., & van Lieshout, P. (2013). Normative Voice Data for Younger and Older Adults. *Journal of Voice*, *27*(5), 545–555. http://doi.org/10.1016/j.jvoice.2013.03.002

Grahn, J. A. (2012). See what I hear? Beat perception in auditory and visual rhythms. *Experimental Brain Research*, *220*(1), 51–61. http://doi.org/10.1007/s00221-012-3114-8

Grahn, J. A., & Schuit, D. (2012). Individual differences in rhythmic ability: Behavioral and neuroimaging investigations. *Psychomusicology: Music, Mind, and Brain*, 22(2), 105.

Gray, D. E. (2013). *Doing research in the real world*. Sage Publications Ltd.: London.

Greenberg, S., & Arai, T. (2001). The relation between speech intelligibility and the complex modulation spectrum. In INTERSPEECH (pp. 473-476).

Groen, M. A., Alku, P., & Bishop, D. V. M. (2008). Lateralisation of auditory processing in Down syndrome: A study of T-complex peaks Ta and Tb. *Biological Psychology*, *79*(2), 148–157. http://doi.org/10.1016/j.biopsycho.2008.04.003

Groß, W., Linden, U., & Ostermann, T. (2010). Effects of music therapy in the treatment of children with delayed speech development - results of a pilot study. *BMC Complementary and Alternative Medicine*, 10(2001), 39. http://doi.org/10.1186/1472-6882-10-39

Habib, M., & Besson, M. (2009). What do music training and musical experience teach us about brain plasticity? *Music Perception*, *26*(3), 279–286.

Hall, D., & Gathercole, S. E. (2011). Serial recall of rhythms and verbal sequences: Impacts of concurrent tasks and irrelevant sound. *Quarterly Journal of Experimental Psychology* (2006), 64(8), 1580–1592. http://doi.org/10.1080/17470218.2011.564636

Hallam, S. (2010). The power of music: Its impact on the intellectual, social and personal development of children and young people. *International Journal of Music Education*, *28*(3), 269–289. http://doi.org/10.1177/0255761410370658

Hallam, S. (2016). Musicality. In McPherson, G. E. (Ed.). *The child as musician: A handbook of musical development (2nd Ed.,67-80)*. Oxford: Oxford University Press.

Hallam, S., Creech, A. and McQueen, H. (2016) 'The perceptions of non music staff and senior management of the impact of the implementation of the Musical Futures approach on the whole school', British Journal of Music Education, 33(2), pp. 133–157. doi: 10.1017/S0265051716000139.

Hamilton, C. (1993). Investigation of the articulatory patterns of young adults with Down syndrome using electropalatography. *Down Syndrome Research and Practice*, 1(1), 15-28.

Hannon, E. E., & Trehub, S. E. (2005). Metrical categories in infancy and adulthood. *Psychological Science*, *16*(1), 48–55. http://doi.org/10.1111/j.0956-7976.2005.00779.x

Hargreaves, D. J. (1986). *The developmental psychology of music*. Cambridge: Cambridge University Press.

Hausen, M., Torppa, R., Salmela, V. R., Vainio, M., & Särkämö, T. (2013). Music and speech prosody: A common rhythm. *Frontiers in Psychology*, 4(SEP), 1–16. http://doi.org/10.3389/fpsyg.2013.00566

Hawkins S. 2014 Situational influences on rhythmicity in speech, music, and their interaction. *Philosophical Transactions of the Royal Society*. B 369: 20130398. http://dx.doi.org/10.1098/rstb.2013.0398

Heath, M., Welsh, T. N., Simon, D. A., Tremblay, L., Elliott, D., & Roy, E. A. (2005). Relative processing demands influence cerebral laterality for verbal-motor integration in persons with Down syndrome. *Cortex*, 41(1), 61-66.

Heman-Ackah, Y. D., Heuer, R. J., Michael, D. D., Ostrowski, R., Horman, M., Baroody et al. (2003). Cepstral peak prominence: A more reliable measure of dysphonia. *Annals of Otology, Rhinology and Laryngology*, *112*(4), 324–333.

Henley, D. (2011). *Music education in England: a review by Darren Henley for the Department for Education and the Department for Culture, Media and Sport*. London: Stationery Office.

Heselwood, B. C., Bray, M., & Crookston, I. (1995). Juncture, rhythm and planning in the speech of an adult with Down's syndrome. *Clinical Linguistics & Phonetics*, *9*(2), 121–137. http://doi.org/10.3109/02699209508985328

Ho, Y.-C., Cheung, M.-C., & Chan, A. S. (2003). Music training improves verbal but not visual memory: cross-sectional and longitudinal explorations in children. *Neuropsychology*, *17*(3), 439–450. http://doi.org/10.1037/0894-4105.17.3.439

Honda, K. (1983). Relationship between pitch control and vowel articulation. Haskins Laboratories Status Report on Speech Research, SR, 73, 269-82.

Honda, K., Hirai, H., Estill, J., Tohkura, Y. I., Fujimura, O., & Hirano, M. (1995). Contributions of vocal tract shape to voice quality: MRI data and articulatory modeling. *Vocal fold physiology: Voice quality control*, 23-38.

Hooper, J., Wigram, T., Carson, D., & Lindsay, B. (2008). A review of the music and intellectual disability literature. *Music Therapy Perspectives*, *26*, 66–79.

Hooper, J., Wigram, T., Carson, D., & Lindsay, B. (2010). The practical implication of comparing how adults with and without intellectual disability respond to music. *British Journal of Learning Disabilities*, 22–28. http://doi.org/10.1111/j.1468-3156.2010.00611.x

Houlahan, M., & Tacka, P. (2008). *Kodály today: a cognitive approach to elementary music education*. Oxford: Oxford University Press.

Howard, D. M., Brereton, J., Welch, G. F., Himonides, E., DeCosta, M., Williams, J., & Howard, A. W. (2007). Are real-time displays of benefit in the singing studio? An exploratory study. *Journal of voice*, 21(1), 20-34.

Howell, P. (2010). Speech control in children with Down syndrome. *Down Syndrome Research and Practice*, *12*(2), 98–102. http://doi.org/10.3104/reviews.2095

Huss, M., Verney, J. P., Fosker, T., Mead, N., & Goswami, U. (2011). Music, rhythm, rise time perception and developmental dyslexia: perception of musical meter predicts reading and phonology. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, 47(6), 674–89. http://doi.org/10.1016/j.cortex.2010.07.010

Iacono, T., Torr, J., & Wong, H. Y. (2010). Relationships amongst age, language and related skills in adults with Down syndrome. *Research in Developmental Disabilities*, 31(2), 568–76. http://doi.org/10.1016/j.ridd.2009.12.009

Ilari, B. (2014). Rhythmic engagement with music in early childhood: A replication and extension. *Journal of Research in Music Education*, 62(4), 332–343. http://doi.org/10.1177/0022429414555984

Ilari, B. S. (2002). Music perception and cognition in the first year of life. *Early Child Development and Care*, 172(3), 311-322.

Iversen, J. R., Patel, A. D., & Ohgushi, K. (2004). Perception of nonlinguistic rhythmic stimuli by American and Japanese listeners. In *Proceedings of the International Congress of Acoustics*, Kyoto.

Jackson, S. A., Treharne, D. A., & Boucher, J. (1997). Rhythm and language in children with moderate learning difficulties. *European Journal of Disorders of Communication*, 32(1), 99-108.

Jakobson, L. S., Cuddy, L. L., & Kilgour, A. R. (2003). Time tagging: A key to musicians' superior memory. Music Perception: *An Interdisciplinary Journal*, 20(3), 307-313.

Jäncke, L. (2012). The relationship between music and language. *Frontiers in Psychology*, *3*(APR), 2011–2012. http://doi.org/10.3389/fpsyg.2012.00123

Jarrold C, Baddeley AD (1997). Short-term memory for verbal and visuo-spatial information in Down's syndrome. *Cognitive Neuropsychiatry*. 1997;2:101- 122.

Jarrold, C., Baddeley, A. D., & Hewes, A. K. (2000). Verbal short-term memory deficits in Down syndrome: A consequence of problems in rehearsal?. *Journal of Child psychology and Psychiatry*, 41(2), 233-244.

Jarrold, C., Baddeley, A. D., Hewes, A. K., Leeke, T. C., & Phillips, C. E. (2004). What links verbal short-term memory performance and vocabulary level? Evidence of changing relationships among individuals with learning disability. *Journal of Memory and Language*, 50(2), 134–148. http://doi.org/10.1016/j.jml.2003.10.004

Jarrold, C., Baddeley, A., & Phillips, C. (1999). Down syndrome and the phonological loop: The evidence for, and importance of, a specific verbal short-term memory deficit. *Down Syndrome Research and Practice*, 6(2), 61-75.

Jarrold, C., Thorn, a. S. C., & Stephens, E. (2009). The relationships among verbal short-term memory, phonological awareness, and new word learning: Evidence from typical development and Down syndrome. *Journal of Experimental Child Psychology*, 102(2), 196–218. http://doi.org/10.1016/j.jecp.2008.07.001

Jeffries, K. J., Fritz, J. B., & Braun, a R. (2003). Words in melody: an H(2)15O PET study of brain activation during singing and speaking. *Neuroreport*, 14(5), 749–54. http://doi.org/10.1097/01.wnr.0000066198.94941.a4

Jellison, J.A. (2016). Inclusive music classrooms: a universal approach. In. McPherson, G. E. (Ed.), *The child as musician: A handbook of musical development* (2nd Ed., pp. 361-372). Oxford: Oxford University Press.

