Copyright Statement

The digital copy of this thesis is protected by the Copyright Act 1994 (New Zealand).

This thesis may be consulted by you, provided you comply with the provisions of the Act and the following conditions of use:

- Any use you make of these documents or images must be for research or private study purposes only, and you may not make them available to any other person.
- Authors control the copyright of their thesis. You will recognize the author's right to be identified as the author of this thesis, and due acknowledgement will be made to the author where appropriate.
- You will obtain the author's permission before publishing any material from their thesis.

General copyright and disclaimer

In addition to the above conditions, authors give their consent for the digital copy of their work to be used subject to the conditions specified on the Library Thesis Consent Form and Deposit Licence.
ASSESSING AND TRAINING AFFECTIVE PROSODIC PERCEPTION AND PRODUCTION IN CHILDREN WITH AUTISM SPECTRUM DISORDER:
A BEHAVIOURAL AND NEUROPHYSIOLOGICAL EXPLORATION INTO THE USE OF COMPUTER-BASED ACTIVITIES AND REMOTE MICROPHONE HEARING AIDS FOR AUDITORY PROCESSING

JOAN HUAN LEUNG

A THESIS SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY IN PSYCHOLOGY
THE UNIVERSITY OF AUCKLAND 2017
Abstract

Aims: This thesis comprised a series of studies that examined the assessment and training of affective prosody perception and production in children with high-functioning Autism Spectrum Disorder (HF-ASD), using a range of methodologies. The studies are motivated by evidence that individuals with ASD experience significant auditory processing difficulties, which hinders their perception of the subtle cues in speech that govern emotion expression (affective prosody). This is thought to also affect the production of affective prosodic cues. A number of studies have reported successful behavioural training of emotion perception in children and adults with ASD, especially when delivered via a computer- or technology-based platform. One aim of this research was to investigate whether enhancing the incoming auditory signal by using remote microphone hearing aids (RMHAs) would further improve the effects of computer-based emotion training. Another focus of this research was to explore the neurophysiological patterns underlying affective prosody perception, to determine whether this is different between typically developing (TD) and HF-ASD individuals, and whether these are influenced by emotion training and RMHA use.

Methods: The first study implemented a 3-week computer-based emotion training intervention with 24 children with HF-ASD. Twelve of those children underwent computer-based training only. The other 12 children wore RMHAs during the computer-based training sessions, and in addition trialled the RMHAs at school for the duration of the intervention period. The HF-ASD participants were assessed on prosody perception and production, twice before and twice after the intervention. Fourteen TD children made up the control group, and were assessed at one time point. Prosody production was analysed objectively by examining acoustic parameters of recorded emotional speech productions, and subjectively via third
party listeners rating the conveyed emotion and intensity of emotion. The listeners were blinded to the intended emotion and the recording session (pre- versus post-intervention). Pre- and post-intervention effects were examined for both HF-ASD groups, and results from the computer training only group were compared to those from children who received the combined intervention. Data from the HF-ASD children were also compared to those from their TD peers. Furthermore, HF-ASD children from the RMHA intervention group, along with their teachers, completed questionnaires to evaluate whether wearing the RMHAs influenced classroom listening experiences and behaviours.

The second study involved the recording of cortical auditory evoked potentials (CAEPs) from TD adult participants. Natural speech stimuli with angry, happy, sad, and neutral emotional tones were presented in a passive oddball paradigm. Angry, happy, and sad served as deviant stimuli, occurring less frequently in the sequence of stimuli mainly made up of the neutral-standard. Subtracting the CAEP response of the standard from that of a deviant, results in a derived mismatch response (MMR) representing pre-attentive auditory change detection of affective prosody. CAEP and MMR morphology was examined for the adult participants, and investigated for differences between emotions, with the purpose of evaluating the feasibility of the novel natural speech stimuli introduced in this study.

The third study involved the additional recording of CAEPs and deriving of MMRs from TD children participants (on one occasion), as well as HF-ASD participants who were recorded before and after they underwent computer-based emotion training combined with RMHAs. CAEP and MMR morphology was compared between 1) TD adults and children, 2) TD children and HF-ASD children, and 3) HF-ASD children before and after intervention.

**Results:** Outcomes from the first study showed that children with HF-ASD were able to engage in the computer-based emotion training intervention, and that repeated exposure to
the activities improved their abilities in both facial expression recognition and affective prosody perception. There were no large differences between children who went through the intervention with the RMHAs, and those who completed the training without the RMHAs. However, only the RMHA intervention group maintained a stable, high performance on affective prosody perception during the 2-week follow up period post-intervention, suggesting better maintenance of training effects when training is combined with RMHAs (effectively increasing the ‘auditory dose’).

The objective acoustic analyses and the subjective raters’ judgements of affective speech productions did not reveal substantial intervention effects for either group of HF-ASD children. However, correlation results indicated that several acoustic factors – namely mean fundamental frequency (pitch), various measures of pitch range, mean intensity (loudness), and various measures of intensity range – were correlated with raters’ emotion perceptions indicating that these acoustic features could be considered to contribute to, or even drive, the subjective perception of emotion in speech.

Children with HF-ASD who trialled the RMHA at school reported significantly improved listening experiences in the classroom, in particular in a noisy environment. Their teachers also reported significantly improved listening behaviours, and evaluated the effectiveness of the RMHA positively.

Joint outcomes from the second and third studies showed that the auditory stimuli, implemented for the first time in a neurophysiology experiment, were successful in evoking typical adult and children CAEP responses that corroborated the existing literature, although latencies were generally later than what is typically reported, which could reflect the complexity of the stimuli. In addition, all participants showed a late N400 response in their standard and deviant waveforms. In previous studies N400 has been associated with the
cognitive appraisal of semantic and emotional information, thus the presence of N400 suggests that the auditory stimuli successfully engaged neural processes associated with emotion processing.

There were significant differences in CAEP and MMR morphology between TD children and children with HF-ASD. Recordings from the children with HF-ASD before they underwent training did not show any significant separation between mismatch difference waves for angry, happy and sad emotions. This suggests that, unlike their TD peers and the TD adults, children with HF-ASD are to detect changes in the ongoing stimulus stream when a deviant was presented, but did not differentiate the emotional sounds based on either their acoustic features or emotional content. Recordings from the same HF-ASD children after intervention revealed significant differentiation between emotions. However, unlike their TD peers, both late and early MMRs showed differences between emotions. It is hypothesised that this may reflect an increase in perceptual effort or attention required for participants with HF-ASD to perceive changes in the auditory stimulus sequence compared to their TD peers who had access to more pre-attentive automatic processes for differentiating the emotional deviants from the neutral standard stimuli.

**Conclusion:** It is the hope that this series of studies will lead to further research investigating atypical affective prosody and other aspects of auditory processing in children with ASD, using both behavioural and neurophysiological methods. Results suggest the potential benefits to social communication of developing interventions targeting social perception of audio-visual stimuli for children with ASD. The recording and analysis of CAEPs can potentially be considered alongside traditional diagnostic practices of clinical observation and behavioural assessments to screen for early markers of affective prosody and auditory processing dysfunction associated with ASD.
Acknowledgements

First and foremost, I would like to sincerely thank my truly wonderful supervisor, Professor Suzanne Purdy. You have been the most amazing mentor, confidante, and inspiration throughout this entire journey. Thank you for always being overly generous with your time, wisdom, and support. I am grateful to have been given the opportunity to work with you, and to have come out with experiences and knowledge that will last me a lifetime.

I would also like to thank my great co-supervisor, Associate Professor Paul Corballis. Thank you for going above and beyond everything that was asked of you in this role, and for your constant words of encouragement. Thank you for your patience and invaluable guidance through the toughest parts of this thesis.

I am indebted to the generosity of Phonak AG, and in particular Mr. Stefan Launer, for providing me with a scholarship for this PhD, as well as providing me with the hearing amplification equipment for this research. I would also like to thank Mr. David Crowhen from Phonak New Zealand for generously giving up his time to provide logistical support.

Many thanks to Dr. Erin Schafer from the University of North Texas for unreservedly sharing the resources that she created for introducing children to personal amplification systems. And also many thanks to Bill Keith and Jenny Coutts from SoundSkills Auckland for helping with volunteer recruitment to film the demonstration video.

Thanks to Autism New Zealand, the Wilson Home Trust, the New Zealand Herald, and the Minds 4 Minds Autism network for their help with spreading the word about this research and participant recruitment.
A huge appreciative thank you to all my participants and their parents who have given me their valuable time and support. Thank you for believing in my work, and for your winning attitude towards promoting and getting involved in research in the field of Autism. This thesis is for you, and I hope its outcomes will contribute to advocating for more awareness and dedication of services for your aid.

Many thanks to my fellow PhD-ers: Abin, Oscar, Rose, Sylvia, Areej, Tai, Boon, Giulia, Sreekari, Phil, Kristina, and Ying. It would not have been the same without your friendship and daily encouragement. Thank you for being there through thick and thin.

And to my other pillars of support in New Zealand, as well as all around the world: Shiga, Cat, Sam, Renique, Caroline, Kerry, Kylie, Martin, Zdenda, and Monicka – thank you.

Much thanks to all the Speech Science and Psychology support staff: Sue, Sheryl, Adeline, Kamalini, Bronwyn, Evgeniya, and Kevin.

And last but not least, thank you to my family for your unconditional love. Thank you for supporting me mentally, physically, emotionally, and gastronomically. And especially thank you to Mum and Dad for teaching me to respect the value of further education, to attempt everything with conviction, to persevere with my passions, and to use the best of my abilities to help others.
# Table of Contents

Abstract .......................................................................................................................... ii  
Acknowledgements ......................................................................................................... vi  
Table of Contents ........................................................................................................... viii  
List of Tables .................................................................................................................. xii  
List of Figures ................................................................................................................ xiv  
Co-Authorship Forms ..................................................................................................... xvi  
Chapter 1. Introduction ................................................................................................ 1  
  Autism Spectrum Disorder: A theoretical framework ............................................ 2  
  Prevalence and diagnosis ......................................................................................... 3  
  New Zealand guidelines on diagnostic procedures for ASD .............................. 6  
  Thesis structure and chapter outline .................................................................... 8  
Chapter 2. Socio-emotional and prosody processing, auditory processing, and training interventions in Autism Spectrum Disorder .............................................. 12  
  Socio-emotional perception in ASD: Faces and facial expressions .................. 13  
  Socio-emotional perception in ASD – Speech and communication .................. 14  
    Prosody .................................................................................................................. 15  
    Prosody perception in ASD ................................................................................. 18  
    Prosody production in ASD ................................................................................. 21  
  Assessment of social perception in ASD ............................................................... 22  
  Auditory processing and ASD ................................................................................. 25  
  Training of social communication and auditory management in ASD .......... 28  
    Training of social skills and the use of computer-based technology .......... 29  
    Auditory management and training with RMHAs ........................................... 32  
Chapter 3. Documentation of the development of training material and the implementation of remote microphone hearing aids for the study .................................................. 35  
  The development of the emotion training sessions ........................................... 36  
  Remote microphone hearing aids (RMHAs) ......................................................... 41  
  Verification procedure for RMHA fitting ........................................................... 42  
  Familiarisation procedure for RMHA system (adapted from Schafer et al., 2013) 43  
  Production of the demonstration and introduction video to RMHAs .............. 45
Chapter 4. Assessing the effects of a computer-based training intervention, and a trial of remote microphone hearing aids, with children with high-functioning Autism Spectrum Disorder: A study of prosody perception and production, and listening experiences in the classroom ................................................................. 48

Introduction .................................................................................................................. 49

Method ......................................................................................................................... 56

- Study design ............................................................................................................. 56
- Participants ............................................................................................................... 58
- Assessment Tasks .................................................................................................... 63

Editing speech recordings and generating acoustic values ............................................ 66

Raters’ judgements of the speech samples .................................................................... 67

Rater reliability ............................................................................................................. 69

Computer-based training tasks .................................................................................... 70

Statistical analyses ..................................................................................................... 71

Results ......................................................................................................................... 72

- Affective social and prosody perception .................................................................. 72
- Affective prosody production: Acoustic analyses ....................................................... 77
- Affective prosody production: Raters’ judgements .................................................... 80
- Correlations between acoustic parameters and raters’ judgements ......................... 83
- Evaluation of RMHA trial in the classroom ............................................................... 85

Discussion .................................................................................................................... 87

- Emotion and prosody perception ............................................................................ 87
- Prosody production: Acoustics and raters’ judgements ............................................ 90
- Correlations between acoustics and raters’ judgements ........................................... 91
- Trial of RMHA systems in the classroom .................................................................. 92
- Limitations and future directions ............................................................................. 93

Chapter 5. Cortical auditory evoked potentials and Mismatch responses ...................... 96

- Cortical auditory evoked potentials (CAEPs) and Mismatch negativity (MMN) .......... 97
- CAEP components: Acquisition, interpretation, and limitations ............................. 97
- The Mismatch Negativity/Response (MMN/MMR) effect ........................................ 102

MMRs in clinical populations ....................................................................................... 110

- Potential limitations and possible solutions ........................................................... 114
- CAEPs and MMRs in ASD ..................................................................................... 115

Studies 3 and 4: Background methodology .................................................................. 121

- The making and selection of the speech stimuli ....................................................... 121
- Speech stimuli presentation settings ....................................................................... 122
Chapter 6. Exploring change detection in affective prosody in adults through cortical auditory evoked potentials and mismatch responses

Introduction .................................................................................................................. 123
Methods .......................................................................................................................... 124
Participants ...................................................................................................................... 127
Cortical auditory evoked potential recordings ................................................................. 128
Data analysis ................................................................................................................... 131
Results ............................................................................................................................ 134
1) CAEP components from all emotions presented as standard stimuli ......................... 134
2) MMRs in ‘Emotion as deviant’ – ‘Neutral as standard’ waveforms .............................. 138
3) MMRs in ‘Neutral as deviant’ – ‘Emotions as standard’ waveforms ......................... 140
4) MMRs in ‘Deviant - Standard’ waveforms for each of the four emotions ................. 143
Discussion ..................................................................................................................... 145

Chapter 7. Exploring change detection in affective prosody through cortical auditory evoked potentials and mismatch responses: Investigating differences between neurotypical controls and individuals with high-functioning Autism Spectrum Disorder, and evaluating the effects of emotion perception training and auditory amplification

Methods .......................................................................................................................... 150
Participants ...................................................................................................................... 158
Cortical auditory evoked potential recordings ................................................................. 159
Experimental setup ......................................................................................................... 160
Data processing .............................................................................................................. 161
Data analysis ................................................................................................................... 162
Results ............................................................................................................................ 165
1) CAEP components from ‘neutral as standard’ and ‘emotions as deviants’ waveforms 165
2) Magnitude, polarity, and timing of MMRs between participant groups ...................... 174
3) Emotion effects within each MMR between participant groups ................................. 179
Discussion ..................................................................................................................... 182
Between group differences for CAEPs .......................................................................... 182
Between group differences in emotion change detection, exhibited by MMRs ............ 185
Limitations and future directions .................................................................................... 189

Chapter 8. General Discussion and Conclusion

Overview of research ..................................................................................................... 191
Summary of key findings ............................................................................................... 193
Novel aspects of this research ....................................................................................... 199
Limitations of the research ................................................................. 201
Future directions ............................................................................... 203
Appendices ....................................................................................... 207
Appendix 1: Ethics approval ............................................................... 208
Appendix 2: Ethics amendment approval ........................................... 209
Appendix 3: Recruitment advertisement (Study 1) ........................... 210
Appendix 4: Recruitment advertisement (Study 2) ........................... 211
Appendix 5: Recruitment advertisement (Controls) .......................... 212
Appendix 6: Participant information sheet (Study 1) – for ASD children ................................................................. 213
Appendix 7: Participant information sheet (Study 1) – for parents/caregivers of ASD ................................................................. 215
Appendix 8: Participant information sheet (Study 2) – for ASD children ................................................................. 221
Appendix 9: Participant information sheet (Study 2) – for parents/caregivers of ASD ................................................................. 224
Appendix 10: Participant information sheet (Study 2) – for Control children ................................................................. 231
Appendix 11: Participant information sheet (Study 2) – for parents/caregivers of Controls ................................................................. 233
Appendix 12: Participant information sheet (Study 2) – for Adult controls ................................................................. 238
Appendix 13: Participant information sheet (Study 2) – for Teachers ................................................................. 243
Appendix 14: Participant questionnaire ................................................ 248
Appendix 15: Childhood Autism Rating Scale (CARS-2) – Parents/Caregivers questionnaire ................................................................. 249
Appendix 16: Childhood Autism Rating Scale (CARS-2) – High functioning version scoring guide ................................................................. 255
Appendix 17: Children’s Communication Checklist (CCC-2) ................. 264
Appendix 18: Children’s Communication Checklist (CCC-2) – Score sheet .............................................................................. 268
Appendix 19: Listening Inventory for Education – Student version (LIFE-7, adapted for NZ) ................................................................. 269
Appendix 20: Listening Inventory for Education – Teacher version (LIFE-UK) ................................................................. 274
Appendix 21: Listening Inventory for Education – Teacher version (LIFE-R) ................................................................. 276
Appendix 22: Advanced Clinical Solutions – Social Perception subtests score sheet ................................................................. 277
Appendix 23: Social story 1 for RMHA familiarisation ................................................................. 279
Appendix 24: Social story 2 for RMHA familiarisation ................................................................. 280
Appendix 25: Script for RMHA demonstration and introduction video ................................................................. 281
References .......................................................................................... 283
List of Tables

Table 1. Studies in this PhD thesis and the participants involved in each..............................11

Table 2. Distortion product optoacoustic emission (DPOAE) signal-to-noise ratios (SNR) measured in decibels (dB) from TD and HF-ASD children (computer and RMHA groups). DPOAEs were collected from the left and right ears of each participant, across five frequency levels ranging from 2-6 kHz .................................................................................................................................61

Table 3. Group average, minimum, and maximum scores for Childhood Autism Rating Scale (CARS-2) from HF-ASD participants; and General Communication Composite (GCC) and Social Interaction Deviance Composite (SIDC) scores from the Childhood Communication Checklist (CCC-2), from all participants ........................................................................................................................................63

Table 4. Post hoc analyses results for the Affect Naming and Social Perception Prosody scores from the ACS Social Perception subtests. Means (M) and standard deviations (SD) are shown in parentheses, and p-values are shown for each comparison. Session A and B indicate the two sessions in each pairwise comparison. * Bonferroni corrected p<.008 statistical significance. ........................................................................................................................................73

Table 5. One-way ANOVA and post hoc analyses results for the Affect Naming and Social Perception Prosody scores. F-statistics (degrees of freedom with and without error), and p-values for between group effects are reported. Means (M) and standard deviations (SD) are shown for each participant group, followed by p-values for post hoc comparisons at each session. Participant groups are coded 1 for TD children, 2 for HF-ASD Computer, and 3 for HF-ASD RMHA. ........................................................................................................................................76

Table 6. One-way ANOVA and significant post hoc comparison results for between group effects on acoustic parameters measured for angry, happy, and sad across assessment sessions. ........................................................................................................................................78

Table 7. Percentage of time raters judged the intended emotion of the participants correctly, and when they confused it with the other two alternative emotions ........................................82

Table 8. Significant correlations between acoustic parameters and accuracy of raters’ judgements for each emotion, calculated separately for TD and HF-ASD groups. ..........................84

Table 9. Total number of stimulus presentations per emotion, in their two different forms (as standard and as deviant) ........................................................................................................................................129

Table 10. Repeated measures ANOVA results comparing MMR magnitudes between emotions at identified regions of mismatch. df (e) – degrees of freedom (error), M – mean, SD – standard deviation ........................................................................................................................................141

Table 11. Latency range and average amplitudes of MMRs for each emotion ..................144
Table 12. DPOAE results (signal-to-noise ratios, SNR) measured in decibels (dB) from TD and HF-ASD children. DPOAEs were collected from the left and right ears of each participant, across five frequencies ranging from 2-6 kHz.

Table 13. Independent samples t-test results showing significant differences between typically developing adults and children in CAEP response component amplitudes and latencies, measured for the neutral-standard stimulus condition.

Table 14. Independent samples t-test results showing significant differences between typically developing adults and children in CAEP response component amplitudes and latencies, measured for the emotional-deviant stimulus conditions (combined across angry, happy, and sad CAEP recordings).

Table 15. Paired samples t-test results showing significant differences between HF-ASD children before and after they received the intervention in CAEP response component amplitudes and latencies, measured in the emotional-sounding deviant stimuli condition (combined across Angry, Happy, and Sad CAEP recordings).

Table 16. Timeline format presentation of mismatch latency windows for each participant group, derived by conducting single t-tests at each time point of group average emotion-combined difference waveforms, and identifying contiguous periods of significant deviance from 0 uV.

Table 17. Independent samples t-test, and paired samples t-test results showing significant differences in MMR average amplitude and peak latency for three between group comparisons. Average amplitude is measured in microvolts (uV) and latency in milliseconds (ms). The numbering of the MMR is based on Table 16.

Table 18. Within-subject repeated measures ANOVA results showing significant differences in amplitude and latency between each emotional deviance (angry, happy, and sad) within identified MMR regions from each participant group. Average amplitude is measured in microvolts (uV) and latency in milliseconds (ms). Bonferroni adjusted p<.017 was applied to post hoc result interpretations.
List of Figures

Figure 1. Depiction of computer-training program, task presentation and navigation ..........40

Figure 2. Group average scores from the Affect Naming task of the ACS Social Perception subtests. The score for TD control group obtained for one test occasion only is shown as a dashed line. Scores are shown for the four assessment sessions for the HF-ASD computer and RMHA groups (Sessions 1 and 2 for pre-intervention and Sessions 3 and 4 for post-intervention) .................................................................................................................................74

Figure 3. Group average scores from the Social Perception Prosody tasks (sum of Face- and Pair-matching tasks) of the ACS Social Perception subtests. The score for TD control group obtained for one test occasion only is shown as a dashed line. Scores are shown for the four assessment sessions for the HF-ASD computer and RMHA groups (Sessions 1 and 2 for pre-intervention and Sessions 3 and 4 for post-intervention) .................................................................................................................................74

Figure 4. Pre- and post- RMHA trial at school LIFE questionnaire scores, self-reported levels of listening difficulty by HF-ASD children. The dashed lines represent mean scores for each factor of the questionnaire calculated from normative New Zealand data collected by Purdy and colleagues in 2011. Listening in noise = 2.14±0.83; Listening in quiet = 1.23±0.46; Focused listening = 1.60±0.59; Overall = 1.73±0.51; * represents p≤.05 statistical significance. .....................................................................................................................................85

Figure 5. Pre- and post- RMHA trial at school LIFE questionnaire scores, reflecting the evaluation by the teacher of listening behaviours in the classroom, supporting the positive outcomes of the RMHA trial........................................................................................................................................86

Figure 6. Electrical time waveforms (Adobe Audition CS6) depicting the monosyllabic /ba/ stimulus presented in four emotions. dBV represents relative sound intensity in decibel-voltages, plotted against time in milliseconds. ...............................................................................................................129

Figure 7. Grand average CAEP waveforms generated for each emotion, for the four electrode sites (Cz, Fz, F4, F3) ........................................................................................................................................137

Figure 8. Grand average difference waves averaged across emotions (in black), and underneath, separated grand average difference waves for each emotion (in colour). Presented for the four electrode sites (Cz, Fz, F4, F3). Dev – deviant, Std – standard ..................139

Figure 9. Grand average difference waves averaged across emotions (in black, top), and grand average difference waves for each emotion (in colour, below). Results are shown for the four electrode sites (Cz, Fz, F4, F3). Dev – deviant, Std – standard ..................142

Figure 10. Deviant-standard difference waveforms, averaged across electrode sites, for each emotion. The grey shaded areas indicate latency windows. ................................................................................................................144
Figure 11. Column A shows grand average responses to standard (thick line) and deviant (thin line) /ga/ stimulus. Column B shows the difference response (deviant-standard) with a dotted line. The vertical line in Column B shows where the negative MMR offset latency was identified at 250 ms. Reproduced with permission from M. Sharma and colleagues (2006).

Figure 12. Electrical time waveforms (Adobe Audition CS6) depicting the monosyllabic stimulus /ba/ stimulus presented in four emotions. dBV represents relative sound intensity in decibel-voltages, plotted against time in milliseconds.

Figure 13. CAEP waveforms from typically developing adults and children, for both neutral-as-standard and emotions-as-deviant stimulus presentation conditions. Response potentials are measured in microvolts (uV) and time in milliseconds (ms), across four electrode sites (Cz, Fz, F4, and F3).

Figure 14. CAEP waveforms from TD children and HF-ASD children (recorded before intervention), for both neutral-as-standard and emotions-as-deviant stimulus presentation conditions. Response potentials are measured in microvolts (uV) and time in milliseconds (ms), across four electrode sites (Cz, Fz, F4, and F3).

Figure 15. CAEP waveforms from HF-ASD children recorded before and after intervention, for both neutral-as-standard and emotions-as-deviant stimulus presentation conditions. Response potentials are measured in microvolts (uV) and time in milliseconds (ms), across four electrode sites (Cz, Fz, F4, and F3).

Figure 16. Group average difference waveforms (neutral-as-standard minus emotions-as-deviants), combined across emotions and electrode sites, resulting in a single grand average difference waveform for each group – typically developing adults, children, and HF-ASD children before and after intervention. Response amplitudes are measured in microvolts (uV) and time in milliseconds (ms).

Figure 17. Grand average difference waveforms for each emotional deviance, collapsed across electrodes, presented for each participant group. Response potentials are measured in microvolts (uV) and time in milliseconds (ms).
Co-Authorship Forms

This form is to accompany the submission of any PhD that contains published or unpublished co-authored work. Please include one copy of this form for each co-authored work. Completed forms should be included in all copies of your thesis submitted for examination and library deposit (including digital deposit), following your thesis Acknowledgements. Co-authored works may be included in a thesis if the candidate has written all or the majority of the text and had their contribution confirmed by all co-authors as not less than 65%.

Please indicate the chapter/section/pages of this thesis that are extracted from a co-authored work and give the title and publication details or details of submission of the co-authored work.

Chapter 4
Assessing the effects of a computer-based training intervention, and a trial of remote microphone hearing aids, with children with high-functioning Autism Spectrum Disorder: A study of prosody perception and production, and listening experiences in the classroom
In preparation for manuscript submission

<table>
<thead>
<tr>
<th>Nature of contribution by PhD candidate</th>
<th>Author of manuscript, design of experiment, collection of data, analysis of data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extent of contribution by PhD candidate (%)</td>
<td>95</td>
</tr>
</tbody>
</table>

CO-AUTHORS

<table>
<thead>
<tr>
<th>Name</th>
<th>Nature of Contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suzanne C. Purdy</td>
<td>Supervisor, advice on experimental design and data analysis, feedback on manuscript</td>
</tr>
</tbody>
</table>

Certification by Co-Authors

The undersigned hereby certify that:

- the above statement correctly reflects the nature and extent of the PhD candidate’s contribution to this work, and the nature of the contribution of each of the co-authors; and
- that the candidate wrote all or the majority of the text.

<table>
<thead>
<tr>
<th>Name</th>
<th>Signature</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suzanne C. Purdy</td>
<td></td>
<td>1 November 2016</td>
</tr>
</tbody>
</table>

Last updated: 19 October 2015
This form is to accompany the submission of any PhD that contains published or unpublished co-authored work. Please include one copy of this form for each co-authored work. Completed forms should be included in all copies of your thesis submitted for examination and library deposit (including digital deposit), following your thesis Acknowledgements. Co-authored works may be included in a thesis if the candidate has written all or the majority of the text and had their contribution confirmed by all co-authors as not less than 65%.

Please indicate the chapter/section/pages of this thesis that are extracted from a co-authored work and give the title and publication details or details of submission of the co-authored work.

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Exploring change detection in affective prosody in adults through cortical auditory evoked potentials and mismatch responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>In preparation for manuscript submission</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Nature of contribution by PhD candidate</th>
<th>Author of manuscript, design of experiment, collection of data, analysis of data</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Extent of contribution by PhD candidate (%)</th>
<th>90</th>
</tr>
</thead>
</table>

---

### CO-AUTHORS

<table>
<thead>
<tr>
<th>Name</th>
<th>Nature of Contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suzanne C. Purdy</td>
<td>Supervisor, advice on experimental design and data analysis, feedback on manuscript</td>
</tr>
<tr>
<td>Paul M. Corballis</td>
<td>Co-supervisor, advice on data analysis, feedback on manuscript</td>
</tr>
</tbody>
</table>

---

### Certification by Co-Authors

The undersigned hereby certify that:
- the above statement correctly reflects the nature and extent of the PhD candidate’s contribution to this work, and the nature of the contribution of each of the co-authors; and
- that the candidate wrote all or the majority of the text.

<table>
<thead>
<tr>
<th>Name</th>
<th>Signature</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suzanne C. Purdy</td>
<td></td>
<td>28 October 2016</td>
</tr>
<tr>
<td>Paul M. Corballis</td>
<td></td>
<td>1 November, 2016</td>
</tr>
</tbody>
</table>

Last updated: 19 October 2015
This form is to accompany the submission of any PhD that contains published or unpublished co-authored work. Please include one copy of this form for each co-authored work. Completed forms should be included in all copies of your thesis submitted for examination and library deposit (including digital deposit), following your thesis Acknowledgements. Co-authored works may be included in a thesis if the candidate has written all or the majority of the text and had their contribution confirmed by all co-authors as not less than 65%.

Please indicate the chapter/section/pages of this thesis that are extracted from a co-authored work and give the title and publication details or details of submission of the co-authored work.

<table>
<thead>
<tr>
<th>Nature of contribution by PhD candidate</th>
<th>Extent of contribution by PhD candidate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author of manuscript, design of experiment, collection of data, analysis of data</td>
<td>90</td>
</tr>
</tbody>
</table>

### CO-AUTHORS

<table>
<thead>
<tr>
<th>Name</th>
<th>Nature of Contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suzanne C. Purdy</td>
<td>Supervisor, advice on experimental design and data analysis, feedback on manuscript</td>
</tr>
<tr>
<td>Paul M. Corbellis</td>
<td>Co-supervisor, advice on data analysis, feedback on manuscript</td>
</tr>
</tbody>
</table>

### Certification by Co-Authors

The undersigned hereby certify that:
- the above statement correctly reflects the nature and extent of the PhD candidate’s contribution to this work, and the nature of the contribution of each of the co-authors; and
- that the candidate wrote all or the majority of the text.

<table>
<thead>
<tr>
<th>Name</th>
<th>Signature</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suzanne C. Purdy</td>
<td>[Signature]</td>
<td>28 October 2016</td>
</tr>
<tr>
<td>Paul M. Corbellis</td>
<td>[Signature]</td>
<td>1 November, 2016</td>
</tr>
</tbody>
</table>

Last updated: 19 October 2015
Autism Spectrum Disorder: A theoretical framework

There is a general consensus in the field of psychology that Autism Spectrum Disorder (ASD) should be viewed as a lifelong, pervasive, neurodevelopmental condition (Diagnostic and Statistical Manual of Mental Disorders – 5th edition, DSM-V; American Psychiatric Association, 2013). The considerable complexity of ASD has contributed to the current agreement that there are multiple causes of the condition (Sealey et al., 2016). Many theoretical frameworks proposed in the literature have links to cognitive, behavioural, and biological aspects of ASD.

One of the dominant theories suggests that individuals with ASD lack Theory of Mind abilities, which explains their difficulties in recognising and understanding the mental states and intentions of others (McCann & Peppé, 2003; Rutherford, Baron-Cohen & Wheelwright, 2002; Tager-Flusberg, 1999). It is thought that atypical reduction in joint attention, observed during the earliest years of life (Schertz & Odom, 2004), underlies the underdevelopment of a social awareness of others. Infants and toddlers with ASD show diminished sharing of attention towards events or objects with other people (Charman et al., 1997; Dawson et al., 2004), which is hypothesised to lead to impairments in social communication and relationship-forming later in life (Mundy, 1995; Mundy, Sigman & Kasari, 1994).

Another theory hypothesises that individuals with ASD have an intrinsic lack of intention to use oneself as a social agent to influence the behaviours and attitudes of other people (Prizant & Wetherby, 1985). The absence of motivation to communicate is thought to be interconnected with the difficulties in language processing exhibited by those with ASD (Frith, 1989; Frith & Happé, 1994; Minshew, Goldstein & Siegel, 1995; Tager-Flusberg, 1996). Abnormal development and activation of a range of neural regions have been associated with dysfunction in face (fusiform gyrus), speech (superior temporal sulcus), and
emotion processing (anterior cingulate cortex) in children with ASD (Amaral, Schumann & Nordahl, 2008; Minshew & Keller, 2008).

Another view of ASD is that impaired executive functioning underlies many of the observed restrictive and repetitive behavioural symptoms (Hill, 2004). Difficulty with impulse control and action monitoring, as well as a reluctance to continuously shift one’s mind set, are commonly observed in children and adults with ASD (Hughes, Russell & Robbins, 1994; Ozonoff, Pennington & Rogers, 1991). Abnormalities in the development and neural activation of the pre-frontal lobe have been associated with the manifestation of these dysfunctional behaviours (Stuss & Alexander, 2000; Tager-Flusberg, 2008). Elisabeth Hill (2004) noted that there are some difficulties with this view of ASD, concluding that little is known about neuroanatomical correlates of executive function in people with ASD and that investigating executive functions is complex in ASD due to the impact of other potential deficits such IQ and attention.

The series of studies presented in this PhD have approached ASD from a neurodevelopmental perspective – focusing on investigating atypical neural processes that may underlie behavioural deficits, and comparing these effects between children and adults. It was hypothesised that fundamental impairments in auditory processing and processing of speech prosody may contribute to the language, social communication, and Theory of Mind deficits that have been observed in people with ASD.

**Prevalence and diagnosis**

ASD has one of the fastest increasing prevalence rates amongst other clinically recognised conditions (Anagnostou et al., 2014). Between 1966 and 1998, approximately 1 in 220 children assessed had some diagnosis of autism (Fombonne, 1999). Since then, prevalence figures in the early 21st century have ranged from 1/132 to 1/160, with developed
and developing countries differing in levels of awareness, the availability of diagnostic procedures, and governmental policies around the special needs population, and consequently reporting different prevalence rates (Elsabbagh et al., 2012; Fombonne, 2002; Fombonne, 2005; Baxter et al., 2014). Studies on small population cohorts have reported rates as high as 1/64 (Cambridgeshire, UK; Baron-Cohen et al., 2009), 1/86 (South Thames, UL; Baird et al., 2006), 1/87 (Stockholm, Sweden; Idring et al., 2012), and 1/111 (Wisconsin, USA; Harrington, 2010). In smaller countries, such as New Zealand, autism is thought to affect the lives of around 77,500 people, which is a rate of approximately 1/58 (Autism New Zealand, 2016). Despite widely varying prevalence rates, the ratio of every 4 boys to 1 girl being diagnosed with some form of autism, remains relatively stable (Fombonne, 2003).

Change in diagnostic criteria across the years is a significant influencing factor in the documented prevalence of autism. An increase in prevalence during the last two decades of the 20th century likely, in part, stemmed from having one category of childhood autism in the 3rd edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-III, American Psychiatric Association, 1980), to having a wider variety of different conditions under the same umbrella (autistic disorder, Asperger disorder, and pervasive developmental disorder – not otherwise specified (PDD-NOS) in DSM-IV, text revised in 2000). Although the most recent progression from DSM-IV to DSM-V has returned to a single diagnosis of autism along a spectrum of severity (ASD), the prevalence did not seem to be affected, with recent rates still around 2.5%. On one hand, the collective inclusion of individuals displaying similar symptoms has been proposed to reflect the true prevalence of ASD (DSM-V, APA, 2013). Studies have reported that many individuals were successfully re-diagnosed under the new criteria (Kim et al., 2014; McPartland, Reichow & Volkmar, 2012). On the other hand, challenges have been raised with regards to the increased stringency of the criteria. Studies have also reported that cognitively able individuals of different ages (Matson, Belva,
Horovitz, Kozlowski & Bamburg, 2012a; Matson, Kozlowski, Hattier, Horovitz & Sipes, 2012c), who were previously diagnosed with Asperger disorder or PDD-NOS (McPartland et al., 2012), failed to be diagnosed under the new criteria, compared to when they were screened with alternative assessments (for example, using the Childhood Autism Rating Scale (CARS), or the Checklist for Autism Spectrum Disorder (CASD); Mayes et al., 2014).

Which leads to another major issue that researchers and clinicians face when studying and diagnosing ASD - the lack of a single established tool that can be used for an unequivocal diagnosis. Due to many factors that remain unknown about the condition, studies throughout the years have based their recruitment of participants on different criteria, some of which vary greatly depending on underlying theoretical frameworks and how each country’s governments may structure their disability and special needs policies. This results in an absence of consistency between research goals, expectations, methods, participants and result interpretations across the published literature. The current recommended approach to identifying ASD involves a multidisciplinary evaluation with input from different sources (Anagnostou et al., 2014), including language and social communication assessments (Tager-Flusberg & Caronna, 2007, Tager-Flusberg et al., 2009), as well as genetic testing and neuroimaging (Tchaconas & Adesman, 2013).