Johnstone, T., & Scherer, K. R. (1999). The Effects of emotions on voice quality. *Proceedings of the 14th International Conference of Phonetic Sciences*, 2029–2032.

Jou, J., & Harris, R. J. (1992). The effect of divided attention. *Bulletin of the Psychonomic Society*, 30(June 1991), 301–304. http://doi.org/10.3758/BF03330471

Jungers, M. K., Palmer, C., & Speer, S. R. (2002). Time after time: The coordinating influence of tempo in music and speech. *Cognitive Processing*, 2(614), 21–35.

Katz, G., & Lazcano-Ponce, E. (2008). Intellectual disability: definition, etiological factors, classification, diagnosis, treatment and prognosis [Discapacidad intelectual: definición, factores etiológicos, clasificación, diagnóstico, tratamiento y prognosis.]. salud pública de méxico, 50. Retrieved from

http://www.scielosp.org/scielo.php?script=sci_arttext&pid=S0036-36342008000800005&lng=en&nrm=iso&tlng=en

Kay Raining-Bird, E., & Chapman, R. S. (1994). Sequential recall in individuals with Down syndrome. *Journal of Speech, Language, and Hearing Research*, 37(6), 1369-1380.

Kay-Raining Bird, E., Cleave, P. L., & McConnell, L. (2000). Reading and phonological awareness in children with Down syndrome: A longitudinal study. *American Journal of Speech-Language Pathology*, 9(4), 319-330.

Keating, P., Garellek, M., & Kreiman, J. (2015). Acoustic properties of different kinds of creaky voice. *Proceedings of the 18th International Congress of Phonetic Sciences.*

Keller-Bell, Y., & Fox, R. A. (2007). A Preliminary study of speech discrimination in youth with Down syndrome. *Clinical linguistics & phonetics*, 21(4), 305-317.

Kent, R. D. (2000). Research on speech motor control and its disorders: A review and prospective. *Journal of Communication disorders*, 33(5), 391-428.

Kent, R. D., & Vorperian, H. K. (2013). Speech impairment in Down syndrome: A review. *Journal of Speech, Language, and Hearing Research*, *56*, 178–210. http://doi.org/10.1044/1092-4388(2012/12-0148

Keough, D., & Jones, J. a. (2009). The sensitivity of auditory-motor representations to subtle changes in auditory feedback while singing. *The Journal of the Acoustical Society of America*, 126(2), 837–846. http://doi.org/10.1121/1.3158600

Khawaja, M. A., Ruiz, N., & Chen, F. (2008, December). Think before you talk: An empirical study of relationship between speech pauses and cognitive load. In *Proceedings of the 20th Australasian Conference on Computer-Human Interaction: Designing for Habitus and Habitat* (pp. 335-338). ACM. - find and add

Kilgour, A. R., Jakobson, L. S., & Cuddy, L. L. (2000). Music training and rate of presentation as mediators of text and song recall. *Memory & cognition*, 28(5), 700-710.

Kirschner, S., & Tomasello, M. (2010). Joint music making promotes prosocial behavior in 4-year-old children. *Evolution and Human Behavio*r, 31(5), 354–364. http://doi.org/10.1016/j.evolhumbehav. 2010.04.004

Klein, S.K. & Rapin, I. (1993). Intermittent hearing loss and language development. In Bishop, D., & Mogford-Bevan, K. (Eds.), *Language development in exceptional circumstances* (pp. 96-109). Hove: Psychology Press.

Koelewijn, T., Shinn-Cunningham, B. G., Zekveld, A. A., & Kramer, S. E. (2014). The pupil response is sensitive to divided attention during speech processing. Hearing research, 312, 114-120.

Koenig, L. L., Fuchs, S., & Lucero, J. C. (2011). Effects of consonant manner and vowel height on intraoral pressure and articulatory contact at voicing offset and onset for voiceless obstruents. *The Journal of the Acoustical Society of America*, 129(5), 3233–3244. http://doi.org/10.1121/1.3561658, 2011

Kolb, B. and Whishaw, I.Q. (2006). *An introduction to brain and behaviour* (2nd Ed.). New York: Worth Publishers.

Kolinsky, R., Cuvelier, H., Goetry, V., Peretz, I., & Jose, M. (2009). Musical training facilitates lexical stress processing. *Music Perception*, *26*(3), 235–246.

Kolinsky, R., Lidji, P., Peretz, I., Besson, M., & Morais, J. (2009). Processing interactions between phonology and melody: Vowels sing but consonants speak. *Cognition*, 112(1), 1-20.

Korenman, L. M., & Peynircioglu, Z. F. (2007). Individual Differences in Learning and Remembering Music: Auditory versus Visual Presentation. *Journal of Research in Music Education*, *55*(1), 48–64. http://doi.org/10.1177/002242940705500105

Kramer, E., Linder, R., & Schönweiler, R. (2013). A Study of subharmonics in connected speech material. *Journal of Voice*, 27(1), 29-38.

Kraus, N., & Chandrasekaran, B. (2010). Music training for the development of auditory skills. *Nature Reviews. Neuroscience*, *11*(8), 599–605. http://doi.org/10.1038/nrn2882

Kraus, N., Slater, J., Thompson, E. C., Hornickel, J., Strait, D. L., Nicol, T., & White-Schwoch, T. (2014). Music enrichment programs improve the neural encoding of speech in at-risk children. *The Journal of Neuroscience*, *34*(36), 11913–8. http://doi.org/10.1523/JNEUROSCI.1881-14.2014

Kraus, N., Strait, D. L., & Parbery-Clark, A. (2012). Cognitive factors shape brain networks for auditory skills: spotlight on auditory working memory. *Annals of the New York Academy of Sciences*, 1252, 100–7. http://doi.org/10.1111/j.1749-6632.2012.06463.x

Krause, V., Schnitzler, A., & Pollok, B. (2010). Functional network interactions during sensorimotor synchronization in musicians and non-musicians. *Neuroimage*, 52(1), 245-251.

Krivokapić, J. (2012). Prosodic planning in speech production. In Fuchs, S., Weihrich, M., Pape, D., Perrier, P (Eds.), *Speech planning and dynamics* (pp. 157-190). Peter Lang.

Krumhansl, C. L. (2000). Rhythm and pitch in music cognition. *Psychological Bulletin*, 126(1), 159.

Kuhlman, K., & Schweinhart, L. J. (1999). Timing in Child Development. *High/Scope Educational Research Foundation*, 1–10.

Kumin, L. (2006). Speech intelligibility and childhood verbal apraxia in children with Down syndrome. *Down Syndrome Research and Practice*, 10(1), 10-22.

Kumin, L., & Adams, J. (2000). Developmental apraxia of speech and intelligibility in children with Down syndrome. *Down Syndrome Quarterly*, 5(3), 1-7.

Kumin, L., Councill, C., & Goodman, M. (1994). A longitudinal study of the emergence of phonemes in children with Down syndrome. *Journal of Communication Disorders*, 27(4), 293-303.

Ladefoged, P. (2005). *A course in phonetics* (4th ed.). Orlando, Fl: Harcourt.

Laibsirinon, S., Jarusurin, N., Kokoi, C., & Manakiatichai, T. (2012). Pulmonary Function and Chest Expansion in Thai Boys with Down Syndrome, *Thammasat Medical Journal-***ธรรมศาสตร์ เวช สาร***,* 12(2), 269-275

Lamont, A. (2009). Music in the school years. In Hallam, S., Cross, I. & Thaut, M. (Eds.), The Oxford handbook of music psychology (pp. 235-243). Oxford: Oxford University Press.

Lanfranchi, S., & Carretti, B. (2012). The increase in colored progressive matrices test performance in individuals with down syndrome: a qualitative and quantitative review. *Journal of Cognitive Education and Psychology*, 11(2), 143-158.

Lanfranchi, S., Baddeley, A., Gathercole, S., & Vianello, R. (2012). Working memory in Down syndrome: is there a dual task deficit?. *Journal of Intellectual Disability Research*, 56(2), 157-166.

Lanfranchi, S., Carretti, B., Spanò, G., & Cornoldi, C. (2009). A specific deficit in visuospatial simultaneous working memory in Down syndrome. *Journal of Intellectual Disability Research*, 53(5), 474-483.

Lanfranchi, S., Jerman, O., Dal Pont, E., Alberti, a, & Vianello, R. (2010). Executive function in adolescents with Down Syndrome. *Journal of Intellectual Disability Research : JIDR*, *54*(4), 308–19. http://doi.org/10.1111/j.1365-2788.2010.01262.x

Latash, M. L. (2007). Learning motor synergies by persons with Down syndrome. *Journal of Intellectual Disability Research*, 51(12), 962–971. http://doi.org/10.1111/j.1365-2788.2007.01008.x

Latash, M. L., Kang, N., & Patterson, D. (2002). Finger coordination in persons with Down syndrome: atypical patterns of coordination and the effects of practice. *Experimental Brain Research*, 146(3), 345-355.

Latash, M., Wood, L., & Ulrich, D. (2008). *What is currently known about hypotonia, motor skill development, and physical activity in Down syndrome*? Down Syndrome Education Online. doi: 10.3104/reviews.2074 https://www.down-syndrome.org/reviews/2074/?page=1

Laver, J. (1980). The phonetic description of voice quality. *Cambridge Studies in Linguistics London*, 31, 1-186.