Despite the issues raised about the new DSM-V criteria, the two main areas of difficulty proposed in the DSM-V by the APA (2013) – social communication deficits, and restrictive, repetitive patterns of behaviour – are widely recognised and hence relatively robust characteristics of ASD. Other assessment tools for ASD have focused primarily on these aspects, and in particular, address problems with social interaction, emotional responses and regulation, verbal and nonverbal communication, body and object use, sensory responses, and anxiety (CARS; Schopler, Van Bourgondien, Wellman & Love, 2010; CASD; Mayes, 2012). The newly termed ‘spectrum disorder’ refers to the DSM-V severity scale (Clinician-
Rated Severity of Autism Spectrum and Social Communication Disorders) that spans three levels, and can be complemented with specific information highlighting further impairments in language, or associated medical, genetic, neurodevelopmental, mental, or behavioural conditions. Instead of assuming that all the deficits listed are part of the condition (Prelock, 2015), as was the case previously, the current approach provides the diagnostic process with more flexibility when evaluating such a heterogeneous clinical population. Additional criteria also examine the presence of symptoms at an early developmental period, impairment in various areas of functioning, and determine whether the symptoms cannot be better explained by intellectual disability or global developmental delay (DSM-V, APA, 2013).

**New Zealand guidelines on diagnostic procedures for ASD**

According to the New Zealand ASD guidelines (Ministries of Health and Education, 2016), there are 3 main age groups corresponding to different developmental stages, that each have documented a most likely diagnostic pathway that would have been followed.

Children aged 1-3 years would most likely present language delay, social impairments, communication impairments, and behavioural and sensory issues. New Zealand currently enforces a developmental surveillance program shared between parents and health professionals, namely a Well Child nurse. Parents are encouraged to proactively monitor and discuss concerns raised by themselves, or other contacts with their child, such as early childhood teachers. Children exhibiting questionable behaviour are usually referred to a paediatrician, or a child development/early intervention specialist. Children with milder presentations of the issues mentioned above may not raise concerns until they are exposed to increased social interactions as they enter a primary school environment aged 4-8 years. The key people at this stage would include teachers who report their concerns to parents, who may follow it up with a health practitioner.
Children older than 8 years are predicted to display emotional or behavioural issues that may be picked up by Child, Adolescent and Family Mental Health Services; or school performance issues pick up by Special Education Needs Coordinators, Resource Teachers of Learning and Behaviour, of the Ministry of Education. Unfortunately, although there are multiple potential referral points and services for families, New Zealand does not currently have a consistent referral and assessment pathway for children with ASD.

An ideal process would start early, and would commence with an evaluation by a developmental specialist or paediatrician. New Zealand’s Ministry of Health set elective guidelines stating that referrals should be seen within 6 months of initiation. A mandatory hearing assessment rules out hearing loss as a potential cause of the language problems observed in the child. The New Zealand guidelines (2016) then suggest a multidisciplinary assessment for ASD, that consists of 1) the child’s family and developmental history (sometimes done via semi-structured interview tools such as the Autism Diagnostic Interview – Revised (ADI-R) schedule); 2) observations of the child’s behaviour taken from more than one setting, for example at home and in school, to determine their levels of communication, social, and play skills (sometimes done using tools such as the Autism Diagnostic Observation Schedule (ADOS)); 3) observation of the child’s behaviour in unstructured settings with familiar and unfamiliar people, to determine their patterns of interaction and social functioning; 4) an assessment of the child in an educational setting; 5) an assessment of the child’s cognition; 6) an assessment of the child’s speech and language competency, carried out by speech language therapists with specific experience with ASD; 7) an assessment of the child’s mental health; 8) an assessment of the child’s physical health, and a range of medical and DNA tests to investigate comorbid conditions such as Fragile X syndrome or epilepsy; and finally 9) an assessment of other factors such as unusual sensory, motor, and coordination issues that affect self-care and daily functioning.
Thesis structure and chapter outline

This thesis is comprised of a series of studies that were undertaken to examine the assessment and training of affective prosody perception and production in children with HF-ASD using a range of methodologies.

Chapter 2 reviews the literature around impaired social processing of facial expressions and affective prosodic cues in speech demonstrated by individuals with ASD. It also links prosody perception difficulties with available literature around auditory processing difficulties experienced by individuals with ASD, and explores the co-morbidity of ASD and auditory processing disorder (APD). A review of existing evidence of the effects of social training follows, focusing on training delivered from a computer- or technology-based platform. The remote microphone hearing aid (RMHA) technology that was utilised in this research is then introduced, and is followed by a review of the literature around auditory training and RMHA interventions for individuals with HF-ASD, as well as hearing loss, APD, and other learning disorders.

In Chapter 3, there is documentation of the processes that were involved in creating a series of computer-based emotion training tasks that were utilised in the training studies. It also includes information about the adaptation of a familiarisation procedure developed by Schafer and colleagues (2013) for trialling RMHAs with children with HF-ASD, steps for RMHA fitting and verification, and the development of the RMHA demonstration video.

The first study is documented in Chapter 4, which examined the effects of implementing the computer-based emotion training program on prosody perception and production, by assessing participants twice before and twice after a 3-week intervention period. Two groups of children with HF-ASD were assessed - one group went through the intervention with just the computer training, and the other group went through computer
training in conjunction with a trial of a RMHA system. Several published studies report positive auditory processing and academic outcomes for RMHAs in children with ASD (Rance, Saunders, Carew, Johansson & Tan 2014; Schafer et al., 2013), but none has examined affective prosody in children with ASD using RMHAs. The children with HF-ASD were also provided with the option to trial the RMHAs at school during classroom teaching time for the 3-week period, and 11 of the 12 participants did this. Research questions for this study include:

- Does the implementation of an emotion training program have an effect on affective prosody perception in children with HF-ASD?
- Does the implementation of an emotion training program have an effect on the acoustic or subjective properties of affective prosodic production from children with HF-ASD?
- Are the effects different between children who received computer training only, and those who received computer training combined with RMHAs?
- Does affective perception and production of participants with HF-ASD pre- and post-intervention differ from typically developing (TD) peers?
- Does wearing a RMHA system at school have an effect on students’ reports of listening experiences, and teachers’ reports of listening behaviours in the classroom?

Chapter 5 reviews the literature around the neurophysiological recording of cortical auditory evoked potentials (CAEPs), as well as the derivation of mismatch negativities (MMNs) and mismatch responses (MMRs). It covers published evidence that has tracked maturational and stimulus manipulation related changes in the morphology of these neural responses, as well as reports of atypical or absent auditory change detection responses from individuals with HF-ASD, and general concerns about recording CAEPs with challenging populations. At the end of the chapter, the background methodology section outlines the
processes that were involved in creating and validating new natural speech stimuli spoken with angry, happy, sad, and neutral emotional expressions. These stimuli have not been utilised in any previous research. Procedures undertaken to set up the experimental environment, including calibrating the intensity of the stimulus output and positioning speakers relative to the participant’s seating location, are also outlined at the end of this chapter.

CAEP data collected from TD adults are reported in Chapter 6. The main purpose of this was to conduct a feasibility study was to establish that the novel stimuli being tested in a complex oddball paradigm evoked robust CAEP and MMRs. The research questions underlying this study were:

- What is the morphology of the CAEPs recorded in response to each emotional speech stimuli?
- In what ways are these similar or different to what is reported in the literature?
- Does the morphology of derived MMRs differ between emotions?

Chapter 7 contains the final study, which examined the effects of prosody training and RMHA use on speech-evoked CAEPs and MMRs in response to changes in affective prosodic cues in the auditory stimuli (angry, happy, and sad). CAEPs were recorded, and MMRs derived, for TD control children and children with HF-ASD, and were compared. Children with HF-ASD had CAEPs and MMRs recorded on two occasions, before and after they participated in the computer training with RMHA intervention. The research questions for this study were:

- Are there any differences in CAEP and MMR morphology between TD adults and children?
- Are there any differences in CAEP and MMR morphology between TD children and children with HF-ASD (before intervention)?

- Does the combined intervention of computer-based emotion training and RMHA use have an effect on the CAEPs and MMRs of children with HF-ASD?

The final chapter (Chapter 8) summarises the key findings from this thesis, highlights the novel aspects of this research, addresses the limitations that arose from these studies, and proposes future research ideas that are clinically relevant and would contribute to the understanding of affective prosody and auditory processing in ASD from both a behavioural and neurophysiological perspective. Table 1 illustrates the three studies and the different participant groups involved in each.

Table 1.
Studies in this PhD thesis and the participants involved in each.

<table>
<thead>
<tr>
<th>Participant Groups</th>
<th>Study 1</th>
<th>Study 2</th>
<th>Study 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>HF-ASD children (Computer intervention only) Pre-intervention assessments</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HF-ASD children (Computer intervention only) Post-intervention assessments</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HF-ASD children (Computer + RMHA intervention) Pre-intervention assessments</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>HF-ASD children (Computer + RMHA intervention) Post-intervention assessments</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>TD children One assessment</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TD adults One assessment</td>
<td></td>
<td>✓</td>
<td></td>
</tr>
</tbody>
</table>
Chapter 2. Socio-emotional and prosody processing, auditory processing, and training interventions in Autism Spectrum Disorder
Socio-emotional perception in ASD: Faces and facial expressions

Being one of two principle ASD-indicative factors, ‘social communication’ is with little doubt a crucial area of focus for intervention, learning, and improvement. Research in this field has shown both behavioural and neurological differences in individuals with ASD compared to their typically developing (TD) peers.

Studies looking into nonverbal communicative behaviour have mainly focused on the visual perception of emotional facial expressions due to its universality (Dawson, Webb & McPartland, 2005; Kimbarow, Quach & Meyerson, 2010). Vocal-facial affect and vocal-affect naming matching tasks are significantly more difficult for people with ASD compared to controls (Boucher, Lewis & Collis, 2000). Individuals with ASD have demonstrated deficits in processing facial identity and expression, finding it more difficult to differentiate faces if emotional expressions are similar, which has been hypothesised to be due to their preferences for restricted and repeated patterns (Robel et al., 2004). It has also been shown that children with ASD are significantly slower than TD controls at recognising basic emotions, and that the severity of their symptoms and attention to the eyes were related to their accuracy (Bal et al., 2010). Although eye gaze habits of individuals with ASD have been documented to be typical at birth, decline in eye contact at ages 2-6 months was reported by Jones and Klin (2013) to be indicative of a future diagnosis of ASD, and this area of research has been widely discussed in the popular media. Poor facial expression perception in individuals with ASD has been linked to tendencies to fixate on the lower features of a face, like the mouth (Klin, Jones, Schultz, Volkmar & Cohen, 2002), and the lack of holistic processing to perceive a coherent facial expression (Gross, 2004).

Differential neurophysiological responses (Webb et al., 2011), and altered activation of neural regions associated with face and emotion processing (Minshew & Keller, 2010),
have also been documented to be consistent with the atypical processes that individuals with ASD seem to engage in when doing this type of task. Visual change detection experiments demonstrate that adults with ASD have a heightened response to different stimuli, regardless of whether the changes are minor or major, which suggests hyper-attention towards individual details instead of applying generalisation rules (Cléry et al., 2013). High-functioning adults with ASD are reported to display later event related potentials (ERPs) to faces and facial features compared to their TD peers, and yet this difference is not seen when viewing objects (O’Connor, Hamm & Kirk, 2007). Poorer accuracy in emotion identification has also been correlated with late and smaller evoked potentials (O’Connor, Hamm & Kirk, 2005). The under-development of the amygdala in the brain of people with ASD has been posited to have a negative effect on the growth and activation ability of neural regions associated with face processing (fusiform face area) and social perception (Schultz, 2005). It has been reported that individuals with ASD lack variance in amygdala activation when viewing fearful faces of escalating emotional intensities, and instead atypically engage areas (anterior cingulate gyrus and superior temporal cortex) related to more effortful social cognitive processes, as opposed to autonomous ones (Ashwin, Baron-Cohen, Wheelwright, O’Riordan & Bullmore, 2007; Zilbovicius et al., 2006).

**Socio-emotional perception in ASD – Speech and communication**

Studies investigating verbal communicative behaviour in people with ASD have mainly focused on spoken language ability (Tager-Flusberg et al., 2009) and prosody - the minimal distinctions in spoken language that determine what emotion and intent one conveys (Crystal, 1976; Kalathottukaren, Purdy & Ballard, 2015). Although prosody related deficits have been clinically documented in ASD since the early years of work in this area (Kanner, 1943; Baltaxe & Simmons, 1985), this area started off crucially under-researched, with only
16 studies published between 1980 and 2002 (McCann & Peppé, 2003). However, with apparent increase in prevalence rates for ASD, and DSM-V’s emphasis on social-emotional reciprocity which relies heavily on spoken communication and perception and production of prosody, there has been a substantial increase in attention to this field, from both the behavioural and neurological front.

**Prosody**

The four aspects of prosody that contribute to successful communication have been categorised as 1) ‘indexical’ factors, for example raising one’s intonation at the end of a sentence to signify a question; 2) ‘affective’ factors, the feelings and attitudes behind the words of the speaker; 3) ‘grammatical’ factors, punctuation and syllabic stress to differ the meaning of words (e.g. REcord vs. reCORD); and 4) ‘pragmatic’ factors, stress within the sentence which enables the speaker to direct focus towards the intended subject (e.g. BOB wants coffee vs. Bob wants COFFEE) (Peppé, 2009). Measurement of neural responses shows that our brains demonstrate an immediate engagement in prosodic cues while processing spoken language (Steinhauer, Alter & Friederici, 1999), and a bilateral temporo-frontal network with right-hemispheric lateralisation has consistently been reported to activate in response to the perception of affective prosody (Witteman, Van Heuven & Schiller, 2012).

The perception of communicative intent is thought to depend on the manipulation of a number of acoustic parameters identified as the foundation blocks of prosodic form (Järvinen-Pasley, Peppé, King-Smith & Heaton, 2008; Peppé, 2009; Wilson & Wharton, 2006). There is general consensus that variations in fundamental frequency, voice intensity, voice quality, and temporal aspects of speech, contribute to the differential conveyance of intent and emotion (Johnstone & Scherer, 2000; Juslin & Laukka, 2001; Murray & Arnott,
Acoustic measures of fundamental frequency and intensity have been shown to be relevant across languages in determining the indexical features of a sentence (Lin, Daniell & Al Busaidi, 2013). ‘Fundamental frequency’ (F0) refers to the number of times the vocal folds open and close, perceived as the pitch of the voice during phonation. ‘Voice intensity’ is a measure of loudness, and reflects a combination of respiratory effort and phonation action. ‘Voice quality’ encompasses the richness and strength of the tone and accuracy of articulation. ‘Temporal aspects’ include measurements of the rate of speech and durations of silence between words (Juslin & Laukka, 2001). These aspects of speech can be readily quantified using computer software that is used widely for voice analysis (Mehta & Hillman, 2008).

There have been a number of attempts to establish reliable acoustic patterns specific to different spoken emotions. This has relevance to human social communication as well as artificial speech and computer-based speech recognition algorithms (Ji, Zhou & Zhu, 2015). Studies have reported that anger and happiness are in general characterised by increased mean F0, larger F0 range, and higher vocal intensity; rate of speech is slower and inter-articulation silence is longer for sadness; and increased mean F0 and articulation rates have been associated with fear (Banse & Scherer, 1996; Murray & Arnott, 1993; Yildirim et al., 2004).

Variations in contact time between the vocal folds, measured by electroglottography, result in different phonation and demonstrate a quantifiable hierarchy across eight basic emotions with a change in hierarchical order between male and female voices (Waaramma & Kankare, 2013). Alternatively, longer durations of articulation have been attributed to the conveyance of reluctance; whereas other acoustic parameters showed poor reliability between speakers (van Zyl & Hanekon, 2012). Both general (e.g. pitch contour) and small
modifications of a range of acoustic parameters form the basis of synthetic speech algorithms (Murray & Arnott, 1993).

However, in the absence of an accepted framework for acoustic aspects of affective prosody, acoustic differences are merely descriptive, and there remains little consensus on their interpretation for an objective perspective on affective prosody (Belyk & Brown, 2014). Instead of targeting specific emotions, studies have had more success in determining consistent patterns of acoustic codes along dimensions of arousal and valence, proposed by the theory of core affect (Russell, 2003), as well as the quality and authenticity of vocal expressions (Belyk & Brown, 2014; Johnstone & Scherer, 2000). For example, the International Affective Digitized Sounds (IADS) database of 111 sounds are categorised by valence, arousal, and dominance, and there are data from this database that are comparable to standardised affective visual stimuli (Stevenson & James, 2008). The Montreal Affective Voices dataset, which consist of 90 nonverbal bursts of sound, also are characterised by ratings of valence, arousal, and intensity dimensions in order to validate emotional expression (Belin, Fillion-Bilodeau & Gosselin, 2008). Some argue that this approach would be better for studying vocal affect, as it is less skewed by factors like the differential perception and production of ‘hot’ and ‘cold’ emotions (e.g. Anger) (Banse & Scherer, 1996; Johnstone & Scherer, 2000); and the various effects of acoustic change across different languages (Wang & Lee, 2014) and cultures (Jürgens, Drolet, Pirow, Scheiner & Fischer, 2013), and in-group biases (Scherer, Banse & Wallbot, 2001).

It has further been proposed that the inherent purpose of vocalising different acoustic patterns is not to reflect an inner emotional state, but to influence the emotional arousal of the listener in order to influence their responses towards the speaker (Bachorowski & Owren, 2003). From this perspective, studies recognising the significance of both the deliverer and receiver in a communicative exchange have taken to pairing acoustic analyses with
perceptual ratings from third-party listeners to evaluate emotional speech (Banse & Scherer, 1996; Belyk & Brown, 2014; Liscombe, Venditti & Hirschberg, 2003; Monnot, Orbelo, Riccardo, Sikka & Ross, 2003). There is some consensus that this combination of the objective and subjective seems of offer a more well-rounded evaluation of prosody production.

Despite all these approaches of study, prosody is still considered relatively hard to research and discuss explicitly because of the inconclusiveness surrounding it (Peppé, 2009). As a result, full knowledge about how individuals with ASD comprehend prosody, and how their apparent impairments in perceiving prosodic cues affect their prosodic expression, is still far from being established (McCann & Peppé, 2003).

**Prosody perception in ASD**

Clinical observations and research has shown that individuals with ASD demonstrate considerably more difficulties with the semantic, interpretive, and pragmatic aspects of language, compared to the linguistic aspects (Frith, 1989; Frith & Happé, 1994; Minshew et al., 1995; Tager-Flusberg, 1996). In other words, they may be able to express themselves effectively with perfectly constructed speech, but would lack the ability to convey and understand the meaning and purpose behind their own words, and those spoken by other. Studies report particularly evident social communication impairments when individuals with ASD are faced with more complex settings involving the processing of metaphors (Melogno, D’Ardia, Pinto & Levi, 2012), as well as irony, incongruent emotions, jealousy, social blunders, and other’s intentions (Kaland, Mortensen & Smith, 2011). It has been proposed that these deficits stem from an inherent inability to recognise and understand the mental states and intentions of others, namely a lack of Theory of Mind skills (McCann & Peppé, 2003; Rutherford et al., 2002; Tager-Flusberg, 1999). Prosody, which influences the meaning
of spoken communication, is a particularly strong indicator of Theory of Mind comprehension, however much more research is needed to confirm and understand their connection (Frith & Happé, 1994).

A substantial foundation of literature has described atypical prosody perception and production in children and adults with ASD (McCann & Peppé, 2003). The PEPS-C (Profiling Elements of Prosodic Systems in Children; Peppé & McCann, 2003) is a computer-based assessment tool that evaluates children’s abilities to understand, discriminate, and express different forms of prosody by manipulating factors related to indexical purposes (questions and statements), affect, and pauses and stress intonations that shift contextual focus. The PEPS-C has robust normative results from children as young as 4 years of age (Gibbon & Smyth, 2013), and has demonstrated sensitivity to impairments in stress, pitch, and affect discrimination in children with ASD (Peppé, McCann, Gibbon, O’Hare & Rutherford, 2006). High-functioning adults with ASD reportedly perform poorly in detecting differences in syllabic stress patterns, irrespective of communicative ability (Kargas, López, Morris & Reddy, 2016). With only affective prosodic cues available, children with ASD have more difficulties correctly identifying the emotions and mental states of speakers (Rutherford et al., 2002).

There is also evidence for reduced neural potentials in response to syllabic change in children with ASD, where electrophysiological results correlate with a measured behavioural preference for non-speech compared to real speech signals (Kuhl, Coffey-Corina, Padden & Dawson, 2005). A lack of significant neural activity measured in response to changes in syllables spoken in different emotional tones has also been reported for adults with ASD, with weaker potentials correlated with more severe autistic symptoms (Fan & Cheng, 2014). Another study with a similar paradigm showed delayed neural responses approximately 100-200 milliseconds after stimulus presentation in the autism group compared to TD controls.
(Korpilahti et al., 2007). However, neural associations with behavioural prosody perception results remain inconclusive. There are some reports of increased activity in inferior frontal and temporal regions reflecting more processing effort in ASD individuals (Wang et al., 2006); and alternative reports of decreased activation in similar regions correlated with poorer integration of social information (Groen et al., 2010).

On the other hand, other studies have failed to find differences in prosody processing between people with ASD and controls. Brennand and colleagues found no significant differences between ASD and TD individuals in emotion and valence identification (Brennand, Schepman & Rodway, 2011), and emotional prosody utilisation when cues were presented in the absence of context (Le Sourn-Bissaoui, Aguert, Girard, Chevreuil & Laval, 2013). It has also been reported that individuals with high-functioning ASD perceive affective prosody and lexical stress (Grossman, Bemis, Skwerer & Tager-Flusberg, 2010), grammatical pauses and indexical intonation (Chevallier, Novek, Happé & Wilson, 2009; Järvinen-Pasley et al., 2008), as well as their TD peers.

This discrepancy may be related to task complexity as individuals with ASD have greater difficulty processing prosodic cues when presented with more than one aspect of communication to focus on. For example, despite being able to perform as well as TD individuals in prosody-related tasks with single words, it has been reported that those with ASD do significantly poorer on tasks involving sentences and complex intonation (Järvinen-Pasley et al., 2008). In more complex situations, people with ASD tend to show reduced attention towards prosody-based details, instead preferring to focus on content-related features of sentences presented to them (Brooks & Ploog, 2013). Also, unlike their TD peers, when presented with syntactically ambiguous sentences, school-aged individuals with ASD did not use grammatical prosodic cues to help make sense of the sentences (Diehl, Bennetto, Watson, Gunlogson & McDonough, 2008). Given more than one aspect of language to
simultaneously process, for example affective and semantic cues, performance of participants with ASD were poorer than their TD peers (Le Sourn-Bissaoui et al., 2013; Singh & Harrow, 2014). Additionally, people with ASD are less likely to integrate their knowledge of contextual information with their prosodic perception, leading to inaccurate judgements of misleading intent such as irony (Wang, Lee, Sigman & Dapretto, 2006).

**Prosody production in ASD**

Due to the complexities associated with measuring prosody, the number of studies documenting atypical speech and prosodic patterns in individuals with ASD is considerably less than those investigating prosody perception (Shriberg et al., 2001). Nevertheless, reasonably consistent patterns of deficits have been reported in the literature.

Vocalisations with exaggerated F0, pitch range, and pitch contours have been recorded in the speech of individuals with ASD (Green & Tobin, 2009; Sharda et al., 2010). The increased pitch range does not match typical pitch variation, and third-party listeners can have difficulties rating the speech of speakers with ASD as a result of the non-conventional prosodic patterns (Nadig & Shaw, 2012). However, other studies reported that children with ASD do not differ significantly from TD peers in speech rate, loudness, or pitch, and instead only show atypical misplacement or reduction of stress across the syllables of words (McAlpine, Plexico, Plumb & Cleary, 2014). The utilisation of grammatical, pragmatic, and affective prosody has also been reported to be comparable to TD controls, with impairments solely evident for stress production in participants with ASD in another study (Paul, Augustyn, Klin & Volkmar, 2005). There is consensus in the latter two studies that stress patterns are altered in the speech of people with ASD. This is consistent with the work of Shriberg and colleagues (2001) who reported that ASD is associated with inappropriate
pause, stress, and resonance, coupled with significantly more articulation errors and speech constraints.

As a result of reduced or inaccurate use of stress and intonation patterns, individuals with HF-ASD may not accurately communicate their purpose or intention to other listeners (Fine, Bartolucci, Ginsberg & Szatmari, 1991). They also perform below competence levels on affect expression, use of indexical signals to form questions and statements, verbal phrasing, purposeful stress placement, and prosodic imitation – as assessed by PEPS-C tasks (Peppé, Cleland, Gibbon, O’Hare & Castilla, 2011). In both these studies, the subgroup identified as having Asperger’s syndrome (diagnosed pre-DSM-V) displayed fewer or even no deficits in the tasks compared to other participants with ASD. This suggests that prosody production deficits are not homogenous and may differ depending on ASD severity and other aspects of the individual’s language development (Fine et al., 1991; Peppé et al., 2011).

Atypical indexical and pitch contours spoken by an individual with ASD have also been shown to affect how an interacting psychologist would adjust their voice quality to mimic their patient’s communicative abnormalities, with acoustic cues correlating with condition severity (Bone et al., 2014). This has considerable implications for clinicians and therapists working with this population.

**Assessment of social perception in ASD**

Deficits in social perception, with regards to both faces and prosody, have been a dominant observation of health and teaching professionals, parents and caregivers. They have driven a wide range of approaches used with the aim to quantify, track, and improve the abilities of individuals with ASD. To evaluate social and emotional perception, many propose a multimodal approach usually consisting of visual and auditory material, as these are inextricably linked and co-occur in real life communication (de Gelder & Vroomen, 2000;
Swerts, 2009). Bidirectional effects have been reported, with perceptual biases observed, in both facial and vocal emotion identification tasks (de Gelder & Vroomen, 2000). Ratings of vocal affect along arousal and valence dimensions are comparable to ratings of pictures that elicit emotion, which are similar to uncontrolled physiological responses like heart rate, skin conductance, eye startle reflexes, and facial movements (Bradley & Lang, 2000). Ratings of the emotional valence and arousal of images by neurotypical people seem to be linked to physiological responses through the engagement of thalamic, frontomedial, and anterior parietal regions in the brain (Anders, Lotze, Erb, Grodd & Birbaumer, 2004). There is much overlap in the neural bases of language processing and social attention, particularly in the superior temporal region (Redcay, 2008), with dynamic audio-visual stimuli shown to successfully activate this region bilaterally (Robins, Hunyadi & Schultz, 2009).

One of the predominant theories supported by the literature is that individuals with ASD fair worse than their TD peers in the integration of audio and visual input. Experiments with the McGurk effect (McGurk & MacDonald, 1976) showed that children with ASD were significantly less likely to use visual cues, like movement of the mouth, to aid their auditory perception of syllables (Irwin, 2007), resulting in the absence of the illusionary effect, and suggesting that integrative processing in individuals with ASD is impaired. In addition, people with ASD perform worse on audio-visual integration tasks that involve human faces and voices as stimuli, yet this deficit in performance compared to TD controls is not apparent for the same tasks involving non-human stimuli (Mongillo et al., 2008). ASD-related abnormalities in the anatomy and neural activation patterns of the superior temporal region, whose function is widely implicated in the simultaneous processing of social-based auditory and visual information, have also been reported (Redcay, 2008; Stevenson, VanDerKlok, Pisoni & James, 2011). Acknowledgement of this audio-visual connection has been the basis for widely used social perception assessment tools, such as the Cambridge Mindreading face-
voice battery (CAM; Golan, Baron-Cohen & Hill, 2006), the face and voice modules of the Diagnostic Analysis of Nonverbal Accuracy – 2 (DANVA-2; Nowicki, 2004), the Wechsler Advanced Clinical Solutions (ACS) Social Perception subtest (Pearson, 2009), and the affective prosody section of the PEPS-C (Peppé & McCann, 2003).

As discussed above, individuals with ASD demonstrate behavioural perceptual deficits in processing of facial and vocal emotion, particularly for more complex tasks that require emotion identification as well as judgements of mental state (Golan et al., 2006; Golan, Baron-Cohen, Hill & Rutherford, 2007; Rutherford et al., 2002). This is supported by reports of delayed neural responses associated with slower processing speed for affective stimuli (Korpilahti et al., 2007; Lerner, McPartland & Morris, 2013), and reduced neural activation (Groen et al., 2010).

Examples of multisensory interference is also demonstrated by a correlation between increasing background auditory noise, and worsening abilities of children with ASD to integrate recorded speech with video recorded movies of speakers (Brandwein et al., 2013; Foxe et al., 2013). Performance differences between ASD and TD participants have been detected in neural responses early on in the multisensory integration process, which are hypothesised to contribute to the consequential social cognitive deficits seen behaviourally (Brandwein et al., 2013). Although, it has been reported that these impairments decrease significantly as children with ASD enter into adolescence, which has encouraging implications for early intervention and training of multisensory processing (Foxe et al., 2013). This leads on to a discussion of the auditory processing issues observed in ASD, and its direct and indirect effects on social perception.
Auditory processing and ASD

One dominating hypothesis behind the observed and measured deficits in socio-emotional processing has been in support of atypical auditory processing, characteristics of which individuals with ASD have been documented to exhibit over the last two decades (reviews; Haesen, Boets & Wagemans, 2011; O’Connor, 2012). In large samples of children with hearing loss, a large proportion will display additional global impairments in motor and communicative abilities, as well as atypical cognitive and neurological functioning (Fitzpatrick, Lambert, Whittingham & Leblanc, 2014). The most common additional diagnosis in children with hearing loss is ASD (Cupples et al., 2013), and a close look at behavioural symptoms identified in ASD and paediatric hearing loss reveals substantial overlap. Impaired auditory input contributes to the lack of motivation to communicate socially, and vice versa, and often presents complications for clinical diagnostic procedures (Camarata, 2013).

However, unlike those with hearing loss, many individuals with ASD show little or no manifestation of physiological hearing disadvantages, as measured by brain stem responses, otoacoustic emission screening of inner ear function, and acoustic reflexes (Tharpe et al., 2006). On the contrary, some individuals with ASD demonstrate comparable or enhanced perception of simple sounds, and may have pure tone behavioural thresholds that are better than normal thresholds (Tharpe et al., 2006). They may also supersede their typically developing peers in pitch discrimination and pitch categorisation tasks (Bonnel et al., 2003; Ouimet, Foster Tryfon & Hyde, 2012). However, although some children with ASD demonstrate above chance performance at making same-different judgements on both music and speech sounds, it has been shown that they do not show a typical response advantage to speech (Järvinen-Pasley & Heaton, 2007), suggesting that they may engage in a more low-
level “perceptual blanket” strategy that focuses more on identifying salient patterns, instead of the communicative significance of the auditory information.

Group differences between those with ASD and their typically developing peers emerge more prominently when incoming information becomes more complex (Ouimet et al., 2012). Performance is reduced on tasks involving sentences and complex intonation (Järvinen-Pasley et al., 2008). Individuals with ASD also display a reduced ability to take advantage of timing dips in background noise to facilitate their speech recognition abilities in noise (so called “glimpsing”) (Alcántara, Weisblatt, Moore & Bolton, 2004; Alcántara, Cope, Cope & Weisblatt, 2012), and instead have especially exacerbated auditory processing difficulties in the presence of background noise (Alcántara et al., 2004). Behavioural studies demonstrate that auditory filtering is one of the significantly poorer sensory processing factors for children with ASD (Tomcheck & Dunn, 2007), and they also underperform when asked to recall speech stimuli presented in competition with each other (Carpenter, Estrem, Crowell & Edrisinha, 2014), when prosody-based cues are in competition with content-based features (Brooks & Ploog, 2013), or when affective and semantic cues are required to be processed simultaneously (Singh & Harrow, 2014). Individuals with ASD are reportedly worse at manipulating and maintaining their attention towards auditory stimuli (Corbett & Constantine, 2006).

Electrophysiological data show that the neural responses of children with ASD when presented with speech stimuli in a quiet environment do not differ from neural responses of TD peers when presented with speech in noise, suggesting that those with ASD are engaging in what is perhaps their maximum effort to process speech in minimal noise, and the increase in noise may only serve to degrade their speech perception even more (Russo et al., 2009). In addition to reduced neural connectivity in the fusiform face area, amygdala, and hippocampus – temporal areas affiliated with face and socio-emotional processing – individuals with ASD
display decreased connectivity in parietal and occipital regions of the brain involved in the perception and integration of sensory information, similar to those with sensory processing disorder (Chang et al., 2014), which supports the presence of a more fundamental underlying deficit in the processing of incoming sensory information.

With much consistency amongst reports of diminished efficiency in the processing of temporal modulations in sound, filtering, and specialisation for auditory information with speech properties (DePape, Hall, Tillmann & Trainor, 2012), a striking overlap between ASD (and its related language, reading, and attentional problems) and auditory processing disorder (APD) has not gone unnoticed (Dawes & Bishop, 2009). APD is characterised by an impairment in auditory perception, especially in background noise, along with difficulties in understanding and following verbal instructions and messages, localising sound, and maintaining auditory focus (Dawes, Bishop, Sirimanna & Bamiou, 2008; de Wit et al., 2016). Multiple associations and societies worldwide have developed position statements for APD - from the American Speech and Hearing Association (1995-2005), to the American Audiological Association (2010), to the British Society of Audiology (2011), to the National Acoustics Laboratory in Australia (2015), and most recently the New Zealand guidelines (Keith, Purdy, Baily & Kay, 2017) which are still under development. Close to half of a large sample of school children with learning disabilities were found to have suspected APD after failing a substantial number of psychoacoustic assessments (Iliadou, Bamiou, Kaprinis, Kandylis & Kaprinis, 2009), and there are reports of those with APD displaying poorer cognitive abilities in terms of sustained attention, auditory working memory, and nonverbal intelligence (Tomlin, Dillon, Sharma & Rance, 2015). Being a relatively new addition to the International Classification of Diseases (2016, United States, code ICD-10 H93.25), the validation of APD as a clinical entity is still considered controversial. One school of thought argues that listening difficulties experienced by individuals with APD may result from
cognitive, language, and attentional problems (top-down processes), as opposed to bottom-up processes with its bases in the auditory modality (de Wit et al., 2016; Moore & Hunter, 2013). However, upon closer inspection, a dissection of the literature finds dissimilarities between attention-based and auditory processing issues, and does not support new findings that show successful interventions for APD targeting bottom-up processes like amblyaudia (Moncrieff, 2017). Furthermore, although APD is found to be comorbid with ADHD, ASD, and language and reading impairment, cognitive and attentional deficits are not unequivocally present in all diagnoses of APD (Keith et al., 2017).

Nevertheless, APD-associated deficits, like auditory filtering, have also been shown to significantly contribute to academic underachievement in children with ASD, and to aggravate other issues the children may have, like inattention, hypersensitivity, hyperactivity, and oppositional behaviour (Ashburner, Ziviani & Rodger, 2008). The investigation into a likely comorbid relationship between ASD and auditory processing disorder (APD), and how they could potentially share the same auditory management strategies, is a new and rapidly growing field of study.

**Training of social communication and auditory management in ASD**

There is some agreement that, although the ability to utilise social and cognitive factors in pragmatic communication is weakened in ASD, it is unlikely to be completely disabled (Loukusa & Moilanen, 2009). Social skills training has had an increasing presence in ASD intervention studies. More recently, with the comorbidity between APD and ASD being more evident, different methods of auditory management have also been investigated, in particular the use of remote microphone hearing aids (RMHAs).
Training of social skills and the use of computer-based technology

Improvements in the comprehension, production, and spontaneous use of language have been reported following speech and sign language interventions with children with moderate to severe ASD (Layton, 1988), and newly acquired words and gestures were retained for months after the end of the training, regardless of verbal ability. Participation in a 12-week social skills training program also yielded significant improvements in self-reported social competence and problem behaviours from adolescents with ASD, complemented by their parents’ reports that the newly attained social skills were being generalised to non-research-related situations (Tse, Strulovitch, Tagalakis, Meng & Fombonne, 2007). In addition, training prosocial behaviour has been shown to lead to increased participation and subsequent inclusion of children with ASD into mainstream classrooms, accompanied by academic, learning style, and relational improvements (Butler, 2011).

There have also been reports of behavioural training through repeated matching-to-sample tasks that resulted in the successful acquisition and generalisation of accurate emotional vocal and facial pairings (Matsuda & Yamamoto, 2013). Recurring exposure to behavioural models, as occurs in the video modelling method, has proven effective in training individuals with ASD for specific skills; for example, conversational speech about a particular topic (Charlop & Milstein, 1989), or gesture imitation (Cardon, 2013). Watching a video recording of oneself performing a particular behaviour or attaining a particular skill (video is doctored, with separate clips put together to form what looks like a successful behavioural sequence) has similar effects to watching a model (Smith, Hand & Dowrick, 2014). Apart from having a very long skills-retention period, most of the time, the newly learned abilities also exhibit extensive generalisation effects that span out to different topics,
different conversational partners (Charlop & Milstein 1989), and increased expressive communication and play (Cardon, 2013).

With influences from studies of audio-visual integration skills in ASD (e.g. Irwin, 2007; Mongillo et al., 2008) and interventions (e.g. Matsuda & Yamamoto, 2013), coupled with the exponential development of mobile technology, individuals with ASD increasingly have access to new forms of augmentative and alternative means of communication and technology based training that are portable and accessible all the time (Shane et al., 2012). The relatively recent availability of touchscreen technology provides visual support for speech and language perception and production (Shane et al., 2012), and underlies the development of hardware and software optimised for the spectrum of atypical characteristics found in the ASD population (Mejia-Figueroa & Juarez-Ramirez, 2013).