Laws, G. (1998). The use of nonword repetition as a test of phonological memory in children with Down syndrome. *Journal of Child Psychology and Psychiatry*, 39(8), 1119-1130.

Laws, G., & Hall, A. (2014). Early hearing loss and language abilities in children with Down syndrome. *International Journal of Language & Communication Disorders*, 49(3), 333-342.

Lecanuet, J. P. (1996). The fetal stage. In Deliege, L; Sloboda, J.(Eds.), *Musical Beginnings*. Origins and Development of Musical Competence. Oxford, 3-36.

Lee, E. K., & Son, Y. I. (2005). Muscle tension dysphonia in children: Voice characteristics and outcome of voice therapy. *International Journal of Pediatric Otorhinolaryngology*, *69*(7), 911–917. http://doi.org/10.1016/j.ijporl.2005.01.030

Lee, M. T., Thorpe, J., & Verhoeven, J. (2009). Intonation and phonation in young adults with Down syndrome. *Journal of Voice : Official Journal of the Voice Foundation*, *23*(1), 82–7. http://doi.org/10.1016/j.jvoice.2007.04.006

Lee, Y., Lu, M., & Ko, H. (2007). Effects of skill training on working memory capacity. *Learning and Instruction*, *17*(3), 336–344. http://doi.org/10.1016/j.learninstruc.2007.02.010

Lense, M. D., & Dykens, E. M. (2011). Musical interests and abilities in individuals with developmental disabilities. *International review of research in developmental disabilities*, 41, 265-312.

Levitin, D. J. (1994). Absolute memory for musical pitch: evidence from the production of learned melodies. *Perception & Psychophysics*, 56(4), 414–423. http://doi.org/10.3758/BF03206733

Levitin, D. J., & Cook, P. R. (1996). Memory for musical tempo: Additional evidence that auditory memory is absolute. *Perception & Psychophysics*, 58(6), 927-935.

Lidji, P., Jolicœur, P., Kolinsky, R., Moreau, P., Connolly, J. F., & Peretz, I. (2010). Early integration of vowel and pitch processing: A mismatch negativity study. *Clinical neurophysiology*, 121(4), 533-541.

Lim, H. A. (2010). Effect of " Developmental Speech and Language Training Through Music " on Speech Production. *Journal of Music Therapy*, *XLVII*(1), 2–26.

Lin, K. W. E., Anderson, H., Hamzeen, M. H. M., & Lui, S. (2014). Implementation and Evaluation of Real-Time Interactive User Interface Design in Self-learning Singing Pitch Training Apps. Ann Arbor, MI: Michigan Publishing, University of Michigan Library.

Lin, E., Jiang, J., Noon, S. D., & Hanson, D. G. (2000). Effects of head extension and tongue protrusion on voice perturbation measures. *Journal of Voice*, 14(1), 8-16.

Liu, F., Patel, A. D., Fourcin, A., & Stewart, L. (2010). Intonation processing in congenital amusia: discrimination, identification and imitation. *Brain: A Journal of Neurology*, *133*(Pt 6), 1682–93. http://doi.org/10.1093/brain/awq089

Lucero, J. C. (2005). Dynamics of the Vocal Fold Oscillation. *Tendencias Em Mathematica Aplicada E Computacional*, *6*(1), 11–20. http://doi.org/10.1121/1.406922

Lucero, J. C., Lourenço, K. G., Hermant, N., Van Hirtum, A., & Pelorson, X. (2012). Effect of source–tract acoustical coupling on the oscillation onset of the vocal folds. *The Journal of the Acoustical Society of America*, *132*(1), 403. http://doi.org/10.1121/1.4728170

Ludlow, C. L., & Loucks, T. (2003). Stuttering: A dynamic motor control disorder. *Journal of Fluency Disorders*, 28(4), 273–295. http://doi.org/10.1016/j.jfludis.2003.07.001

Määttä, T., Määttä, J., Tervo-Määttä, T., Taanila, A., Kaski, M., & Iivanainen, M. (2011). Healthcare and guidelines: a population-based survey of recorded medical problems and health surveillance for people with Down syndrome. Journal of Intellectual and Developmental Disability, 36(2), 118-126.

MacDonald, R. A., O'Donnell, P. J., & Davies, J. (1999). An empirical investigation into the effects of structured music workshops for individuals with intellectual disabilities. *Journal of Applied Research in Intellectual Disabilities*, 12(3), 225-240.

Mackenzie Beck, J. (2010). Organic Variation of the Vocal Apparatus. In Hardcastle, W.J., Laver, J. and Gibbon, F.E. (Eds), *The handbook of phonetic sciences* (2nd Ed., pp. 153-201). Chichester: Blackwell Publishing Ltd.

Magne, C., Schön, D., & Besson, M. (2006). Musician children detect pitch violations in both music and language better than nonmusician children: behavioral and electrophysiological approaches. *Journal of Cognitive Neuroscience*, 18(2), 199–211. http://doi.org/10.1162/089892906775783660

Mahoney, G., & Robenalt, K. (1986). A comparison of conversational patterns between mothers and their Down syndrome and normal infants. *Journal of Early Intervention*, *10*(2), 172-180.

Malak, R., Kotwicka, M., Krawczyk-Wasielewska, A., Mojs, E., & Samborski, W. (2013). Motor skills, cognitive development and balance functions of children with Down syndrome. *Annals of Agricultural and Environmental Medicine : AAEM*, *20*(4), 803–6. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/24364457

Mang, E. (2001). A cross-language comparison of preschool children's vocal fundamental frequency in speech and song production. *Research Studies in Music Education*, 16(1), 4–14. http://doi.org/10.1177/1321103X010160010201

Mang, E. (2006). The effects of age, gender and language on children's singing competency. *British Journal of Music Education*, 23(02), 161. http://doi.org/10.1017/S0265051706006905

Marcell, M. (1995). Relationships between hearing and auditory cognition in Down syndrome youth. *Down Syndrome Research and Practice*, *3*(3), 75–91. http://doi.org/10.3104/reports.54

Marcell, M. M. & Weeks, S. L. (1988). Short-term memory difficulties and Down's syndrome. *Journal of Mental Deficiency Research*, 32, 153-162

Marcell, M.M., Cohen, P.S., Weathers, P., Wiseman, P., Croen, P. and Sewell, D.H. (1990). *Hearing abilities of Down Syndrome*. Paper presented at the Biennial Conference on Human development, Richmond, VA.. http://doi.org/10.1017/CBO9781107415324.004

Marie, C., Magne, C., & Besson, M. (2011). Musicians and the metric structure of words. *Journal of Cognitive Neuroscience*, *23*(2), 294–305. http://doi.org/10.1162/jocn.2010.21413

Maris, M., Wojciechowski, M., Van de Heyning, P., & Boudewyns, A. (2014). A cross-sectional analysis of otitis media with effusion in children with Down syndrome. *European Journal of Pediatrics*, *173*(10), 1319–1325. http://doi.org/10.1007/s00431-014-2323-5

Markides, A. (1983). *The speech of hearing-impaired children*. Manchester: Manchester University Press.

Marozeau, J., Innes-Brown, H., & Blamey, P. J. (2013). The acoustic and perceptual cues affecting melody segregation for listeners with a cochlear implant. *Frontiers in Psychology*, *4*(NOV), 1–11. http://doi.org/10.3389/fpsyg.2013.00790

Maryn, Y., Corthals, P., Van Cauwenberge, P., Roy, N., & De Bodt, M. (2010). Toward improved ecological validity in the acoustic measurement of overall voice quality: combining continuous speech and sustained vowels. *Journal of voice*, 24(5), 540-555.

Mason-Apps, E., Stojanovik, V., & Houston-Price, C. (2011). Early word segmentation in typically developing infants and infants with Down syndrome: a preliminary study. In *Proceedings of the 17th International Congress of Phonetic Sciences* (pp. 1334-1337).

Massaro, D. W., & Light, J. (2004). Using visible speech to train perception and production of speech for individuals with hearing loss. *Journal of Speech, Language and Hearing Research (JSLHR)*, 47(April), 304–320. http://doi.org/10.1044/1092-4388(2004/025)

Mathieson, L. (2001). *Mathieson's the voice and its disorders* (6th Edition). London and Philadelphia: Whurr Publishers.

Matsubara, M., Terasawa, H., & Hiraga, R. (2014). The effect of musical experience on rhythm perception for hearing-impaired undergraduates. In *Systems, Man and Cybernetics* (SMC), 2014 IEEE International Conference on (pp. 1666-1669). IEEE.

Matsuyama, K. (2005). Correlation between musical responsiveness and developmental age among early age children as assessed by the Non-Verbal Measurement of the Musical Responsiveness of Children. *Medical Science Monitor*, 11(10), CR485-CR492.

Mattys, S. L., & Wiget, L. (2011). Effects of cognitive load on speech recognition. *Journal of Memory and Language*, 65(2), 145–160. http://doi.org/10.1016/j.jml.2011.04.004

Mertens, P. (2004). The prosogram: Semi-automatic transcription of prosody based on a tonal perception model. In Speech Prosody 2004, International Conference.

Miendlarzewska, E., & Trost, J. W. (2013). How musical training affects cognitive development: rhythm, reward and other modulating variables. *Frontiers in Neuroscience*, 7(279).

Mishra, J. (2007). Correlating musical memorization styles and perceptual learning modalities. *Visions of Research in Music Education*, *9*(1), 1–19. Retrieved from http://users.rider.edu/~vrme/v9n1/vision/Mishra Final.pdf

Mithen, S. (2006). *The Singing Neanderthals: The origins of music, language, mind and body.* London: Orion Publishing Company.

Monsen, R. B. (1979). Acoustic qualities of phonation in young hearing-impaired children. *Journal of Speech, Language, and Hearing Research*, 22(2), 270-288.

Montague, J.C., & Hollien, H. (1973). Perceived voice quality disorders in Down's Syndrome children. *Journal of Communication Disorders*, 6, 6-87.

Moran, M. J. (1986). Identification of Down's syndrome adults from prolonged vowel samples. *Journal of Communication Disorders*, 19(5), 387-394.

Moran, M. J., and Gilbert, H. R. (1978). Speaking fundamental frequency characteristics of institutionalized adults with Down's syndrome. American Journal of Mental Deficiency. 83~248-252

Moreno, S., & Bidelman, G. M. (2014). Examining neural plasticity and cognitive benefit through the unique lens of musical training. *Hearing Research*, 308, 84–97. http://doi.org/10.1016/j.heares.2013.09.012

Morris, J. K., & Springett, A. (2011). *The national Down syndrome cytogenetic register for England and Wales: 2010 annual report*. London, United Kingdom: Queen Mary University of London, Barts and The London School of Medicine and Dentistry.

Moura, C. P., Cunha, L. M., Vilarinho, H., Cunha, M. J., Freitas, D., Palha, M., Peuschel, S.M. & Pais-Clemente, M. (2008). Voice parameters in children with Down syndrome. *Journal of Voice*, 22(1), 34-42.

Murphy, A. T , & Simons, R. F. (1959). Music therapy for the speech-handicapped. Elementary School Journal, 59, 39-45.

Næss, K. A. B., Lyster, S. A. H., Hulme, C., & Melby-Lervåg, M. (2011). Language and verbal short-term memory skills in children with Down syndrome: A meta-analytic review. *Research in Developmental Disabilities*, 32(6), 2225-2234.

Newton, 1999. Speech Processes repetition. In 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*. John Wiley & Sons.

Nooteboom, S. (1997). The prosody of speech: melody and rhythm. In Hardcastle, W.J. and Laver, J. (Eds), *The handbook of phonetic sciences* (pp. 640-673). Oxford: Blackwell.

Northern, J. L., & Downs, M. P. (2002). *Hearing in children*. Philadelphia: Lippincott Williams & Wilkins. Retrieved from:https://books.google.co.uk/books?hl=en&lr=&id=1CUFXsBbBgoC&oi=fnd&pg=PR5&dq=Northern,+J.+L.,+%26+Downs,+M.+P.+(2002).+Hearing+in+children.&ots=HVJ-7glJvf&sig=xoxJfx3jwvuhE8G8-dFo5XFUc9Y#v=onepage&q=Northern

%2C%20J.%20L.%2C%20%26%20Downs%2C%20M.%20P.%20(2002).%20Hearing%20in%20children.&f=false

Novak A. The voice of children with Down's syndrome. Folia Phoniatrica (Basel). 1972; 24:182–194.

Núñez, B. F., Suárez, N. C., Muñoz, P. C., Baragaño, R. L., Alvarez, Z. M., & Martínez, F. A. (1999). Spectrographic study of voice disorders: subharmonics. *Acta otorrinolaringologica espanola*, 51(1), 52-56. Abstract retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10799933

Ockelford, a., Welch, G., Jewell-Gore, L., Cheng, E., Vogiatzoglou, A., & Himonides, E. (2011). Phase 2: Gauging the music development of children with complex needs. *European Journal of Special Needs Education*, *26*(2), 177–199. http://doi.org/10.1080/08856257.2011.563606

Office for Standards in Education, Children's Services and Skills [Ofsted](2012). Music in schools: wider still, and wider. Manchester: Ofsted. Accessed 14th October 2016: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/413347/Music_in_schools_wider_still__and_wider.pdf

Office for Standards in Education, Children's Services and Skills [Ofsted](2013). Music in schools: what hubs must do. Manchester: Ofsted. Accessed 14th October 2016: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/413107/Music_in_schools_what_hubs_must_do.pdf

Ogawa, M., Hosokawa, K., Yoshida, M., Yoshii, T., Shiromoto, O., & Inohara, H. (2014). Immediate effectiveness of humming on the supraglottic compression in subjects with muscle tension dysphonia. *Folia Phoniatrica et Logopaedica*, *65*(3), 123–128. http://doi.org/10.1159/000353539

Olakunbi, D., Bamiou, D. E., Stewart, L., & Luxon, L. M. (2010). Evaluation of musical skills in children with a diagnosis of an auditory processing disorder. *International Journal of Pediatric Otorhinolaryngology*, 74(6), 633-636.

Olander, L., Smith, A., & Zelaznik, H. N. (2010). Evidence that a motor timing deficit is a factor in the development of stuttering. *Journal of Speech, Language, and Hearing Research*, 53(4), 876-886.

Osberger, M. J., & McGarr, N. S. (1982). Speech production characteristics of the hearing impaired. *Speech and language: Advances in basic research and practice*, 227-267.

Overy, K. (2000). Dyslexia, temporal processing and music: The potential of music as an early learning aid for dyslexic children. *Psychology of Music*, 28(2), 218-229.

Oxenham, A. J. (2008). Pitch perception and auditory stream segregation: implications for hearing loss and cochlear implants. *Trends in Amplification*, 12(4), 316-331.

Oxenham, A.J. (2013). Perception of musical tones. In Deutsch, D. (Ed.), Psychology of music (3rd Ed., pp. 1-33). London: Elsevier

Özdemir, E., Norton, A., & Schlaug, G. (2006). Shared and distinct neural correlates of singing and speaking. *NeuroImage*, *33*(2), 628–635. http://doi.org/10.1016/j.neuroimage.2006.07.013

Palisano, R. J., Walter, S. D., Russell, D. J., Rosenbaum, P. L., Gémus, M., Galuppi, B. E., & Cunningham, L. (2001). Gross motor function of children with Down syndrome: creation of motor growth curves. *Archives of Physical Medicine and Rehabilitation*, 82(4), 494-500.

Palmer, C. (2013). Music performance: Movement and Coordination. In Deutsch, D. (Ed.), *Psychology of music* (3rd Ed., pp. 405-422). London: Elsevier. .

Papousek, H. (1996). Musicality in infancy research: biological and cultural origins of early musicality. In Deliège, I., & Sloboda, J. A. (Eds.). *Musical beginnings: Origins and development of musical competence (pp. 37-51)*. Oxford: Oxford University Press.

Parbery-Clark, A., Anderson, S., & Kraus, N. (2013). Musicians change their tune: how hearing loss alters the neural code. *Hearing Research*, *302*, 121–31. http://doi.org/10.1016/j.heares.2013.03.009

Parbery-Clark, A., Anderson, S., Hittner, E., & Kraus, N. (2012). Musical experience strengthens the neural representation of sounds important for communication in middle-aged adults. *Frontiers in Aging Neuroscience*, 4, 30.

Parbery-Clark, A., Skoe, E., Lam, C., & Kraus, N. (2009). Musician enhancement for speech-in-noise. *Ear and Hearing*, *30*(6), 653–61. http://doi.org/10.1097/AUD.0b013e3181b412e9

Parncutt, R. (2016). Prenatal development. In McPherson, G. E. (Ed.), *The child as musician: A handbook of musical development* (2nd Edition, pp. 1-31). Oxford: Oxford University Press.

Patel, A. D. (2008). *Music, language, and the brain*. Oxford university Press: Oxford.

Patel, A. D. (2011). Why would Musical Training Benefit the Neural Encoding of Speech? The OPERA Hypothesis. *Frontiers in Psychology*, 2(June), 142. http://doi.org/10.3389/fpsyg.2011.00142

Patel, A. D. (2014). Can nonlinguistic musical training change the way the brain processes speech? The expanded OPERA hypothesis. *Hearing Research*, *308*, 98–108. http://doi.org/10.1016/j.heares.2013.08.011

Patel, S., Scherer, K. R., Björkner, E., & Sundberg, J. (2011). Mapping emotions into acoustic space: The role of voice production. *Biological Psychology*, 87(1), 93-98.

Patterson, T., Rapsey, C. M., & Glue, P. (2013). Systematic review of cognitive development across childhood in Down syndrome: Implications for treatment interventions. *Journal of Intellectual Disability Research*, *57*(4), 306–318. http://doi.org/10.1111/j.1365-2788.2012.01536.x

Pekkonen, E., Osipova, D., Sauna-Aho, O., & Arvio, M. (2007). Delayed auditory processing underlying stimulus detection in Down syndrome. *Neuroimage*, 35(4), 1547-1550.

Pentz Jr, A. L., & Gilbert, H. R. (1983). Relation of selected acoustical parameters and perceptual ratings to voice quality of Down syndrome children. *American Journal of Mental Deficiency*, 88(2), 203-210.

Peppé, S. J. (2009). Why is prosody in speech-language pathology so difficult?. *International Journal of Speech-Language Pathology*, 11(4), 258-271.