The delivery of learning material for children with ASD and other special needs via a computer-based platform has received widespread implementation (Ploog, Scharf, Nelson & Brooks, 2013). An entire database of facial, vocal, and emotion perception training tutorials and games (Baron-Cohen, Hill, Golan & Wheelwright, 2002) has been developed and made accessible to families, schools, and clinicians via DVD-ROM. The immediate effects of training children with and without ASD to attend to sentence content and prosodic intonation, have also been validated and measured in a video game format, with relatively high retention rates testifying to the engaging experimental paradigm (Ploog, Banerjee & Brooks, 2009).

Strategies to cope with daily routines and social anxiety (Mintz, Branch, March & Lerman, 2012), as well as pro social skills - such as collaboration, turn-taking, understanding others’ emotions, and sharing interests with another (Hourcade, Bullock-Rest & Hansen, 2012), were reported to have improved for students with ASD after exposure to applications and target activities available on smartphones and tablets which they used in the classroom.
In another study, perceptually challenging natural speech presented in conjunction with increasing background noise was accompanied by a speaking face (visual cue) that was integrated into an app-based training program. This yielded significantly enhanced performance on auditory speech-in-noise tasks for children with ASD (Irwin, Preston, Brancazio, D’Angelo & Turcios, 2014). Another research group demonstrated that after watching an interactive animated program everyday as part of a 4-week intervention, children with ASD improved on emotional vocabulary and recognition to a level comparable to their TD peers (Golan et al., 2010). However, results from these studies should not be generalised unassumingly, as there is minimal consistency in experimental designs between studies. Reports have been based on four case studies (Irwin et al., 2014), as well as sample groups of 18-26 individuals (Golan et al., 2010; Hourcade et al., 2012). Some researchers were stringent in affirming the children’s ASD diagnosis above a certain evaluation cut off point (Golan et al., 2010; Irwin et al., 2014), while others adopted an all-inclusive approach and had participants ranging from non-verbal, picture-communication dependent children with ASD, to highly functioning children with ASD who fully participate in mainstream schools (Hourcade et al. 2012). Furthermore, some studies only measured one group of children with ASD across multiple time points (Hourcade et al., 2012), whereas others had opted to include a TD control group, as well as an ASD no-intervention control group (Golan et al., 2010).

Nevertheless, these studies suggest that optimum outcomes may result from a combination of interactive interfaces and a structured intervention with activities that target specific behaviours (Mejía-Figueroa & Juárez-Ramírez, 2013). However, there is still much debate around assessment procedures, and intervention type and intensity (Warren, Fey & Yoder, 2007). Although there is a general consensus that learning is optimised when interventions are spread across many sessions over time, instead of amassed into one (Warren et al., 2007), reports of statistical effect sizes of different treatments lack consistency (Zeng,
Law & Lindsay, 2012). Dosages of four 1-hour sessions a week have had comparable results to lesser intervention intensities (Hargrove, 2013). The dose of treatment is highly variable across studies, and the optimal dose is likely to vary across individuals there is a multitude of relevant variables likely to impact on the outcome including the client, clinician, condition severity, and health services and policies (Baker, 2012), as well as the intervention approach and the number of targeted outcomes (Hargrove, 2013). In addition, despite the enthusiastic integration of computer- and app-based technologies into teaching, social interventions, and assessments for individuals with ASD, their efficacy and clinical impact are yet to be rigorously and scientifically established to the extent that they could replace non-technological based approaches (Ploog et al., 2013). The literature on technology based interventions and audio-visual perception of affective prosody informed the design of the computer-based training programme used in this research.

**Auditory management and training with RMHAs**

In recent years, evidence in support of using RMHAs to manage the auditory functioning of children with APD, ASD, ADHD, and other language and learning disorders, has accumulated substantially (Schafer et al., 2014). They are currently the dominant amplification treatment option that has been shown to effectively help children with APD improve in both objective and subjective measures of auditory processing, as well as evaluations of academic performance and psychosocial issues (Keith & Purdy, 2014). However, other studies have also reported similar benefits for school children with hearing loss (Socklingam, Pinard, Caissie & Green, 2007), and reading delay (Purdy, Smart, Baily & Sharma, 2009), as well as improved speech-in-noise performance for stroke patients recovering from acquired auditory processing deficits (Koohi et al., 2016).
A recurring concern that underlies much of the research in this area addresses the effects of poor classroom acoustics that affect student learning – namely noise from people and equipment, sound reverberation, and the environment surrounding the school (Berg, Blair & Benson, 1996). Large scale sound-field amplification has demonstrated relatively positive effects on the classroom performance of children with and without hearing loss (Socklingam et al., 2007), with the emergence of additional enhancements when amplification is coupled with interventions directed at particular language skills and reading (Good and Gillon, 2014). However, for individuals with more specific auditory issues, the availability of a personal amplification system seems to offer more appropriate support (Schafer et al., 2014). RMHAs have been linked to significant advantages in speech perception in noisy classroom environments, academic performance, and psychosocial wellbeing for children with APD (Johnston, John, Kreisman, Hall & Crandell, 2009). RMHA-related benefits were also reported to be maintained after prolonged use, and continue even when the individual’s hearing is unaided, due to apparent auditory training effects of device use (Johnston et al., 2009; Keith & Purdy, 2014). There is evidence for reduced listening difficulties based on self-report from students with APD, their teachers, and their parents, following a RMHA trial (Smart, Purdy & Kelly, 2013).

Being a comparatively new concept, only a handful of projects have investigated the implementation of RMHAs with individuals with ASD. A study involving children with ASD, ADHD, and those diagnosed with both, revealed significant enhancements in speech recognition in noise, on-task behaviours in the classroom, and teacher-rated listening behaviours, after trialling personal frequency modulation (FM) systems (Schafer et al., 2013). Following a 6-week trial with RMHA usage for up to seven hours a day, school aged children with ASD have also demonstrated developments in auditory spatial and temporal processing, resulting in significant improvements in speech-in-noise discrimination, ease of
communication, and lessened perceived effects of background noise on their hearing (Rance et al., 2014).

In terms of pairing the use of RMHAs with auditory training, the literature is even sparser. Interventions targeting both bottom-up discriminative auditory processes, as well as top-down language processes, have been shown to be beneficial for children with APD, along with enhanced post-training performance for those groups allocated with an accompanying personal FM system (M. Sharma, Purdy & Kelly, 2012). Although studies involving computer-based auditory training have reported successes in facilitating on-task behaviours, and generalising speech perception and cognitive skills (Henshaw & Ferguson, 2013), most of the work has only been done with individuals with hearing loss. A brief attempt at providing dichotic listening training for children with ASD has revealed improvements in language and auditory processing tasks (Denman, Banajee & Hurley, 2015), yet it remains a challenging feat to engage children with ASD in auditory training, and has not been attempted with RMHAs.

Based on the promising results from studies evaluating the effectiveness of RMHAs for children with auditory processing difficulties, we embarked on this study with the aim to help close the gap in this field of work, specifically for children with ASD. We set out to combine computer-based emotion training intervention with RMHA-use, to hopefully target and enhance prosody processing and classroom listening experiences for students with ASD.
Chapter 3. Documentation of the development of training material and the implementation of remote microphone hearing aids for the study
The development of the computer-based emotion training sessions

Each of the nine training sessions were programmed by the researcher on Microsoft Excel. Each session had three parts to be completed; namely “Which Emotion Is It?”, “Match the Emotion”, and “Build an Emotion”. For the first part, participants had to choose the correct facial expression out of three available options that matched a written emotion. There were 30 to complete for each session. The second part required participants to listen to a recorded speech sample and choose the correct matching facial expression out of three. There were 12 to complete for each session. Within the second part, there was also another task that required participants to listen to two recorded speech samples, and choose whether they matched or not in underlying emotion (forced choice task). There were again 12 to complete for each session. For the third part, participants were presented with randomised facial parts (eyes, nose, and mouth), and were required to put the pieces together to form four facial expressions. The tasks were kept in the same order, and all the participants completed the nine sessions in sequence.

Figure 1 depicts how the tasks were presented to the participants, and how the programming for each training session was organised. Coloured textboxes and images were assigned different macro instructions so that they would lead to different parts of the session. Participants were given immediate feedback as to whether they answered correctly. This came as a visual presentation of a large tick with “Correct!”, or a cartoon face with “Oops, no”. There was no audio pairing to the feedback, because the researcher did not want to risk the participants becoming fixated on a particular sound, which could then, for example, lead to deliberately answering wrongly just to hear the sound. Although some of the visual and auditory material were repeated in different sessions, there were no instances where the correct pairing of a facial expression with a speech sample, or two affectively-matched speech samples, were duplicated. Participants were allowed to re-try an item that they got
wrong, and were instructed to complete the tasks at their own pace. There was no set delay between trials, as the computer-based tasks were not pre-programmed as automatic presentation sequences.

The material for the training sessions was gathered from several different sources. Images of faces with different emotional expressions were sampled from the Mindreading Emotions Library (Baron-Cohen et al., 2002). The researcher purchased the CD version of the assessment tool that included various picture and video files. These were sorted into folders labelled with different emotions. The researcher selected 20 different characters of both genders and from a wide range of ages, each portraying a variety of targeted and alternative facial expressions. Targeted emotions that corresponded with a correct answer in the tasks were angry, afraid, annoyed, disgusted, happy, neutral, sad, surprised, and worried. A randomly sampled selection of alternative facial expressions was used as the other available options in the forced-choice tasks.

Speech material for part 2 of the training sessions was recorded from nine speakers of both genders (4 male, 5 female), across different ages (ranges from 14 – 65 years). The speakers were asked to say different sentences, sampled from the Bamford-Kowal-Bench (BKB) sentence lists (Bench, Kowal & Bamford, 1979), in six emotions – angry, afraid, happy, neutral, sad, and surprised. The researcher selected the sentences with the criteria that they had little, if no, emotional content. For example; “children like strawberries” would not have been selected; whereas “the house had nine rooms” would. Each speaker was given a list of 30 sentences to say in the different emotions, which totalled 180 speech samples recorded for each speaker. These samples were not put through a rating process with third-party listeners. The researcher selected the samples to be included in part 2 of all the nine training sessions. Approximately equal samples were used from each speaker for each emotion.
The objective of these tasks was not to provide adaptive training. Participants’ responses were not registered, nor did the sessions increase in level of difficulty. The decided approach of training was more exposure-based, as the goal was to show the participants a variety of ways different people would convey the same emotion. Training via repeated exposure aims for the sensitisation of stimuli, and has been shown to be effective in, for example, auditory frequency discrimination of identical tones, which improved following repeated exposure at a level that is neither too easy nor too difficult to perceive (Amitay, Irwin & Moore, 2006). The authors reported learning effects that were equal to those demonstrated by participants who underwent adaptive training. It was implied that the salience of a stimulus increases as a result of modality- and dimension-specific attention and general arousal of interest.

These nine training sessions formed a novel set of tasks for the participants, who will not have encountered any of the material from the assessments, or previous computer and iPad games and applications. For copyright purposes, the sessions were solely used for the purpose of this research, and there are no plans to disseminate the training material publicly.

As mentioned above, treatment dosage likely to vary across individuals depending on the client, clinician, condition severity, and health services and policies (Baker, 2012), as well as the intervention approach and the number of targeted outcomes (Hargrove, 2013). Nine sessions across 3 weeks was chosen for this study as the researcher endeavoured to find balance between pragmatic feasibility and consistency with other training studies. Golan and colleagues (2000) administered a passive intervention that required participants with ASD to watch at least 3 episodes of an interactive program for 4 weeks. A more active paradigm saw Irwin and colleagues (2014) asking children with ASD to complete 10 minutes (2 blocks) of speech-in-noise adaptive training, 3 times per week for 12 weeks. This study had similar factors to both examples. The intervention was active, in that the participants with HF-ASD
had to spend 20-30 minutes completing one training session on the computer. However, it was not adaptive, in that the training sessions did not increase in difficulty and the main purpose was exposure-based training for the children. With the additional factor of the researcher needing to schedule 3 home visits per participant per week, often staggering 3 children at any one time, the chosen dosage of training in this study was considered to be optimal given all circumstances.

Both Rance and colleagues (2014), and Schafer and colleagues (2016) implemented 6-week RMHA trial programs (daily device usage) with children with ASD. Having only implemented half the time of RMHA exposure compared to other studies, may have contributed to the evident, but not strong, advantages of receiving additional input from the hearing devices. This is acknowledged as a limitation to the study and warrants revisions to the study design in future studies.
Figure 1. Depiction of computer-training program, task presentation and navigation
Remote microphone hearing aids (RMHAs)

RMHAs are a form of hearing technology that focus on ameliorating hearing experiences in noise, by capturing speech signals close to the source (through a microphone with a transmitter) and wirelessly sending the speech to receivers worn by the recipient of the communicative exchange. This direct transmission aims to overcome background noise, distance, and poor room acoustics to deliver optimum audio information to the listener. For this study, hearing technology manufacturers – Phonak, provided the researcher with their Roger wireless devices to trial with a sample of children with HF-ASD. Users of Roger transmitting technology paired with hearing aids (Thibodeau, 2010; Thibodeau, 2014) as well as cochlear implants (Wolfe et al., 2013), have demonstrated significantly better performances at speech-in-noise sentence recognition tasks, especially when background noise was simultaneously presented at high levels (80 dB), compared to users of older technology like dynamic or traditional frequency modulating (FM) transmitting devices. Study participants who trialled the three types of RMHA technology (participants were blinded to what they were using) in four different real-life noisy environments also reported that they preferred the Roger devices over dynamic and traditional FM systems (Thibodeau, 2014).

According to the Phonak product reference guide, the Roger Inspiro transmitter has built into it an audio chip with proprietary programming that wirelessly sends signals at a 2.4 GHz (giga-Hertz) frequency band to connected receivers in approximately 17 milliseconds. It has a potential transmission range of 20 meters, compared to older FM technologies which can transmit over 30 meters in distance. The Roger system’s more moderate range partially reduces the absorption of audio signals by obstructing animate and inanimate bodies. However, body absorption levels are still a problem in comparison to short range hearing
aids, and limiting the appropriate situations where the technology can be optimally used. So far, the Roger Inspiro has been deemed most appropriate for classroom or lecture use, with the lapel microphone designed specifically for educational settings. Due to its fixed 2.4 GHz frequency band, users of the Roger system are also less likely to experience transmission interference from overlapping frequencies, and do not need to be concerned about constantly managing the frequency band of their own devices, which was a problem with the older FM technologies.

In addition, the Roger system has an in-built dynamic noise-adaptation function called EasyGain. Depending on immediate noise levels in the environment, the transmitter is capable of measuring the signal-to-noise ratio (SNR), and automatically adjusting how much signal gain is required to produce optimum audio information to be sent to the connected receivers. This feedback process repeatedly assesses the noise in the environment, and signal gain adjustments are constantly being updated. For this study, the Roger Inspiro transmitter was paired with two Roger Focus receivers, made specifically available by Phonak for individuals with APD, ASD, and somatic symptom disorder (SSD).

**Verification procedure for RMHA fitting**

The verification procedure was provided to the researcher by Phonak (NZ). This was developed based on the American Academy of Audiology hearing assistance technologies guidelines (2008), which were further validated by Schafer and colleagues in a 2014 study with normal hearing children, recommending that real-ear measures were conducted to ensure appropriate SNR gains were met. This procedure was done in the same way for each ear.

Firstly, the researcher used an otoscope to check for any blockages in the ear canal. Once a clear ear canal was confirmed, a probe tube was inserted into the child’s ear. A real ear unaided response (REUR) at a 65 dB SPL stimulus presentation level was recorded on the
AudioScan Verifit system. The corresponding Roger Focus receiver was then inserted into the ear, but left turned off. A real ear aided response (REAR) at the same stimulus level was recorded, and checked to make sure the two electro-acoustic contours more or less overlapped with each other, to indicate minimal hearing disruption due to outer and middle ear obstructions.

The Roger Inspiro transmitter was then switched on and set to verification mode. The lapel microphone was aligned with the reference microphone inside a sealed testing box. The volume on the receiver was adjusted to maximum levels, and an electro-acoustic contour was recorded from a maximum power output (MPO) stimulus presentation. The researcher then asked the child if it was too loud, and adjusted the volume on the receiver accordingly until it was at a comfortable level. If the receiver volume had to be adjusted, another MPO stimulus was presented, and an updated electro-acoustic result was recorded. All participants had their transmitters set to EasyGain to avoid inter-user confound, and to make use of the automatic SNR gain adjustment technology that adjusts to varying classroom noise levels. The researcher then presented two speech sounds produced from the FM chest level, and checked to make sure both electro-acoustic contours did not exceed the MPO one. Key lock was activated on the transmitter and the receivers to prevent anyone from changing the settings while they handled the system.

**Familiarisation procedure for RMHA system (adapted from Schafer et al., 2013)**

With permission from Dr. E. Schafer, from the University of North Texas, an adapted version of her “educational period” was implemented in this study. During this period, the children participants with HF-ASD were given the opportunity to be familiarised with wearing and using the RMHA system, and were briefed on what the intervention would
entail. The familiarisation period took up to one week to ensure that the participants were comfortable and confident with using the RMHA.

Day 1 involved the researcher visiting the participant at home, bringing along an AudioScan Verifit system, and one Roger Focus and Inspiro RMHA set (2 receivers and 1 transmitter). The researcher measured the child’s ear and attached appropriately sized tubing and ear insertion cones onto the two receivers. Following that, the participants underwent a verification procedure for their fitted RMHAs, which is detailed below in the next section.

Day 2 involved the researcher returning to the participant’s home. Firstly, the children were asked to complete the student version of the LIFE questionnaire (LIFE-NZ; Purdy et al., 2009; adapted for New Zealand school children - including 7 questions addressing listening in noise, listening in quiet, and focused listening, and NZ normative results for comparison). They were then encouraged to read two ‘social stories’ together with the researcher – one about “wearing my Roger system”, and the other about “how my Roger system is helping me”. See Appendices 23 and 24 for the stories and accompanying images of the RMHAs that were presented to the children). Lastly, the participants were shown a demonstration and introduction video made by the researcher. The production process of the video is detailed in a separate section below. After watching the video, the children proceeded to practice putting on the two receivers, initially with help from the researcher, and then by themselves under supervision. The researcher made sure that the child could successfully put them on and take them off unaided at least five times before the day was concluded.

For days 3-5 of the familiarisation period, the participants were left with the RMHAs and spare batteries, and were encouraged to use them as much as possible at home before the start of the trial the following week. During those days, the researcher also tried to make an appointment to see the child’s teacher. For all the appointments, the researcher visited the
teachers at school at a time that suited them. A spare RMHA set was brought along to these meetings, and the researcher demonstrated how to operate and wear the transmitter. The teachers were also shown how to reconnect the child’s receivers to the transmitter in the event that the signal was cut off, how to clip the lapel microphone onto their shirt collars, and how to charge the transmitter (although the children were given the responsibility to charge it at the end of each day).