Peretz, I., & Coltheart, M. (2003). Modularity of music processing. *Nature Neuroscience*, *6*(7), 688–691. http://doi.org/10.1038/nn1083

Peretz, I., & Zatorre, R. J. (2003). *The cognitive neuroscience of music*. *Cognitive Psychology*. http://doi.org/10.1093/acprof:oso/9780198525202.001.0001

Peretz, I., Gagnon, L., Hébert, S., & Macoir, J. (2004). Singing in the brain : Insights from cognitive neuropsychology. *Music Perception*, *21*(3), 373–390.

Perry, M. M. R. (2003). Relating improvisational music therapy with severely and multiply disabled children to communication development. *Journal of Music Therapy*, 40(3), 227-246.

Peter, B., & Stoel-Gammon, C. (2008). Central timing deficits in subtypes of primary speech disorders. *Clinical Linguistics & Phonetics*, *22*(3), 171–98. http://doi.org/10.1080/02699200701799825

Peters, J. S. (2000). *Music therapy: an introduction*. Springfield, Illinois: Charles C Thomas Publisher. Retrieved from https://books.google.co.uk/books?id=olOcCQAAQBAJ&printsec=frontcover&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false

Pettinato, M. (2009). The effects of early auditory deprivation – insights from children with cochlear implants. *Down Syndrome Research and Practice*, *12*(3), 176–178.

Pettinato, M., & Verhoeven, J. (2009). Production and perception of word stress in children and adolescents with Down syndrome. *Down Syndrome Research and Practice*, 1–13. http://doi.org/10.3104/reports.2036

Pfordresher, P. Q. (2006). Coordination of perception and action in music performance. *Advances in Cognitive Psychology*, 2(2), 183–198. http://doi.org/10.2478/v10053-008-0054-8

Pfordresher, P. Q., & Brown, S. (2009). Enhanced production and perception of musical pitch in tone language speakers. *Attention, Perception, & Psychophysics*, 71(6), 1385-1398.

Pfordresher, P. Q., & Dalla Bella, S. (2011). DAF and movement. *Journal of Chemical Information and Modeling*, 53, 160. http://doi.org/10.1017/CBO9781107415324.004

Pfordresher, P., & Mantell, J. (2009). Singing as a form of vocal imitation — mechanisms and deficits. *Proceedings of the 7th Triennial Conference of European Society for the Cognitive Sciences of Music (ESCOM 2009)* Jyväskylä, Finland Jukka

Philips-Silver, J. and L.J. Trainor (2007), Hearing What the Body Feels: Auditory Encoding of Rhythmic Movement, *Cognition*, 105 (3): 533–46.

Phillips-Silver, J. (2009). On the Meaning of Movement in Music, Development and the Brain. *Contemporary Music Review*, 28(3), 293–314. http://doi.org/10.1080/07494460903404394

Phillips-Silver, J., & Trainor, L. J. (2005). Feeling the beat: movement influences infant rhythm perception. *Science*, 308(5727), 1430-1430.

Phillips-Silver, J., & Trainor, L.J. (2008). Vestibular influence on auditory metrical interpretation. *Brain and Cognition*, Vol. 67, pp. 94-102

Phillips-Silver, J., Aktipis, C. A., & Bryant, G. A. (2010). The ecology of entrainment: Foundations of coordinated rhythmic movement. *Music Perception: An Interdisciplinary Journa*l, 28(1), 3-14.

Phillips, K. H. (1996). *Teaching kids to sing*. Thomson Learning: Belmont, USA

Picard, B. M. (2009). Music and down's syndrome. Master's thesis, Royal Welsh College Of Music and Drama. Retrieved from http://www.melodymusic.org.uk/publications/4574047824

Pinho, S. M. R., Pontes, P. a L., Gadelha, M. E. C., & Biasi, N. (1999). Vestibular vocal fold behavior during phonation in unilateral vocal fold paralysis. *Journal of Voice*, *13*(1), 36–42. http://doi.org/10.1016/S0892-1997(99)80059-9

Pon-Barry, H. (2008). Prosodic manifestations of confidence and uncertainty in spoken language. *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, 74–77.

Pon-Barry, H., & Shieber, S. M. (2011). Recognizing uncertainty in speech. *EURASIP Journal on Advances in Signal Processing*, 2011(1), 1-11.

Porter, H. L., Grantham, D. W., Ashmead, D. H., & Tharpe, A. M. (2014). Binaural Masking Release in Children With Down Syndrome. *Ear and Hearing*, 35(4), e134-e142.

Porter, M., & Coltheart, M. (2006). Global and local processing in William syndrome, Autism, and Down syndrome: Perception, Attention, and Construction. *Developmental Neuropsychology*, *30*(3), 771–789. http://doi.org/10.1207/s15326942dn3003

Pryce, M. (1994). The voice of people with Down syndrome: An EMG biofeedback study. *Down Syndrome Research and Practice*, 2(3), 106-111.

Purser, H. R., & Jarrold, C. (2005). Impaired verbal short-term memory in Down syndrome reflects a capacity limitation rather than atypically rapid forgetting. *Journal of Experimental Child Psychology*, *91*(1), 1–23. http://doi.org/10.1016/j.jecp.2005.01.002

Purser, H. R., & Jarrold, C. (2005). Impaired verbal short-term memory in Down syndrome reflects a capacity limitation rather than atypically rapid forgetting. *Journal of Experimental Child Psychology*, *91*(1), 1–23. http://doi.org/10.1016/j.jecp.2005.01.002

Quatieri, T., Williamson, J., Smalt, C., Patel, T., Perricone, J., Mehta, D. et al., (2015). Vocal biomarkers to discriminate cognitive load in a working memory task. Interspeech, 2684–2688.

Racette, A., & Peretz, I. (2007). Learning lyrics: to sing or not to sing? *Memory & Cognition*, *35*(2), 242–53. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/17645165

Rajendran, V., Roy, F. G., & Jeevanantham, D. (2012). Postural control, motor skills, and health-related quality of life in children with hearing impairment: a systematic review. *European Archives of Oto-Rhino-Laryngology*, 269(4), 1063-1071.

Ramig, L. O. (1992). The role of phonation in speech intelligibility: A review and preliminary data from patients with Parkinson's disease. *Intelligibility in speech disorders: Theory, measurement and management*, 119-155.

Ratz, C. (2013). Do students with Down syndrome have a specific learning profile for reading? *Research in Developmental Disabilities*, 34(12), 4504–4514. http://doi.org/10.1016/j.ridd.2013.09.031

Redford, M. A. (2015). The handbook of speech production. Chichester: John Wiley & Sons.

Reifinger Jr, J. L. (2006). Skill development in rhythm perception and performance: A review of literature. UPDATE: *Applications of Research in Music Education,* 25(1), 15-28.

Repp, B. H. (2005). Sensorimotor synchronization: A review of the tapping literature. *Psychonomic Bulletin & Review*, 12(6), 969-992.

Repp, B. H., & Doggett, R. (2007). Tapping to a very slow beat: a comparison of musicians and nonmusicians. *Music Perception: An Interdisciplinary Journa*l, 24(4), 367-376.

Repp, B. H., & Su, Y.-H. (2013). Sensorimotor synchronization: a review of recent research (2006-2012). *Psychonomic Bulletin & Review*, 20(3), 403–52. http://doi.org/10.3758/s13423-012-0371-2

Revoile, S. G. (1999). Hearing loss and the audibility of phoneme cues. In Pickett, J.M. (Ed), *The Acoustics of Speech Communication* (pp. 289-323). Boston: Allyn and Bacon.

Riecker, A., Ackermann, H., Wildgruber, D., Dogil, G., & Grodd, W. (2000). Opposite hemispheric lateralization effects during speaking and singing at motor cortex, insula and cerebellum. *Neuroreport*, 11(9), 1997–2000. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10884059

Rigoldi, C., Galli, M., & Albertini, G. (2011). Gait development during lifespan in subjects with Down syndrome. *Research in Developmental Disabilities*, 32(1), 158–163. http://doi.org/10.1016/j.ridd.2010.09.009

Rigoldi, C., Galli, M., Mainardi, L., Crivellini, M., & Albertini, G. (2011). Postural control in children, teenagers and adults with Down syndrome. *Research in Developmental Disabilities*, *32*(1), 170–175. http://doi.org/10.1016/j.ridd.2010.09.007

Ringenbach, S. D., Allen, H., Chung, S., & Jung, M. L. (2006). Specific instructions are important for continuous bimanual drumming in adults with Down syndrome. *Downs Syndrome Research and Practice*, 11(1), 29–36. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/17048807

Rinta, T., & Welch, G. F. (2008). Should singing activities be included in speech and voice therapy for prepubertal children? *Journal of Voice : Official Journal of the Voice Foundation*, *22*(1), 100–12. http://doi.org/10.1016/j.jvoice.2006.08.002

Robertson, S. D., Van Gemmert, A. W. A, & Maraj, B. K. V. (2002). Auditory information is beneficial for adults with Down syndrome in a continuous bimanual task. *Acta Psychologica*, 110(2-3), 213–29. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12102106

Roch, M., & Jarrold, C. (2008). A comparison between word and nonword reading in Down syndrome: The role of phonological awareness. *Journal of Communication Disorders*, 41(4), 305-318.

Rochette, F., Moussard, A., & Bigand, E. (2014). Music lessons improve auditory perceptual and cognitive performance in deaf children. *Frontiers in Human Neuroscience*, 8.

Rodger, R. (2009). *Voice quality of children and young people with Down's Syndrome and its impact on listener judgement*. Unpublished PhD dissertation, Queen Margaret University, Edinburgh, Scotland. Retrieved from: http://etheses.qmu.ac.uk/352/

Rondal, J. A. (2009). Spoken language in persons with Down syndrome: a life-span perspective. *International Journal of Early Childhood Special Education*, 1(2), 138-163.

Rondal, J., & Buckley, S. (2003). *Speech and language intervention in Down syndrome.* London: Whurr Publishers Ltd.

Rosin M.M., Swift E., Bless D., Vetter D.K.. Communication profiles of adolescents with Down syndrome. *Journal of Childhood Communication Disorders*. 1988; 12:49–64.

Roy, N. (2003). Functional dysphonia. *Current Opinion in Otolaryngology & Head and Neck Surgery*, 11(3), 144-148.

Rutkowski, J. (1997). The nature of children's singing voices: Characteristics and assessment. In: B.A. Roberts (Ed.), *The Phenomenon of Singing* (pp. 201-209). St. John's, NF: Memorial University Press

Sacks, B., & Buckley, S. (2003). Motor development for individuals with Down syndrome–An overview. *Down Syndrome Issues and Information*. Retrieved from https://www.down-syndrome.org/information/motor/overview/

Saffran, J. R., Loman, M. M., & Robertson, R. R. (2000). Infant memory for musical experiences. *Cognition*, 77(1), B15-B23.

Saito, S. (2001). The phonological loop and memory for rhythms: An individual differences approach. *Memory* (Hove, England), 9(4), 313–322. http://doi.org/10.1080/09658210143000164

Salgueirinho, C., Venâncio, J., Martín-Nogueras, A. M., & Ribeiro, F. (2015). Pulmonary function in young adults with Down syndrome: A cross-sectional study. *International Medical Review on Down Syndrome*.

Salihović, N., Junuzović-žunić, L., Ibrahimagić, A., & Beganović, L. (2009). *Characteristics of voice in stuttering children*. Acta Medica Saliniana, 38(2), 67-75.

Santurette, S., & Dau, T. (2007). Binaural pitch perception in normal-hearing and hearing-impaired listeners. *Hearing Research*, *223*(1-2), 29–47. http://doi.org/10.1016/j.heares.2006.09.013

Sataloff, R. (2014) Physical examination. In Rubin, J. S., Sataloff, R. T., & Korovin, G. S. (Eds.), *Diagnosis and treatment of voice disorders* (4th Ed., pp. 211 -222). San Diego, CA: Plural Publishing. Retrieved from https://books.google.co.uk/books?id=CZQ1BwAAQBAJ&printsec=frontcover&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false

Schaeffler, F., Beck, J. M., & Jannetts, S. (2015). Phonation stabilisation time as an indicator of voice disorder. *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS 2015)*.

Schellenberg, E. G., & Trehub, S. E. (1999). Culture-general and culture-specific factors in the discrimination of melodies. *Journal of Experimental Child Psychology*, 74(2), 107-127.

Schellenberg, G.E & Weiss, M.W. (2013). Music and cognitive abilities. In D. Deutsch (Ed.), Psychology of music (3rd Ed., pp 499-534). Elsevier

Schellenberg, G.E. (2016). Musical and nonmusical abilities. In G.E. McPherson (Ed.), *The child as musician: A handbook of musical development* (2nd Ed. pp.149-176). Oxford: Oxford Univerity Press

Schieve, L. a., Boulet, S. L., Boyle, C., Rasmussen, S. a., & Schendel, D. (2009). Health of Children 3 to 17 Years of Age With Down Syndrome in the 1997-2005 National Health Interview Survey. *Pediatrics*, *123*(2), e253–e260. http://doi.org/10.1542/peds.2008-1440

Schlaug, G., Marchina, S., & Norton, A. (2008). From singing to speaking: why singing may lead to recovery of expressive language function in patients with Broca's aphasia. *Music Perception: An Interdisciplinary Journa*l, 25(4), 315-323.

Schleuter, Stanley, L., & Schleuter, L. J. (1989). The relationship of rhythm response tasks and PMMA scores with music training, grade level and sex amongst K-3 children. *Bulletin of the Council for Reserach in Music Education*, (100), 1–13.

Schön, D., Boyer, M., Moreno, S., Besson, M., Peretz, I., & Kolinsky, R. (2008). Songs as an aid for language acquisition. *Cognition*, 106(2), 975–83. http://doi.org/10.1016/j.cognition.2007.03.005

Schulze, K., Zysset, S., Mueller, K., Friederici, A. D., & Koelsch, S. (2011). Neuroarchitecture of verbal and tonal working memory in nonmusicians and musicians. *Human Brain Mapping*, 32(5), 771-783.

Sebastian, S., Benedict, A. S., & Balraj, A. (2013). Laryngeal movements in stutterers. *Journal of Laryngology and Voice*, 3(1), 14.

Setter, J., & Stojanovik, V. (2011). Speech prosody in Down's and Williams syndrome: a comparison. In *Proceedings of the XVII International Congress of Phonetic Sciences* (pp. 1798-1801).

Seung, H., & Chapman, R. (2000). Digit span in individuals with down syndrome and in typically developing children: temporal aspects. *Journal of Speech, Language, and Hearing Research*, 43, 609–620.

Seybold, C. D. (1971). The value and use of music activities in the treatment of speech delayed children. *Journal of Music Therapy*, 8, 102-110

Shapiro, B. L. (2003). Down syndrome and associated congenital malformations (pp. 207-214). Springer Vienna.

Shipley, K., & McAfee, J. (2009). *Assessment in speech-language pathology: A resource manual.* (4th ed.). New York, USA: Delmar Cengage Learning.

Shriberg, L. D., Austin, D., Lewis, B. A., McSweeny, 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.

Shu, M., Jiang, J. J., & Willey, M. (2016). The Effect of Moving Window on Acoustic Analysis. *Journal of Voice*, *30*(1), 5–10. http://doi.org/10.1016/j.jvoice.2014.11.008

Shumway-Cook, a, & Woollacott, M. H. (1985). Dynamics of postural control in the child with Down syndrome. *Physical Therapy*, *65*, 1315–1322.

Sigman, M., Ruskin, E., Arbelle, S., Corona, R., Dissanayake, C., Espinosa, M., ... & Robinson, B. F. (1999). Continuity and change in the social competence of children with autism, Down syndrome, and developmental delays. Monographs of the society for research in child development, i-139

Silverman, W. (2007). Down Syndrome: cognitive phenotype. *Mental Retardation and Developmental Disabilities Research Reviews*, *13*, 228–236. http://doi.org/10.1002/mrdd

Simpson, K., & Keen, D. (2011). Music interventions for children with autism: narrative review of the literature. *Journal of Autism and Developmental Disorders*, *41*(11), 1507–14. http://doi.org/10.1007/s10803-010-1172-y

Sloboda, J. A. (1985). *The musical mind: The cognitive psychology of musi*c. Oxford: Oxford University Press.

Slonims, V., & McConachie, H. (2006). Analysis of mother-infant interaction in infants with Down syndrome and typically developing infants. *American Journal on Mental Retardation*, *111*(4), 273-289.

Slough, L. L., Creech, A., Saunders, J., & Welch, G. (2015). Musical Pride: Music education in plural communities. Retrieved January 10, 2016, from: http://leicestershiremusichub.org/uploads/musical-pride-music-education-in-plural-communities.pdf

Smith, E., & Jarrold, C. (2014). Demonstrating the effects of phonological similarity and frequency on item and order memory in Down syndrome using process dissociation. *Journal of Experimental Child Psychology*, *128*, 69–87. http://doi.org/10.1016/j.jecp.2014.07.002

Snyder, B. (2000). *Music and memory: An introduction*. London: The MIT Press.

Snyder, B. (2009). Memory for music. In Hallam, S., Cross, I., and Thaut, M. (Eds.), *The Oxford handbook of music psychology* (pp 171-183). Oxford: Oxford University Press.

Sonneta [computer software]. Version 1.0.1(90). Mint Leaf Software Inc. Toronto: Canada.

Sourtiji, H., Hosseini, S. M. S., Soleimani, F., & Hosseini, S. A. (2010). Relationship between motor and mental age in children with Down syndrome. *Iranian Rehabilitation Journal*, 8(11), 4-7.

Sowiński, J., & Dalla Bella, S. (2013). Poor synchronization to the beat may result from deficient auditory-motor mapping. *Neuropsychologia*, 51(10), 1952-1963.

Sparks, R W, & Holland, A. L. (1976). Method: Melodic intonation therapy for aphasia. Journal of Speech and Hearing Disorders, 41, 287-297

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.

Stager, S. (2011). The Role of the Supraglottic Area in Voice Production. *Otolaryngol S1*, 1–7. http://doi.org/10.4172/2161-119X.S1-001

Stahl, B., Henseler, I., Turner, R., Geyer, S., & Kotz, S. A. (2013). How to engage the right brain hemisphere in aphasics without even singing: evidence for two paths of speech recovery. *Frontiers in Human Neuroscience*, *7*(February), 35. http://doi.org/10.3389/fnhum.2013.00035

Stelmachowicz, P. G., Pittman, A. L., Hoover, B. M., & Lewis, D. E. (2004). Novel-word learning in children with normal hearing and hearing loss. *Ear and Hearing*, 25, 47–56. http://doi.org/10.1097/01.AUD.0000111258.98509.DE

Stephenson, J. (2006). Music Therapy and the Education of Students with Severe Disabilities. *Education and Training in Developmental Disabilities*, *41*(September), 290–299.