At the meeting, the teachers were also asked to complete the teacher’s version of the LIFE questionnaire. They were presented with both the revised version from the United States (LIFE-R; Anderson, Smaldino, & Spangler, 2011), as well as the UK version (LIFE-UK; Canning, 1999) which has separate questionnaires for before- and after- amplification device implementation, and an opportunity for more detailed reporting of observed classroom behaviour (Appendices 20 and 21). At the pre-intervention meeting, the teachers were given one copy of the LIFE-R and the first half of the LIFE-UK, and were asked to return a scanned copy to the researcher by the end of the week. Post-intervention, the researcher did not meet the teachers face-to-face again, but electronically sent another copy of the LIFE-R and the second half of the LIFE-UK to them for completion.

Production of the demonstration and introduction video to RMHAs

With support from SoundSkills (Auckland, NZ), the researcher managed to recruit one boy and two girls with APD, aged 10-12 years, who were users of the Phonak Roger system, to take part in the production of an introductory video that was shown to the study participants with ASD. Both the child and their primary caregiver consented to being filmed for this video, and for their first names to be used. Building from Schafer and colleagues’ (2013) RMHA familiarisation process, it was believed that having a video model of similar aged children using the hearing devices will help those with ASD acclimatise better to the
RMHAs and the subsequent 3-week trial. The video begins with scripted segments on how to wear and operate the system, as well as an overview of how they should be trialled at school and at home while doing the emotion training exercises with the researcher. For the full script, refer to Appendix 25. The video concludes with the Roger system users providing feedback about their experiences with the hearing devices and how it has helped them. The video is structured into individual sections that focus on conveying different information, with ample transition times between each one consisting of a section title and some music. The total video length is 8 ½ minutes.

The boy and one of the girls took part in the filming at the University of Auckland in a mock classroom set up. After introducing themselves and their Roger systems, one actor demonstrated how to put the receiver on, while the other explained the different parts of the system and how the transmitter and receivers operated. They then elaborated on how the system could be used in class and why it was beneficial. This was followed by a reminder to take the receivers off before going out to play at lunch time, as well as bringing the whole system home every day after school. The other girl was filmed in her home, with the scene set up to demonstrate what it will look like when the researcher visits participants at home to conduct the computer-based training exercises. After introducing herself, the actress once again went through how to wear the receivers and operate the system. A repetition of this procedure hoped to have the effect of reassuring the children with ASD, in case they became anxious about not being able to remember all the steps after they watched it the first time.

The actress then elaborated on how the hearing devices would be used in conjunction with the nine emotion training sessions, which the researcher will deliver three times a week for the next three weeks. The microphone from the transmitter was clipped on as close as possible to the built-in speakers of the laptop to enhance the salience of the speech samples from the tasks. Upon review, it was suggested by Phonak (NZ) that the speech signals
projected from the laptop would have been better controlled if the transmitter was connected directly via an audio cable, instead of clipping on the lapel microphone. However, the required equipment was not readily available, and due to the time constraints of the study, the researcher decided to move forward without this. However, this suggestion was acknowledged and would be implemented in future trials.

Following on from the instructional sections, the video proceeded to feature a compilation of comments from the three actors in response to the questions: 1) what was it like before you got your Roger system; 2) what has improved for you since you started wearing it; 3) is it hard or uncomfortable to wear; and 4) what do your friends think about your Roger system? Answers to the first question generally revolved around not being able to focus in class, understand teacher’s instructions, and requiring multiple points of clarification from the teacher before being able to do the work. The three children all reported doing better academically and being more efficient at tasks after they started using RMHAs. One also commented on how it helped her “hear better when the air conditioning or the projector is on” in the classroom. Two of the children reported that they experienced a slight discomfort during the first few days of wearing the hearing devices. However, this was reduced significantly after a few days and they subsequently said that it “felt like a part of them” and “could not even feel it” was in their ear. Finally, all three children recounted positive experiences of wearing their RMHAs in the presence of their friends, which hopefully had the effect of reassuring the children with ASD that they should not be reluctant to wear the hearing devices in fear of being teased by their peers.
Chapter 4. Assessing the effects of a computer-based training intervention, and a trial of remote microphone hearing aids, with children with high-functioning Autism Spectrum Disorder: A study of prosody perception and production, and listening experiences in the classroom

This chapter is written with the intention to submit it as a manuscript for publication.
Introduction

The persistence, frequency, and disturbance of social communication deficits is one of the main factors that determines a diagnosis of ASD based on the latest 5th edition of the DSM (DSM-V; APA, 2013). Impairments in social communication and interaction have been a consistent symptomatic marker of ASD. Since Kanner’s (1943) first production of clinical documentation, there has been much emphasis on social aspects that set ASD apart from other conditions associated with communicative and learning difficulties (Schopler & Mesibov, 1985).

Social cognition research has shown that individuals with ASD differ significantly from their TD peers in the way they utilise context to interpret and infer meaning from what others are communicating (Loukusa & Moilanen, 2009). In terms of visual cues, studies have shown that individuals with ASD demonstrate deficits in processing facial identity and expression, finding it more difficult to differentiate between faces if their emotional expressions are similar (Robel et al., 2004). It has been proposed that increased attention towards individual details (Cléry et al., 2013) results in a lack of holistic processing and an inability to coherently perceive faces and facial expressions (Gross, 2004). Children with ASD are significantly slower than TD controls at recognising basic emotions (anger, disgust, fear, happiness, sadness, and surprise; Ekman, 1970; Ekman & Friesen, 1971), and in one study the severity of their symptoms and deviating eye gaze patterns was correlated with their accuracy emotion recognition (Bal et al., 2010).

In terms of verbal communication, a considerable body of literature has also demonstrated that individuals with ASD exhibit more difficulties with the semantic, interpretive, and pragmatic aspects of language, compared to the syntactic aspects (Frith, 1989; Frith & Happé, 1994; Kaland et al., 2011; McCann & Peppé, 2003; Melogno et al.,
A number of studies have focused on socio-emotional awareness through spoken language, and have investigated how children with ASD perceive prosody (McCann & Peppé, 2003; Peppé et al., 2006). Prosody refers to the small distinctions in spoken language that determine what emotion and intent one conveys (Crystal, 1976). Prosodic cues consist of intonation, syllabic stress, grammatical stress, and affective factors, which when manipulated can generate multiple meanings from the words that are spoken (Peppé, 2009). The main acoustic parameters that form the bases of prosodic variation include ‘fundamental frequency’ (F0) perceived as the pitch of the voice during phonation; ‘voice intensity’ which is a measure of loudness; ‘voice quality’ which encompasses the richness and strength of the tone and accuracy of articulation; and ‘temporal aspects’ which include measurements of the rate of speech and durations of silence between words (Juslin & Laukka, 2001). Arguably the most ambiguous prosodic factor to quantify is the affective one, even though emotion is integral to human social communication. Researchers are yet to reach a consensus regarding how to define emotion, whether it be categorically through various patterns (Banse & Scherer, 1996; Murray & Arnott, 1993; Yildirim et al., 2004), or along subjective dimensions of valence and arousal (Belyk & Brown, 2014; Johnstone & Scherer, 2000; Russell, 2003). This makes it hard to implement consistently in studies, and is likely to be part of the reason why the literature in this area remains relatively sparse (McCann & Peppé, 2003).

Prosody–related deficits in ASD have been clinically documented for many years (Baltau & Simmons, 1985), but have not been consistently demonstrated in the literature. On one hand, impairments in stress (Kargas et al., 2016), pitch, and affect discrimination and identification (Peppé et al., 2006; Peppé et al., 2007, Rutherford et al., 2002) have been documented both behaviourally and neurophysiologically (Fan & Cheng, 2014; Korpilahti et al., 2007; Kuhl et al., 2005) in the ASD population. They have more difficulties, compared to
TD controls, at vocal-affect naming, and vocal-facial affect matching tasks (Boucher et al., 2000); as well as correctly identifying the emotions and mental states of speakers solely by listening to their vocalisations (Rutherford et al., 2002). There are additional reports of social communication impairments in more complex settings, such as those involving the processing of metaphors (Melogno et al., 2012), as well as irony, incongruent emotions, jealousy, social blunders, and others’ intentions (Kaland et al., 2011).

On the other hand, individuals with high-functioning (HF) ASD have also been documented to be as good as TD controls in perceiving affective prosody and lexical stress in sentences (Grossman et al., 2010), as well as grammatical pauses and indexical intonation (Chevallier et al., 2009; Järvinen-Pasley et al., 2008). Other studies have also found that they display no significant deficits in emotion and valence identification (Brennand et al., 2011), or emotional prosody cue utilisation (Le Sourn-Bissaoui et al., 2013).

Difficulties in accurately perceiving prosodic information is thought to contribute to deficits in prosody production. Individuals with ASD display reduced use of relevant patterns that enhance the purpose or intention of the communication (Fine et al., 1991). Most commonly observed are misplacements or reduction of grammatical and pragmatic stress on syllables and words, in both young children (McAlpine et al., 2014) and adolescents and young adults with ASD (Paul et al., 2005). Sue Peppé (2011) showed that individuals with HF-ASD scored the lowest on the affective expression and contrastive stress sections on the PEPS-C (Profiling Elements of Prosody in Speech-Communication), and did significantly poorer on all the expressive prosody tasks, compared to TD controls matched for lexical mental age, as well as chronological age. Along with unusual pause and stress patterns, individuals with ASD also display inappropriate loudness, pitch, and resonance, (Shriberg et al., 2001). In addition, inappropriate pitch exaggeration and atypical pitch contours have been
recorded in children with ASD speaking in English (Nadig & Shaw, 2012), as well as other native languages like Hebrew (Green & Tobin, 2009).

Perceptual, and subsequently production, difficulties are particularly prominent when individuals with ASD are presented with more than one aspect of communication to attend to. Performance is reduced when prosody-based cues are in competition with content-based features (Brooks & Ploog, 2013), or when affective and semantic cues are required to be processed simultaneously (Singh & Harrow, 2014). They also do significantly poorer on tasks involving sentences and complex intonation (Järvinen-Pasley et al., 2008).

A relatively new concept proposes that individuals with ASD may experience auditory processing impairments which underlie their problems with prosody and socio-emotional awareness (Haesen et al., 2011; O’Connor, 2012). Similar to those with auditory processing disorder (APD), there is little or no manifestation of physiological hearing disadvantages in many with ASD, as measured by brain stem responses, otoacoustic emission screening of inner ear function, and acoustic reflexes (Tharpe et al., 2006). Auditory processing difficulties in the presence of background noise have been reported however (Alcántara et al., 2004) as well as difficulties recalling speech stimuli presented in competition with each other (Carpenter et al., 2014), which are common features of APD. Amongst a range of sensory processing difficulties experienced by those with ASD, auditory filtering has also been reported to be one of the most affected factors (Tomcheck & Dunn, 2007).

Electrophysiological data show that the neural responses of children with ASD when presented with speech stimuli in a quiet environment does not differ from neural responses of TD peers when presented with speech in noise, suggesting that those with ASD are engaging in what is perhaps their maximum effort to process speech in minimal noise, and the increase
in noise may only serve to degrade their speech perception even more (Russo et al., 2009). Typical APD-associated deficits, such as understanding and following verbal instructions and messages, localising sound, and maintaining focus (Dawes et al., 2008; de Wit et al., 2016), have also been shown to significantly contribute to academic underachievement and the aggravation of anxiety, hypersensitivity, hyperactivity, and oppositional behavioural issues in children with ASD (Ashburner et al., 2008).

Although the ability to harness the pragmatic uses of language may seem dysfunctional in some children with ASD, it has been argued that it is not disabled and can be ameliorated with training and intervention (Loukusa & Moilanen, 2009). Following behavioural training on matching auditory affective prosody with visual facial expressions, children with ASD have been reported to show the successful generalising of their newly acquired knowledge (Matsuda & Yamamoto, 2013). Improvements in social competence and problem behaviours, and the generalisation of improved skills outside of the research setting, have also been demonstrated after adolescents with ASD and their parents participated in 12 weeks of social skills training (Tse et al., 2007).

The rapid development of computer-based technology has given individuals with ASD increasing access to new forms of visual support for speech and language perception and production (Shane et al., 2012), making it an ideal tool to support language and communication treatments with individuals with ASD (Ploog et al., 2013). Interventions have focused on integrating audio and visual stimuli, to compose a realistic delivery of multisensory information (Irwin, 2007), with influence from methods such as video modelling that consistently show significant benefits to learning novel behaviour and communication skills (Cardon, 2013; Charlop & Milstein, 1989; Smith et al., 2014). Golan and colleagues (2010) also demonstrated that children with ASD improved significantly on emotional vocabulary and recognition, after watching an interactive animated program as part
of a four-week intervention. Pro social skills - such as collaboration, turn-taking, understanding other’s emotions, and sharing interests with another, were also enhanced after exposure to activities delivered via a touchscreen mobile device (Hourcade et al., 2012). Findings seemed to suggest that optimum outcomes result from a combination of interactive interfaces and a structured intervention with activities that target specific behaviours (Mejía-Figueroa & Juárez-Ramírez, 2013; Mintz et al., 2012), but this field still lacks empirical support from research (Ploog et al., 2013).

Training through an iPad app, that integrated speech, faces, and background noise, resulted in improved performance on auditory speech-in-noise tasks (Irwin et al., 2014). Dichotic listening training (Denman et al., 2015) has also been attempted with children with ASD, and resulted in a substantial degree of within-subject improvement. However, such work would have benefited from the inclusion of a comparison no-training control group.

Personal amplification systems, such as RMHAs, that improve the signal to noise ratio in the classroom have been trialled with children with APD (Johnston et al., 2009; Keith & Purdy, 2014; Reynolds, Miller-Kuhaneck & Pfeiffer, 2016; Schafer et al., 2014; Smart et al., 2013), as well as in children with ASD (Rance et al., 2014; Schafer et al., 2013). The general consensus across these studies is that the use of RMHAs improves academic performance in noisy classroom environments, speech recognition, listening difficulties, and auditory spatial and temporal processing. Studies that have explored the simultaneous implementation of auditory training with RMHA use are scarce. A few studies involving participants with APD have reported intervention benefits for both bottom-up discriminative auditory processes and top-down language processes (Sharma et al., 2012), as well as speech perception and generalised cognitive skills (Henshaw & Ferguson, 2013). Sharma et al. (2012) found greater benefits for phonological processing when auditory or language training were combined with RMHA use at school. Currently, there does not appear to be any
literature on combining prosody perception training with RMHAs to improve socio-emotional awareness in individuals with ASD.

This study explored the feasibility of implementing a computer-based emotion training program, either on its own, or in conjunction with a home and classroom trial of a RMHA system with children with HF-ASD. A non-randomised group design was used. The “computer group” of children with ASD received an intervention of 9 sessions of computer-based emotion training activities that spanned 3 weeks (behavioural data reported in previous study). A second group received the same computer-based emotion training but also wore RMHAs when completing the computer activities (the “RMHA group”) at home. In addition, this group used the RMHAs at school during the 3-week intervention period. All participants were evaluated on prosody perception and production. For the RMHA group, outcome measures included self-report and teacher-assessed listening experiences and behaviours in the classroom. Prosody perception and production were compared to results from a group of TD children matched for age range. The study aimed to address the following research questions:

1) What are the effects of the computer-based training with/without RMHA trial on affective facial recognition, and affective face-voice matching task performance, for participants with HF-ASD?

2) What are the effects of the computer-based training with/without RMHA trial on the acoustic properties of affective prosodic productions of participants with HF-ASD?

3) What are the effects of the computer-based training with/without RMHA trial on third-party raters’ judgements of affective prosodic productions from the participants with ASD, in terms of conveyed emotion accuracy and intensity?

4) Are there any correlations between acoustic analyses and raters’ judgements of affective speech? Does this differ between TD and HF-ASD children?
5) Do any of the above results differ significantly between participants with ASD who did the emotion training with and without an accompanying RMHA system?

6) Are there any differences in performance from either/both groups compared to TD controls?

7) Are there differences in the child’s self-report of listening experience in the classroom after trialling the RMHA?

8) Are there differences in the teacher’s report of the student’s listening behaviour in the classroom after trialling the RMHA?

Method

Study design

A prospective longitudinal study design was used to compare treatment outcomes for two groups of children with HF-ASD. Participants with HF-ASD from the computer and RMHA groups were assessed four times, twice before and twice after an intervention period of three weeks. Session 1 involved an introduction to the study for the children and their family members who would be involved in the study. The parents or caregivers completed evaluations of autism symptoms and communication behaviours of the participants, while they took part in a series of social and prosody perception assessments, as well as the recording of a prosody production task. This served as a record of the children’s baseline performance. Two weeks later, Session 2 involved the re-assessment of emotion perception and production using the same measures, and served as a baseline reference point.

The HF-ASD children from the computer group proceeded straight to the 3-week intervention. On the other hand, the RMHA group spent a week prior to the intervention period getting familiarised with the RMHA system. During this week (details of which have been described in Chapter 3), the children had the device fitted and verified for their
individual use. They also received various aids, including social stories, step by step instructions (Appendices 23 and 24), and a demonstration video.

During the 3-week intervention period, all HF-ASD children participated in completing nine computer-based sessions. These were 20-30 minutes in duration, and occurred three times a week. The activities were all completed in the presence of the researcher. Apart from the first session which required task introduction and explanation by the researcher, the ensuing sessions were all self-navigated by the children. Participants in the RMHA group wore their RMHA systems while they were engaged in these training sessions with the researcher.

Upon agreement with each child from the RMHA group to trial the hearing system at school, the researcher consulted their teacher to brief them about the study and to go through how to use the device. The support of parents and teachers was obtained to ensure the child brought the RMHAs to and from school every day, and wore them during class time. Both children and teachers completed the Listening Inventory for Education (LIFE) questionnaire (separate versions for students and teachers) before and after the intervention period to evaluate changes in listening behaviours in the classroom. Out of the 12 children from the RMHA group, one did not trial the system at school. This was in part due to difficult behaviour from the child, and in part due to the reluctance of the teacher to support the child through the trial. This child only used the RMHA system during the computer training sessions.

Session 3 took place the week immediately after the 3-week intervention period ended to evaluate the effects of the interventions, again using the behavioural emotion perception and production measures. Session 4 occurred two weeks later to evaluate the retention of
intervention-related effects. All assessment and training sessions for the HF-ASD participants took place in a quiet room at their homes.

A third group of TD control children were measured on emotion perception and production on one occasion. This session was conducted in a quiet room during a visit to the university. Noted as a limitation of the study design, TD children were not tested on four occasions as per the testing schedule for the children with HF-ASD, and were not considered as a full control group. Instead their main purpose was to provide data that was used for cross-sectional comparison. The researcher was wary of potential ceiling effects, as well as reduced participant retention, if TD volunteers were required to commit to four identical assessment sessions with no purposeful intervention, as the number of recruited participants was already very small.