Stevens, K. (1997). Articulatory-acoustic-auditory relationships. In Hardcastle, W. J. and Laver, J. (Eds.), *The Handbook of Phonetic Science* (pp. 462-506). Oxford: Blackwell.

Stoel-Gammon, C. (2001). Down syndrome phonology: developmental patterns and intervention strategies. *Down's Syndrome, Research and Practice : The Journal of the Sarah Duffen Centre / University of Portsmouth*, *7*(3), 93–100. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11721538

Stojanovik, V. (2011). Prosodic deficits in children with Down syndrome. *Journal of Neurolinguistics*, *24*(2), 145–155. http://doi.org/10.1016/j.jneuroling.2010.01.004

Stratford, B., & Ching, E. Y. Y. (1983). Rhythm and time in the perception of Down's syndrome children. *Journal of Intellectual Disability Research*, 27(1), 23-38.

Stratford, B., & Ching, E. Y. Y. (1989). Responses to music and movement in the development of children with Down's syndrome. *Journal of Intellectual Disability Research*, 33(1), 13-24.

Strik, H., & Boves, L. L. B.-E. (1992). Control of fundamental frequency, intensity and voice quality in speech. *Journal of Phonetics*, *20*, 15–25.

Su, Y. H. (2014). Visual enhancement of auditory beat perception across auditory interference levels. *Brain and Cognition*, 90, 19–31. http://doi.org/10.1016/j.bandc.2014.05.003

Sun, X. (2002). Pitch determination and voice quality analysis using Subharmonic-to-Harmonic Ratio. *2002 IEEE International Conference on Acoustics, Speech, and Signal Processing*, *1*, I–333–I–336. http://doi.org/10.1109/ICASSP.2002.5743722

Swanwick, K. (2004). Mapping Music Education Research in the UK. *Psychology of Music*, 32(3), 239–290.

Sweller, J. (2011). Cognitive load theory. In Mestre, J.P. and Ross, B.H. (Eds.), *The psychology of learning and motivation* (vo.55, pp. 37-76). London: Academic Press

Tan, Y. T., McPherson, G. E., Peretz, I., Berkovic, S. F., & Wilson, S. J. (2014). The genetic basis of music ability. *Frontiers in Psychology,* 5(658), 10-3389.

Tao, C., & Jiang, J. J. (2008). Chaotic component obscured by strong periodicity in voice production system. Physical Review E, 77(6), 061922.

Teixeira, J. P., & Fernandes, P. O. (2015). Acoustic Analysis of Vocal Dysphonia. *Procedia Computer Science*, *64*, 466–473. http://doi.org/10.1016/j.procs.2015.08.544

Thaut, M. H. (2008). *Rhythm, music, and the brain: Scientific foundations and clinical applications*. Routledge.

Thaut, M. H., Peterson, D. A., McIntosh, G. C., & Hoemberg, V. (2014). Music mnemonics aid verbal memory and induce learning–related brain plasticity in multiple sclerosis. *Frontiers in Human Neuroscience*, 8.

Thaut, M. H., Peterson, D. A., Sena, K. M., & Mcintosh, G. C. (2008). Musical structure facilitates verbal learning in multiple sclerosis. *Music Perception: An Interdisciplinary Journal*, 25(4), 325-330.

Thompson, W.F. (2013). Intervals and scales. In D. Deutsch (Ed.), *Psychology of Music* (3rd Ed., pp 107-140). Elsevier.

Thorpe, L. A., & Trehub, S. E. (1989). Duration illusion and auditory grouping in infancy. *Developmental Psychology,* 25(1), 122.

Tierney, A. T., Krizman, J., & Kraus, N. (2015). Music training alters the course of adolescent auditory development. *Proceedings of the National Academy of Sciences*, 201505114. http://doi.org/10.1073/pnas.1505114112

Tierney, A., & Kraus, N. (2013). The ability to move to a beat is linked to the consistency of neural responses to sound. *The Journal of Neuroscience*, 33(38), 14981-14988.

Tierney, A., Dick, F., Deutsch, D., & Sereno, M. (2013). Speech versus song: Multiple pitch-sensitive areas revealed by a naturally occurring musical illusion. *Cerebral Cortex*, *23*(2), 249–254. http://doi.org/10.1093/cercor/bhs003

Titze, I. R. (1994). Workshop on Acoustic Voice Analysis: Summary statement. National Center for Voice and Speech (pp. 1–36).

Titze, I. R. (2008). Nonlinear source–filter coupling in phonation: Theory. *The Journal of the Acoustical Society of America*, *123*(5), 2733. http://doi.org/10.1121/1.2832337

Titze, I. R. (2014). Bi-stable vocal fold adduction: A mechanism of modal-falsetto register shifts and mixed registration. The Journal of the Acoustical Society of America, 135(4), 2091-2101

Trainor, L. J. (2005). Are there critical periods for musical development? *Developmental Psychobiology*, *46*(3), 262–78. http://doi.org/10.1002/dev.20059

Trainor, L. J. and Hannon, E.E. (2013). Musical Development. In Deutsch, D. (Ed.), *The psychology of music*, (3rd Ed., pp. 423-474). London: Elsevier.

Trainor, L. J., & Trehub, S. E. (1992). A comparison of infants' and adults' sensitivity to western musical structure. *Journal of Experimental Psychology: Human Perception and Performance*, 18(2), 394.

Trainor, L. J., Wu, L., & Tsang, C. D. (2004). Long-term memory for music: Infants remember tempo and timbre. *Developmental Scienc*e, 7(3), 289-296.

Traunmüller, H. (2005). Conversion between acoustic and auditory units of pitch and calculation of auditory intervals and distances. http://www2.ling.su.se/staff/hartmut/umrechnung.htm. Accessed 2015-2016.

Treharne, D. (1999) Test of Rhythm and Comprehension of Language. Via Sheffield University.

Trehub, S.E & Degé, F. (2016). Reflections on infants as musical connoisseurs. In McPherson, G. E. (Ed.), *The child as musician: A handbook of musical development* (2nd Edition, pp. 31-51). Oxford: Oxford University Press.

Tsang, C. D., Friendly, R. H., & Trainor, L. J. (2011). Singing development as a sensorimotor interaction problem. *Psychomusicology: Music, Mind and Brain*, 21(1-2), 31.

Tsao, R., & Kindelberger, C. (2009). Variability of cognitive development in children with Down syndrome: Relevance of good reasons for using the cluster procedure. *Research in Developmental Disabilities*, *30*, 426–432. http://doi.org/10.1016/j.ridd.2008.10.009

Tumber, A. K., Scheerer, N. E., & Jones, J. A. (2014). Attentional demands influence vocal compensations to pitch errors heard in auditory feedback. *PLoS ONE*, *9*(10). http://doi.org/10.1371/journal.pone.0109968

Van Borsel, J., & Vandermeulen, A. (2008). Cluttering in Down syndrome. *Folia Phoniatrica et Logopaedica : Official Organ of the International Association of Logopedics and Phoniatrics (IALP)*, *60*(6), 312–7. http://doi.org/10.1159/000170081

Van Houtte, E., Van Lierde, K., & Claeys, S. (2011). Pathophysiology and treatment of muscle tension dysphonia: A review of the current knowledge. *Journal of Voice*, 25(2), 202–207. http://doi.org/10.1016/j.jvoice.2009.10.009, 2011

Vatti, Santurette, Pontoppidan, & Dau, T. (2014). Perception of a sung vowel as a function of Frequency-Modulation rate and excursion in listeners with normal hearing and hearing impairment. *Journal of Speech, Language, and Hearing Research*, 57(October 2014), 1679–1691. http://doi.org/10.1044/2014

Venail, F., Gardiner, Q., & Mondain, M. (2004). ENT and speech disorders in children with Down's syndrome: an overview of pathophysiology, clinical features, treatments, and current management. *Clinical Pediatrics*, 43(9), 783-791.

Vicari S., Bellucci S., Carlesimo G. A. (2005). Visual and spatial long-term memory: differential pattern of impairments in Williams and Down syndromes. *Developmental Medicine and Child Neurology*, 47, 305–311

Vicari S., Carlesimo G. A. (2006). Short-term memory deficits are not uniform in Down and Williams syndromes. *Neuropsychology Review*, 16 87–94

Vicari, S. (2006). Motor development and neuropsychological patterns in persons with Down syndrome. *Behavior Genetics*, *36*(3), 355–364. http://doi.org/10.1007/s10519-006-9057-8

Vicari, S., Bellucci, S., & Carlesimo, G. A. (2000). Implicit and explicit memory: a functional dissociation in persons with Down syndrome. *Neuropsychologia*, 38(3), 240-251.

Vimercati, S. L., Galli, M., Rigoldi, C., Ancillao, A., & Albertini, G. (2013). Motor strategies and motor programs during an arm tapping task in adults with Down Syndrome. *Experimental Brain Research*, *225*(3), 333–338. http://doi.org/10.1007/s00221-012-3373-4

Virji-Babul, N., Kerns, K., Zhou, E., Kapur, A., & Shiffrar, M. (2006). Perceptual-motor deficits in children with Down syndrome: implications for intervention. *Down Syndrome Research and Practice*, 10(2), 74-82.