All participants were assessed during after school hours, either around 3:00-4:00 pm or later in the evening from 5:00-6:00 pm. The researcher endeavoured to schedule two participant visits per day. Training sessions with the HF-ASD groups also occurred during that time of day. With the objective of implementing the intervention and gathering the data as efficiently as possible, the researcher had to cater to the varying schedules of each child, and so there was little control over which days of the week the sessions took place. In doing that, though, training and assessment sessions were scheduled randomly for all participants. This reduced the potential bias of one participant group being tested at the beginning of the school week, with an advantage over another group being tested at the end of the week, who might have been more fatigued.

**Participants**

Participants with HF-ASD came to the study with an existing diagnosis previously made by an individual or multidisciplinary team of health care practitioners. The computer
group consisted of 12 children with HF-ASD. There were 11 males and 1 female ($M_{age}=9.50$ years, $SD_{age}=1.83$ years). Two of the children in this group had comorbid diagnoses of ADHD. The RMHA group consisted of a different group of 12 children with HF-ASD, with 9 males and 3 females ($M_{age}=9.91$ years, $SD_{age}=2.23$ years). Two children in this group had comorbid diagnoses of dyslexia (they were twins), and one child had ADHD. The TD group consisted of 14 children with no current or previous diagnosis of any clinical disorders, as confirmed by their parent/caregiver. There were 4 males and 10 females ($M_{age}=9.43$ years, $SD_{age}=1.87$ years).

Due to pragmatic difficulties in recruiting participants, no volunteers were excluded. The researcher also did not match groups for gender, IQ, or socioeconomic status (SES), but acknowledges that it is a limitation and that it would be important to include these factors in future extensions of this work. Gender differences in emotion perception have yet been defined with conclusive evidence. Belin and colleagues (2008) reported higher accuracy when female listeners were asked to rate female vocalisations from the Montreal Affective Voice Battery. The lowest hit rate occurred for male listeners judging male vocalisations, yet mixed gender ratings produced similar results. There are additional reports of a lack of gender differences in measures of social responsiveness (May, Cornish & Rinehart, 2016) in neuro-typical people; and in a study involving adults with and without ASD, no significant gender differences were found in measures of emotion (facial, vocal, body movements) or social cognition (Philip et al., 2010). The participants in this study were all born in New Zealand, and therefore would have been exposed to verbal cues, and prosodic intonation and expression specific to New Zealand – English. Relatively similar levels of IQ were assumed for the children with HF-ASD compared to the TD children, because they were all enrolled in mainstream school. Two children with HF-ASD required teacher aides to assist them in a group setting (not one-on-one assistance), but otherwise, there was a strong indication that
participants with and without HF-ASD were managing well in class with peers of the same age. A medium to high level of SES was also assumed for the participants in this study, based on the observation that the public schools that they attended (which requires living within the school’s zone) had the resources to cater for students with learning difficulties.

All participants passed an evoked otoacoustic emissions (OAE) screen measured via distortion product using a Grason-Stadler GSI Audioscreener (version 3.21), which indicated that they did not have significant middle ear pathology or damage to the outer hair cells in the cochlea that would be associated with peripheral hearing loss. OAE testing is a recommended hearing screening approach for children (Nozza et al., 1997; Yin et al., 2009). The passing criteria was a signal-to-noise ratio (SNR) above 6 dB across five frequency levels, from 2 – 6 kHz. Group average OAE results are presented in Table 2. As has been reported in the literature (e.g. Sininger & Cone-Wesson, 2004) OAE amplitudes were generally larger in the left ear across the three groups.
Table 2. 
Distortion product optoacoustic emission (DPOAE) signal-to-noise ratios (SNR) measured in decibels (dB) from TD and HF-ASD children (computer and RMHA groups). DPOAEs were collected from the left and right ears of each participant, across five frequency levels ranging from 2-6 kHz.

<table>
<thead>
<tr>
<th>Frequency (kHz)</th>
<th>TD Controls SNR (dB)</th>
<th>HF-ASD Computer SNR (dB)</th>
<th>HF-ASD RMHA SNR (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Left ear</td>
<td>Right ear</td>
<td>Left ear</td>
</tr>
<tr>
<td>2</td>
<td>16.60</td>
<td>15.70</td>
<td>15.74</td>
</tr>
<tr>
<td>3</td>
<td>23.22</td>
<td>19.25</td>
<td>18.23</td>
</tr>
<tr>
<td>4</td>
<td>16.94</td>
<td>15.09</td>
<td>14.90</td>
</tr>
<tr>
<td>5</td>
<td>12.85</td>
<td>23.39</td>
<td>16.04</td>
</tr>
<tr>
<td>6</td>
<td>25.88</td>
<td>22.68</td>
<td>16.40</td>
</tr>
</tbody>
</table>
Caregivers of the HF-ASD participants completed the second edition of the Childhood Autism Rating Scale (CARS-2-QPC; Schopler et al., 2010). The parents and caregivers’ comments were scored with reference to guidelines from the CARS-2-HF (high functioning version). Standardised T-scores are derived from the raw total, and correspond to the following categories: between <20 and 41 indicates that individuals scored in the minimal-no symptoms of ASD range, between 42 and 50 indicates mild-to-moderate symptoms, and 51 and higher indicates severe symptoms of ASD (see Appendix 15). Eight of the 12 children obtained standardised T-scores in the minimal severity group ($M=35.50, SD=2.62$), and four of the children scored in the mild-to-moderate severity group ($M=45.00, SD=2.45$).

Parents/caregivers of all children (TD and HF-ASD) completed the second edition of the Childhood Communication Checklist (CCC-2, Bishop, 2003). The CCC-2 screens for communication problems in children aged 4 to 16 years. Seventy questions make up 10 subscales that assess language structure, vocabulary, discourse, pragmatics (social rules of language), and impaired communicative behaviours commonly displayed by children with ASD. A General Communication Composite (GCC), scaled to individual age groups, indicates whether children may have clinically significant communication issues. In Bishop’s (2003) standardisation data, all children with ASD (and specific language impairment) scored below 55. A Social Interaction Deviance Composite (SIDC), calculated separately, indicates whether an individual child may show a communicative profile that is characteristic of ASD. SIDC values below 0 are most commonly seen in children with autism (Bishop, 2003). Table 3 reports group average, minimum, and maximum CARS-2 scores for the HF-ASD children, and GCC and SIDC scores from the CCC-2 for all TD and HF-ASD participants.

Within the computer group, one outlying individual had a GCC score (59) that exceeded 55, however this participant had the lowest (-22) SIDC score, which reflected the
difficulties he had with oversharing his interests with others, and not acknowledging other people’s perspectives. Four HF-ASD participants (two from each group – computer and RMHA) who did not obtain negative value SIDC scores had relatively low GCC scores (ranging from 14-25). The one TD control child who scored below 0 on the SIDC was home-schooled, and although had a more reserved personality, she scored a very high GCC score (95) which suggested minimal or no communication problems.

Table 3. Group average, minimum, and maximum scores for Childhood Autism Rating Scale (CARS-2) from HF-ASD participants; and General Communication Composite (GCC) and Social Interaction Deviance Composite (SIDC) scores from the Childhood Communication Checklist (CCC-2), from all participants.

<table>
<thead>
<tr>
<th>Participant Group</th>
<th>CARS-2</th>
<th></th>
<th></th>
<th></th>
<th>GCC</th>
<th></th>
<th></th>
<th></th>
<th>SIDC</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ave</td>
<td>Min</td>
<td>Max</td>
<td>Ave</td>
<td>Min</td>
<td>Max</td>
<td>Ave</td>
<td>Min</td>
<td>Max</td>
<td>Ave</td>
<td>Min</td>
<td>Max</td>
</tr>
<tr>
<td>HF-ASD (Computer group)</td>
<td>38.67</td>
<td>33.0</td>
<td>48.0</td>
<td>41.17</td>
<td>23.0</td>
<td>59.0</td>
<td>-7.17</td>
<td>-22.0</td>
<td>4.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HF-ASD (RMHA group)</td>
<td>37.00</td>
<td>29.0</td>
<td>49.0</td>
<td>38.42</td>
<td>14.0</td>
<td>54.0</td>
<td>-8.50</td>
<td>-17.0</td>
<td>5.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TD (Control group)</td>
<td>66.36</td>
<td>53.0</td>
<td>97.0</td>
<td>1.86</td>
<td>-2.0</td>
<td>10.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Assessment Tasks

Affective social and prosody perception

Social perception was assessed using the Wechsler Advanced Clinical Solutions (ACS) Social Perception Subtests (Pearson, 2009). There were three subtests in total.
The ‘affect naming’ task required the participant to identify the emotion that is being expressed on a series of 24 facial photos. They were given a choice out of the following emotions: happy, angry, sad, afraid, surprised, disgusted, and neutral.

The ‘social perception face matching’ task required the participant to listen to an audio recording of a statement, and then select one facial photo out of six that they thought matched the emotional tone behind what was said. There was no need to verbally identify the emotion; participants were instructed to disregard the gender of the voice and the photos, and to focus on the affective facial expressions. There were cases where vocal-gender was incompatible with facial-gender. There were 12 items in this subtest.

The ‘social perception pair matching’ task followed a similar procedure as above, but instead of single faces, each item was accompanied by four photos depicting a scenario with two people. The participant chose one option out of the four, depending on which one they thought matched the emotional tone in the audio recording. Participants were asked to use the characters’ facial expressions and body language to inform their decision. Once again, some items out of the 12 had incompatible vocal- and facial-gender.

The Affect Naming task was scored using the ACS guidelines (Pearson, 2009). Raw scores were used because published normative data are not available for individuals younger than age 16 years (Pearson, 2009). Scores from both the Face- and Pair-matching tasks were added together to produce a ‘Social Perception Prosody’ score.

Average internal consistency across ACS scores is reported as $r = 0.69–0.81$, with test–retest reliability reported as a corrected coefficient of $r = 0.60–0.70$, and inter-rater agreement between 98 and 99% (Pearson, 2009). The Social Perception Subtests were administered to adults with HF-ASD, Asperger’s syndrome, and a typically developing control group in a study by Holdnack, Goldstein & Drozdick (2011). They reported that those
with HF-ASD scored significant worse than controls \((F_{(4, 26)}=8.71, p<.001)\) on overall performance, but that the effect was only evident for social perception face- and prosody-matching tasks \((F_{(6, 98)}=3.92, p<.001)\), and not Affect Naming. These results support the view that the ACS Social Perception Subtests are a valid measure to be used with clinical populations such as ASD.

**Affective prosody production**

The measurement of prosody production was done by recording the participant saying “It is eleven o’clock” (Juslin & Laukka, 2001) in three different emotional tones – angry, happy, and sad. This sentence was used by Juslin and Laukka because of its neutral meaning, in order to minimise potential contextual interpretations. The children were asked to repeat the sentence four times per each emotion, resulting in 12 speech samples. This was done in a randomised order, and they were prompted by the researcher as to which emotion was required before each verbalisation.

The sentences were recorded (44.1 kHz sampling rate, 16 bit) for subsequent acoustic analysis and perceptual rating. The children wore a condenser microphone (AKG C520, MicroMic, Harman International, Austria), with a headset that reached around behind their neck and hooked onto both ears. The microphone was placed at a distance of 5 centimetres from the corner of the mouth, and was connected to a pre-amplifier (M-Audio Mobile Pre USB Interface). The amplifier was connected to a laptop computer (DELL Latitude), and the samples were recorded using Adobe Audition CS6 software.

**Listening Inventory of Education questionnaire**

Children with HF-ASD from the RMHA group who agreed to trial the systems in school for the duration of the intervention \(n=11\) were asked to complete the student version
of the Listening Inventory for Education questionnaire (LIFE-NZ; Purdy et al., 2009; adapted for New Zealand school children) which included 7 questions addressing listening in noise, listening in quiet, and focused listening in the classroom. They did this twice, before and after the three-week intervention period.

The teachers of these children were also asked to complete the teacher version of the LIFE questionnaire (LIFE-R; Anderson et al., 2011), which evaluated classroom listening behaviours of the child. They also completed the questionnaire twice, before and after the intervention. Teachers also completed an additional section of an alternative version of the LIFE questionnaire (LIFE-UK; Canning, 1999) that evaluates the efficacy of implementing an amplification device for a particular individual. Copies of the questionnaires are included in Appendices 20 and 21.

The student version of the LIFE-NZ questionnaire has been validated with typically developing children and children with APD. The seven items in the questionnaire, which consist of 3 identified key factors (listening in noise, listening in quiet, and focused listening) account for 71.8% of the variance in children in the control group, 66.3% of the variance in children with APD, and demonstrated acceptable internal reliability with a Cronbach’s alpha result of $\alpha=.072$ (Purdy, Sharma & Morgan, 2017). Test-retest reliability results of the three key factors are also available for typically developing children with no learning difficulties (Purdy et al., 2009) – with a significant correlation between the two assessment sessions (Spearman’s $r=.63-.74, p<.05$), and no significant difference between repeated ratings ($F_{(1,17)}<.01, p=.92$).

**Editing speech recordings and generating acoustic values**

Each participant had a continuous speech recording per assessment session. The Multi-Dimensional Voice Program (MDVP) from the Sona-Speech II software
(KayPENTAX) was used to edit these recordings into individual segments, as well as cutting out any external noise and prompting from the researcher. Altogether, each child had 48 individual files, which were then acoustically analysed through MDVP. The acoustic measures selected for statistical analysis included: the mean fundamental frequency (Mean F0 Hz), frequency range in Hertz and semitones (Range Hz; Range Semitones), minimum frequency (Min Hz), maximum frequency (Max Hz), standard deviation of frequency in Hertz and semitones (SD Hz; SD Semitone), variance in fundamental frequency (vF0), mean intensity in decibels (Mean Energy dB), intensity range in decibels (Range Energy dB), minimum intensity (Min Energy dB), maximum intensity (Max Energy dB), and standard deviation of intensity in decibels (SD Energy dB).

Raters’ judgements of the speech samples

All the audio files were compiled into one folder, randomised, and renamed as track numbers. Three students from the Speech Science department at the University of Auckland volunteered to rate the speech samples of the participants. They were first trained simultaneously on example recordings that were not part of the actual set. Explanations were given with regards to how each section of the response was to be completed. For each sample, the raters had to indicate which emotion they thought was being conveyed (angry, happy, or sad); and the degree of emotional intensity they thought the utterance had (0=minimal intensity, 1=mild, 2=moderate, 3=intense).

This rating procedure was sourced from work by Ross and colleagues (1997) who developed an assessment battery of prosody production tasks. After completing the training, the raters each wore identical bilateral headphones (Sony MDR7506) and rated all the speech samples from the three participant groups. The raters were blinded to which group the
speaker belonged to, what emotion was intended, and which assessment session the recording
was from.

Although the sentence “it is eleven o’clock” was not from Ross and colleagues’
Aprosodia Battery (1997), the method of acquiring the speech samples and subsequent
acoustic and subjective evaluations of prosody production, has been validated in neuro-
typical control adult participants, patients with aphasia resulting from neurological damage,
patients with Alzheimer’s disease (AD), and patients with Parkinson’s disease (PD). The
subjective method of rating affective speech samples were highly correlated acoustic
measures of variance in fundamental frequency (F0), as reported by Ross and colleagues
(1997) – emotion identification had a correlation co-efficient of \( r = 0.681 (p < .001) \), and emotion
intensity had a correlation co-efficient of \( r = 0.764 (p < .001) \), across both neuro-typical and
neurologically impaired participants. In addition, the measurement of spontaneous affective
prosodic production (as opposed to repeating/imitating example affective utterances) was
significantly correlated with poorer spontaneous speech fluency in patients with right and left
hemispheric brain damage, assessed as part of the Western Aphasia Battery \( (r = 0.703, p = .023) \).
Testa, Beatty, Gleason, Orbelo and Ross (2001) reported that patients with AD demonstrated
significantly less variance in F0 compared to healthy controls, only when asked to imitate
affective productions of sentences \( (F_{1,45} = 10.46, p < .002) \), but not for repetitions of
monosyllabic or dysyllabic utterances. Furthermore, Buxton (2011) reported that raters were
less accurate in judging the emotions conveyed through the speech of people with PD
compared to healthy controls \( (U = 295.0, p = .019) \). The application of this method has yielded
results in line with hypothesised differences between clinical and typical sample groups, and
along with its validation with a range of patient populations, indicates that it is an acceptable
measure of affective prosody production.
Rater reliability

For the accuracy of participants’ affective prosody production based on raters’ judgements: 1) repeated recordings (n=50) were included in the list of audio files. The three raters were not aware of this, and so proceeded with rating all the speech samples. After completing the ratings, the answers for the original 50 recordings were compared to the answers for the repeated 50 recordings, so that intra-rater reliability could be calculated. Intra-rater reliability across the three raters ranged from 78-86%. 2) to determine inter-rater reliability, ‘rater consensus’ was first coded as agreement on emotion type (2/3 or 3/3 of the raters agreed on the perceived emotion). This was very high, at 96%. For the 4% of samples where consensus was not reached, rater 2’s results were used, as rater 2 had the highest intra-rater reliability. The agreed perceived emotions were then compared to the intended emotions, and whether they matched or didn’t match is referred to as ‘accuracy’.

For the intensity of participants’ affective prosody production based on raters’ judgements: a Kappa (K) measure of inter-rater agreement was used to analyse the judgements for degree of perceived emotional intensity. K ranged from .305 to .401 (p<.001). According to Landis and Koch (1977), K values within the range of .21 and .40 represent a fair agreement. Based on these agreement results, judgements for ‘degree’ were averaged across the three raters to derive one rating of intensity for each speech sample. Degree of rated emotional intensity was then further analysed with respect to group, as well as pre- and post-intervention differences. The heterogeneity of vocalisations from the HF-ASD participants, as well as the subjective nature of rating audio files, may have contributed to the ‘fair’ agreement on emotional intensity. Although this warrants improvement in controlling for inter-rater subjectivity, this effect arguably reflects the reality of how individuals perceive
expressive intensity differently, and is also buffered by high inter-rater reliability with respect to accuracy of emotion identification.

**Computer-based training tasks**

The researcher created nine computer-based training sessions that were delivered via Microsoft Excel on a Dell Latitude E6420 laptop. ‘Buttons’ were programmed by recording a macro pathway between various sites in the document to resemble an interactive quiz. Each session took approximately 20 minutes to complete.

Images of faces with different emotional expressions were sampled from the Mindreading face-voice battery, developed by Golan and colleagues in 2006. The purchased CD had various picture and video files available. Speech material was recorded from nine New Zealand speakers of different ages and gender. The speakers were asked to say different sentences, sampled from the Bamford-Kowal-Bench (BKB) sentence lists (Bench et al., 1979), in six emotions – angry, afraid, happy, neutral, sad, and surprised. Combined, these resources formed a novel set of activities for the participants, who will not have previously encountered any of the same material from the assessments, nor other computer and iPad games and applications.