Walicka-Cupryś, K., Przygoda, Ł., Czenczek, E., Truszczyńska, A., Drzał-Grabiec, J., Zbigniew, T., & Tarnowski, A. (2014). Balance assessment in hearing-impaired children. *Research in developmental disabilities*, *35*(11), 2728-2734.

Wan, C., Ruber, T., Hohmann, A., & Schlaug, G. (2010). The therapeutic effects of singing in neurological disorders. *Music Perception*, *27*(4), 287–296.

Wan, C.E. & Schlaug, G. (2013). Brain plasticity induced by musical training. In Deutsch, D. (Ed.), *The psychology of music*, (3rd Ed., pp. 565-582). London: Elsevier.

Welch, G. F. (1985). A schema theory of how children learn to sing in tune. *Psychology of music*, *13*(1), 3-18.

Welch, G. F. (1994). The Assessment of Singing. *Psychology of Music*, 22(1), 3–19.

Welch, G. F. (1998). Early Childhood Musical Development. *Research Studies in Music Education*, 11(1), 27–41. http://doi.org/10.1177/1321103X9801100104

Welch, G. F. (2009). Evidence of the development of vocal pitch matching ability in children 1. *Japanese Journal of Music Education Research*, *21*, 1–13.

Welch, G. F. (2016). Singing and vocal development. In McPherson, G. E. (Ed.), *The child as musician: A handbook of musical development* (pp. 441-461). Oxford: Oxford University Press.

Welch, G. F., & Henley, J. (2014). Addressing the challenges of teaching music by generalist primary school teachers. Revista da ABEM, 22(32).

Welch, G. F., Himonides, E., Saunders, J., Papageorgi, I., & Sarazin, M. (2014). Singing and social inclusion. *Frontiers in Psychology*, *5*(July), 803. http://doi.org/10.3389/fpsyg.2014.00803

Welch, G. F., Howard, D. M., & Rush, C. (1989). Real- time visual feedback in the development of vocal pitch accuracy in singing. Psychology of Music, 17, 146-157.

Welch, G. F., Saunders, J., Edwards, S., Palmer, Z., Himonides, E., Knight, J. et al. (2015). Using singing to nurture children's hearing? A pilot study. *Cochlear Implants International*, 16(sup3), S63-S70.

Welch, G. F., Sergeant, D. C., & White, P. J. (1996). The singing competencies of five-year-old developing singers. *Bulletin of the Council for Research in Music Education*, 155-162.

Welch, G. F., Sergeant, D. C., & White, P. J. (1997). Age, sex, and vocal task as factors in singing" in tune" during the first years of schooling. *Bulletin of the Council for Research in Music Education*, 153-160.

Welch, G. F., Sergeant, D. C., & White, P. J. (1998). The role of linguistic dominance in the acquisition of song. *Research Studies in Music Education*, 10(1), 67-74.

Welch, G., Himonides, E., Saunders, J., Papageorgi, I., Rinta, T., Preti, C. et al., (2008). *Researching the first year of the National Singing Programme in England: An initial impact evaluation of children's singing behaviours and singer identity*. Institute of Education, University of London.

Welsh, T. N., Elliott, D., & Simon, D. A. (2003). Cerebral specialization and verbal-motor integration in adults with and without Down syndrome. *Brain and Language*, 84(2), 152-169.

Welsh, T.N. and Elliott, D. (2000). Preparation and control of goal-directed lib-movements in persons with Down Syndrome. In Weeks, D. J., Chua, R., & Elliott, D. (Eds), *Perceptual-motor behavior in Down syndrome* (pp. 49-70). Champaign, Il: Human Kinetics.

West, M. L. (1992). *Ancient greek music*. Clarendon Press.

White, E. J., Hutka, S. a, Williams, L. J., & Moreno, S. (2013). Learning, neural plasticity and sensitive periods: implications for language acquisition, music training and transfer across the lifespan. *Frontiers in Systems Neuroscience*, *7*(November), 90. http://doi.org/10.3389/fnsys.2013.00090

White, E. J., Hutka, S. A, Williams, L. J., & Moreno, S. (2013). Learning, neural plasticity and sensitive periods: implications for language acquisition, music training and transfer across the lifespan. *Frontiers in Systems Neuroscience*, *7*(November), 90. http://doi.org/10.3389/fnsys.2013.00090

Whiteside, S. P., Grobler, S., Windsor, F., & Varley, R. (2010). An acoustic study of vowels and coarticulation as a function of utterance type: A case of acquired apraxia of speech. *Journal of Neurolinguistics*, *23*(2), 145–161. http://doi.org/10.1016/j.jneuroling.2009.12.002

Wieland, E. A., McAuley, J. D., Dilley, L. C., & Chang, S. E. (2015). Evidence for a rhythm perception deficit in children who stutter. *Brain and Language*, 144, 26-34.

Wilden, I., Herzel, H., Peters, G., & Tembrock, G. (1998). Subharmonics, biphonation, and deterministic chaos in mammal vocalization. *Bioacoustics*, 9(3), 171-196.

Will, E., Fidler, D., & Daunhauer, L. A. (2014). Executive function and planning in early development in down syndrome. *International Review of Research in Developmental Disabilities*, 47, 77–98. http://doi.org/10.1016/B978-0-12-800278-0.00003-8

Williamson, V. J., Baddeley, A. D., & Hitch, G. J. (2010). Musicians' and nonmusicians' short-term memory for verbal and musical sequences: comparing phonological similarity and pitch proximity. *Memory & Cognition*, 38(2), 163–75. http://doi.org/10.3758/MC.38.2.163

Wise, K. J., & Sloboda, J. A. (2008). Establishing an empirical profile of self-defined "tone deafness": Perception, singing performance and self-assessment. *Musicae Scientiae*, 12(1), 3-26.

Wishart, J. (2001). Motivation and learning styles in young children with Down syndrome. *Down syndrome Research and Practice*, 7(2), 47-51.

Wisniewski, M. G., Mantell, J. T., & Pfordresher, P. Q. (2013). Transfer effects in the vocal imitation of speech and song. *Psychomusicology: Music, Mind, and Brain*, 23(2), 82.

Wong, T. P. S., Leung, E. Y. W., Poon, C. Y. C., Leung, C. Y. F., & Lau, B. P. H. (2013). Balance performance in children with unilateral and bilateral severe-to-profound-grade hearing impairment. *Hong Kong Physiotherapy Journal*, 31(2), 81–87. http://doi.org/10.1016/j.hkpj.2013.07.001

Wood, S. (2010). Electropalatography in the assessment and treatment of speech difficulties in children with Down syndrome. *Down Syndrome Research and Practice*, 12(2), 98-102.

Woodhouse J.M. (2005). Vision in children with Down syndrome: a research update. *Down Syndrome News and Update,* 4(3);87-89.

Wu, J., & Morris, J. K. (2013). The population prevalence of Down's syndrome in England and Wales in 2011. *European Journal of Human Genetics*, *21*(9), 1016–1019. http://doi.org/10.1038/ejhg.2012.294

Wuang, Y.-P., & Su, C.-Y. (2011). Correlations of sensory processing and visual organization ability with participation in school-aged children with Down syndrome. *Research in Developmental Disabilities*, *32*(6), 2398–407. http://doi.org/10.1016/j.ridd.2011.07.020

Xue, S. A., Kaine, L., & Ng, M. L. (2010). Quantification of vocal tract configuration of older children with Down syndrome: a pilot study. *International Journal of Pediatric Otorhinolaryngology*, *74*(4), 378–83. http://doi.org/10.1016/j.ijporl.2010.01.007

Yap, T. F. (2012). Speech Production Under Cognitive Load: Effects and Classification. PhD thesis. Retrieved from http://unsworks.unsw.edu.au/fapi/datastream/unsworks:10507/SOURCE02

Yap, T. F., Epps, J., Ambikairajah, E., Choi, E. H. C., & Wales, S. (2015).*Voice source features for cognitive load classification*. School of Electrical Engineering and Telecommunications, ATP Research Laboratory, National ICT Australia (NICTA), Eveleigh 2015, Australia, 5700–5703.

Zampini, L., Fasolo, M., Spinelli, M., Zanchi, P., Suttora, C., & Salerni, N. (2015). Prosodic skills in children with Down syndrome and in typically developing children. *International Journal of Language & Communication Disorders / Royal College of Speech & Language Therapists*, 1–10. http://doi.org/10.1111/1460-6984.12186

Zatorre, R. J., & Baum, S. R. (2012). Musical melody and speech intonation: Singing a different tune. *PLoS Biology*, *10*(7), 5. http://doi.org/10.1371/journal.pbio.1001372

Zentner, M., & Eerola, T. (2010). Rhythmic engagement with music in infancy. *Proceedings of the National Academy of Sciences*, 107(13), 5768-5773.

Zhang, Z., Neubauer, J., & Berry, D. A. (2006). The influence of subglottal acoustics on laboratory models of phonation. *Journal of the Acoustical Society of America*, *120*(3), 1558–1569. http://doi.org/10.1121/1.2225682

Zoller, M. B. (1991). Use of music activities in speech-language therapy. *Language, Speech, and Hearing Services in Schools*, 22(1), 272-276.