Each training session consisted of four different activities. Participants had to 1) choose a correct facial expression, out of three, that matched the emotion stated; 2) listen to a speech sample and choose a correct facial expression, out of three, that matched the emotion conveyed; 3) listen to two speech samples and decide whether the underlying emotions matched or not; and 4) put together eyes, nose, and mouth images arranged as a puzzle, to build four facial expressions of different emotions. The four activities remained constant across all nine training sessions, with the only difference being the images and speech materials used.
Statistical analyses

Statistical analyses were done using IBM SPSS Statistics, v20.0 (Armonk, NY, USA). Within subject repeated-measures ANOVAs were conducted within each of the two HF-ASD groups (computer and RMHA) to explore the effects of the intervention on 1) affective facial recognition, and affective face-voice matching task performance; 2) the acoustic properties of affective prosodic productions; and 3) on third-party raters’ judgements of accuracy and intensity of affective prosodic productions, across the four assessment sessions. Alpha values for post hoc comparisons were Bonferroni-corrected.

One-way ANOVAs were conducted on each outcome measure, at each assessment point, to see whether there were any group differences between TD controls, HF-ASD computer intervention group, and HF-ASD RMHA intervention group. Post hoc comparisons were Bonferroni corrected.

Pearson correlations were conducted between acoustic factors and raters’ judgements of accuracy, within each participant group. The p-values from significant correlations were not corrected for multiple comparisons, as each factor was analysed separately. The aim of this comparison was to see whether there were relationships between particular acoustic parameters and raters’ judgements of emotional expression and to determine whether these relationships differed between TD and HF-ASD groups.

Lastly, pair-samples t-tests were conducted on data from the HF-ASD RMHA group to evaluate whether LIFE scores from the children and their teachers changed significantly post RMHA trial.
Results

Affective social and prosody perception

Within subjects repeated measures ANOVAs revealed that for the Affect Naming task (Figure 2) which involved facial expression emotion identification, there were statistically significant differences in performance across the four assessment sessions for both the HF-ASD computer group ($F_{(3, 33)}=16.02, p<.001$), and the RMHA group ($F_{(3, 33)}=66.96, p<.001$). A similar pattern of results were obtained for the Social Perception Prosody derived score (Figure 3), which involved matching facial expressions to affective speech, for which there were also statistically significant changes over time for the computer group ($F_{(3, 33)}=33.55, p<.001$), as well as the RMHA group ($F_{(3, 33)}=37.45, p<.001$). Post-hoc analyses (Table 4, with a Bonferroni-corrected significance value of $p\leq.008$, showed that significant differences were mostly between pre- and post-intervention sessions, and not between the baseline (1 and 2) and the follow up (3 and 4) sessions. This suggests that before completing the intervention there was a stable baseline, and that the improvements observed after the intervention were retained during the 2-week follow-up period. The stable baseline indicates that the participants did not demonstrate a test-retest practice effect for the task, which supports the use of the prosody tasks to evaluate performance before and after intervention for determining treatment effects.
Table 4. 
Post hoc analyses results for the Affect Naming and Social Perception Prosody scores from the ACS Social Perception subtests. Means (M) and standard deviations (SD) are shown in parentheses, and p-values are shown for each comparison. Session A and B indicate the two sessions in each pairwise comparison. * Bonferroni corrected p<.008 statistical significance.

<table>
<thead>
<tr>
<th>Participant Group</th>
<th>Session A (M, SD)</th>
<th>Session B (M, SD)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Affect Naming</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HF-ASD Computer</td>
<td>1 (16.25, 2.99)</td>
<td>2 (17.33, 4.01)</td>
<td>.800</td>
</tr>
<tr>
<td></td>
<td>1 (21.50, 1.31)</td>
<td>&lt;.001 *</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1 (21.00, 1.81)</td>
<td>.001 *</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2 (4)</td>
<td>.036</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2 (4)</td>
<td>.074</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3 (4)</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>HF-ASD RMHA</td>
<td>1 (15.33, 2.93)</td>
<td>2 (15.42, 1.93)</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>1 (21.25, 0.87)</td>
<td>&lt;.001 *</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1 (21.42, 1.31)</td>
<td>&lt;.001 *</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2 (3)</td>
<td>&lt;.001 *</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2 (4)</td>
<td>&lt;.001 *</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3 (4)</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td><strong>Social Perception Prosody</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HF-ASD Computer</td>
<td>1 (11.92, 3.99)</td>
<td>2 (12.83, 4.57)</td>
<td>1.00</td>
</tr>
<tr>
<td></td>
<td>1 (19.42, 0.90)</td>
<td>&lt;.001 *</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1 (18.50, 2.28)</td>
<td>&lt;.001 *</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2 (3)</td>
<td>.001 *</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2 (4)</td>
<td>.002 *</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3 (4)</td>
<td>.914</td>
<td></td>
</tr>
<tr>
<td>HF-ASD RMHA</td>
<td>1 (13.25, 3.52)</td>
<td>2 (14.25, 2.99)</td>
<td>.710</td>
</tr>
<tr>
<td></td>
<td>1 (19.50, 1.24)</td>
<td>&lt;.001 *</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1 (19.33, 1.56)</td>
<td>&lt;.001 *</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2 (3)</td>
<td>&lt;.001 *</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2 (4)</td>
<td>&lt;.001 *</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3 (4)</td>
<td>1.00</td>
<td></td>
</tr>
</tbody>
</table>
Figure 2. Group average scores from the Affect Naming task of the ACS Social Perception subtests. The score for TD control group obtained for one test occasion only is shown as a dashed line. Scores are shown for the four assessment sessions for the HF-ASD computer and RMHA groups (Sessions 1 and 2 for pre-intervention and Sessions 3 and 4 for post-intervention).

Figure 3. Group average scores from the Social Perception Prosody tasks (sum of Face- and Pair-matching tasks) of the ACS Social Perception subtests. The score for TD control group obtained for one test occasion only is shown as a dashed line. Scores are shown for the four assessment sessions for the HF-ASD computer and RMHA groups (Sessions 1 and 2 for pre-intervention and Sessions 3 and 4 for post-intervention).
One-way ANOVAs revealed that there were significant between group effects at each assessment session for both tasks, except for Social Perception Prosody measured at Session 3, and Affect Naming measured at Session 4. Table 5 reports significant differences between TD children, HF-ASD children with computer-only intervention, and HF-ASD children with RMHA intervention. The one-off assessment scores from the TD children were included in the comparison for each session. Patterns of results are relatively similar for the two HF-ASD groups, when compared to their TD peers. Although not statistically significant following Bonferroni corrections to post hoc p values (p<.017), the RMHA group’s performance at Session 4 for Social Perception Prosody was 0.01, which is essentially the same, albeit fractionally higher than the TD group. This differs from the computer group’s performance, which drops at follow-up, and is not significantly different from the TD group. This trending result suggests that the combined intervention of computer-based training and RMHA may be more effective in prolonging improvements in prosody perception. However, this will need to be investigated further with larger sample sizes per group, as well as the inclusion of an RMHA-only participant group.
Table 5.
One-way ANOVA and post hoc analyses results for the Affect Naming and Social Perception Prosody scores. F-statistics (degrees of freedom with and without error), and p-values for between group effects are reported. Means (M) and standard deviations (SD) are shown for each participant group, followed by p-values for post hoc comparisons at each session. Participant groups are coded 1 for TD children, 2 for HF-ASD Computer, and 3 for HF-ASD RMHA.

### Affect Naming

<table>
<thead>
<tr>
<th>Participant Group</th>
<th>Session 1</th>
<th>Session 2</th>
<th>Session 3</th>
<th>Session 4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M (SD)</td>
<td>p</td>
<td>M (SD)</td>
<td>p</td>
</tr>
<tr>
<td>TD (1)</td>
<td>19.29 (2.23)</td>
<td>1 v. 2 = .022</td>
<td>19.29 (2.23)</td>
<td>1 v. 2 = .268</td>
</tr>
<tr>
<td>HF-ASD Comp (2)</td>
<td>16.25 (2.99)</td>
<td>2 v. 3 = .111</td>
<td>17.33 (4.01)</td>
<td>2 v. 3 = .322</td>
</tr>
<tr>
<td>HF-ASD RMHA (3)</td>
<td>15.33 (2.93)</td>
<td>3 vs.1 = .002*</td>
<td>15.42 (1.93)</td>
<td>3 vs.1 = .004*</td>
</tr>
</tbody>
</table>

### Social Perception Prosody

<table>
<thead>
<tr>
<th>Participant Group</th>
<th>Session 1</th>
<th>Session 2</th>
<th>Session 3</th>
<th>Session 4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M (SD)</td>
<td>p</td>
<td>M (SD)</td>
<td>p</td>
</tr>
<tr>
<td>TD (1)</td>
<td>18.79 (2.29)</td>
<td>1 v. 2 &lt; .001*</td>
<td>18.79 (2.29)</td>
<td>1 v. 2 &lt; .001*</td>
</tr>
<tr>
<td>HF-ASD Comp (2)</td>
<td>11.92 (3.99)</td>
<td>2 v. 3 = .984</td>
<td>12.83 (4.57)</td>
<td>2 v. 3 = .928</td>
</tr>
<tr>
<td>HF-ASD RMHA (3)</td>
<td>13.25 (3.52)</td>
<td>3 vs.1 = &lt;.001*</td>
<td>14.25 (2.99)</td>
<td>3 vs.1 = .005*</td>
</tr>
</tbody>
</table>

Comp = Computer group; * Bonferroni corrected p<.017 statistical significance
Affective prosody production: Acoustic analyses

Within subjects repeated measures ANOVAs showed no significant changes across assessment sessions for any acoustic parameter derived from the prosody production speech samples of the HF-ASD children from the computer group. For those from the RMHA group, analyses showed a positive trend for a pre vs. post measurement of range in fundamental frequency (Hz) for the angry emotion. However, small sample sizes of all participant groups likely contributed to lessened statistical effects and results may reflect low statistical power.

One-way ANOVAs revealed significant differences between groups but did not show systematic differences across the assessment sessions. Table 6 lists the significant statistical results for the between group comparisons of acoustic parameters across assessment sessions. Children with HF-ASD from the computer group consistently had a larger pitch range (measured in semitones) for the angry emotion, compared to the TD controls. They also consistently exhibited smaller measurements of minimum pitch (Hz), and larger measurements of minimum intensity, measured in decibels (dB), for happy; as well as larger measurements in minimum intensity (dB) for sad, compared to their TD peers. Furthermore, children from the computer group consistently demonstrated higher measures of minimum and maximum intensity (dB) for the happy emotion, and higher measures of maximum intensity (dB) for the sad emotion, compared to HF-ASD children from the RMHA group, across pre- and post-intervention sessions. The RMHA group did not differ significantly from the TD controls until the last assessment session, where they had significantly lower values for mean, minimum, and maximum intensity (dB) when they produced happy sounding speech samples.
Table 6.
One-way ANOVA and significant post hoc comparison results for between group effects on acoustic parameters measured for angry, happy, and sad across assessment sessions.

<table>
<thead>
<tr>
<th>Emotion (Acoustic parameter)</th>
<th>Session</th>
<th>$F_{(df, dfe)}$, p</th>
<th>TD $M$</th>
<th>TD $SD$</th>
<th>HF-ASD Computer $M$</th>
<th>HF-ASD Computer $SD$</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Angry (Range in semitones)</td>
<td>1</td>
<td>3.94$_{(2,33), .029}$</td>
<td>11.80</td>
<td>4.28</td>
<td>15.51</td>
<td>3.12</td>
<td>.032</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>7.21$_{(2,32), .003}$</td>
<td>11.80</td>
<td>4.28</td>
<td>17.58</td>
<td>3.87</td>
<td>.002</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>5.86$_{(2,32), .007}$</td>
<td>11.80</td>
<td>4.28</td>
<td>17.50</td>
<td>4.56</td>
<td>.006</td>
</tr>
<tr>
<td>Happy (Minimum F0 in Hz)</td>
<td>1</td>
<td>5.41$_{(2,33), .009}$</td>
<td>195.29</td>
<td>30.02</td>
<td>150.02</td>
<td>42.09</td>
<td>.008</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>6.49$_{(2,33), .004}$</td>
<td>195.29</td>
<td>30.02</td>
<td>151.93</td>
<td>31.59</td>
<td>.004</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>9.35$_{(2,33), .001}$</td>
<td>195.29</td>
<td>30.02</td>
<td>137.41</td>
<td>36.69</td>
<td>.001</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>4.80$_{(2,33), .015}$</td>
<td>195.29</td>
<td>30.02</td>
<td>153.28</td>
<td>38.33</td>
<td>.022</td>
</tr>
<tr>
<td>Happy (Minimum intensity in dB)</td>
<td>1</td>
<td>7.90$_{(2,33), .002}$</td>
<td>45.41</td>
<td>5.28</td>
<td>76.78</td>
<td>6.43</td>
<td>.029</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>6.57$_{(2,33), .004}$</td>
<td>45.41</td>
<td>5.28</td>
<td>50.67</td>
<td>2.60</td>
<td>.029</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>13.90$_{(2,33), &lt;.001}$</td>
<td>45.41</td>
<td>5.28</td>
<td>49.95</td>
<td>4.05</td>
<td>.035</td>
</tr>
<tr>
<td>Sad (Minimum F0 in Hz)</td>
<td>1</td>
<td>10.44$_{(2,33), &lt;.001}$</td>
<td>169.74</td>
<td>27.76</td>
<td>127.36</td>
<td>30.90</td>
<td>.005</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>5.55$_{(2,32), .009}$</td>
<td>169.74</td>
<td>27.76</td>
<td>128.41</td>
<td>31.08</td>
<td>.010</td>
</tr>
<tr>
<td>Sad (Minimum intensity in dB)</td>
<td>1</td>
<td>34.90$_{(2,33), &lt;.001}$</td>
<td>40.23</td>
<td>2.87</td>
<td>49.17</td>
<td>3.79</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>27.95$_{(2,33), &lt;.001}$</td>
<td>40.23</td>
<td>2.87</td>
<td>49.17</td>
<td>3.79</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>28.25$_{(2,32), &lt;.001}$</td>
<td>40.23</td>
<td>2.87</td>
<td>49.44</td>
<td>2.98</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>44.11$_{(2,33), &lt;.001}$</td>
<td>40.23</td>
<td>2.87</td>
<td>49.25</td>
<td>3.38</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>
### HF-ASD Computer vs. RMHA groups

<table>
<thead>
<tr>
<th></th>
<th>HF-ASD Computer</th>
<th></th>
<th>HF-ASD RMHA</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
<td>p</td>
</tr>
<tr>
<td>Happy (Minimum intensity in dB)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>7.90</td>
<td>.002</td>
<td>76.78</td>
<td>6.43</td>
<td>.002</td>
</tr>
<tr>
<td>2</td>
<td>5.71</td>
<td>.007</td>
<td>49.64</td>
<td>2.69</td>
<td>.008</td>
</tr>
<tr>
<td>3</td>
<td>6.57</td>
<td>.004</td>
<td>50.67</td>
<td>2.60</td>
<td>.005</td>
</tr>
<tr>
<td>4</td>
<td>13.90</td>
<td>&lt;.001</td>
<td>49.95</td>
<td>4.05</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Happy (Maximum intensity in dB)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>4.23</td>
<td>.023</td>
<td>78.29</td>
<td>5.57</td>
<td>.023</td>
</tr>
<tr>
<td>3</td>
<td>4.13</td>
<td>.025</td>
<td>80.46</td>
<td>6.27</td>
<td>.022</td>
</tr>
<tr>
<td>4</td>
<td>9.75</td>
<td>&lt;.001</td>
<td>79.76</td>
<td>6.82</td>
<td>.009</td>
</tr>
<tr>
<td>Sad (Minimum intensity in dB)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>34.90</td>
<td>&lt;.001</td>
<td>48.45</td>
<td>2.51</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>2</td>
<td>27.95</td>
<td>&lt;.001</td>
<td>49.17</td>
<td>3.79</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>3</td>
<td>28.25</td>
<td>&lt;.001</td>
<td>49.44</td>
<td>2.98</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>4</td>
<td>44.11</td>
<td>&lt;.001</td>
<td>49.25</td>
<td>3.38</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

### TD controls vs. HF-ASD RMHA group

<table>
<thead>
<tr>
<th></th>
<th>TD</th>
<th></th>
<th>HF-ASD RMHA</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
<td>p</td>
</tr>
<tr>
<td>Happy (Mean intensity in dB)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>6.53</td>
<td>.004</td>
<td>67.92</td>
<td>6.29</td>
<td>.014</td>
</tr>
<tr>
<td>Happy (Minimum intensity in dB)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>13.90</td>
<td>&lt;.001</td>
<td>45.41</td>
<td>5.28</td>
<td>.019</td>
</tr>
<tr>
<td>Happy (Maximum intensity in dB)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>9.75</td>
<td>&lt;.001</td>
<td>76.29</td>
<td>6.82</td>
<td>.009</td>
</tr>
</tbody>
</table>
**Affective prosody production: Raters’ judgements**

Within subjects repeated-measures ANOVAs revealed that third party raters judged happy-sounding speech samples from the HF-ASD computer group to have a higher degree of intensity post-intervention \(F(3, 135)=5.42, p=.002\). Post hoc comparisons between different sessions were not statistically significant following a Bonferroni adjustment to the \(p\) value \((p<.008)\). For the HF-ASD RMHA group, an overall increase in the accuracy of participants’ prosody production based on raters’ judgements (across emotions) was shown \((F(3, 342)=4.21, p=.006)\), though again there were no statistically significant post hoc comparisons. Happy-sounding speech samples from the HF-ASD RMHA group, in particular, were more accurately rated post-intervention \((F(3, 120)=3.66, p=.014)\), specifically between Sessions 1 \((M=0.68, SD=0.47)\) and 3 \((M=0.83, SD=0.38, p=.006)\).

One-way ANOVAs did not show any consistent between group differences for any emotion across assessment sessions. At Session 1, the HF-ASD computer group \((M=0.71, SD=0.46)\) received significantly less accurate judgements of their angry speech samples \((F(1,41)=5.23, p=.027)\), compared to those from the RMHA group \((M=0.91, SD=0.29, p=.019)\). At Session 2, the HF-ASD computer group \((M=1.83, SD=0.77)\) produced significantly more intense happy-sounding speech \((F(8,37)=2.25, p=.046)\), compared to TD controls \((M=1.58, SD=0.35, p=.021)\). Lastly, at Session 3 post-intervention, overall (across emotions) raters judgements of degree of emotional intensity were significantly higher \((F(8,106)=2.22, p=.032)\) for the HF-ASD RMHA group \((M=1.73, SD=0.72)\) compared to their TD peers \((M=1.50, SD=0.67, p=.012)\).

Table 7 depicts a confusion matrix that shows the percentage of the time raters confused the intended emotion with the other two alternatives. For TD children, accurate judgements of intended emotions were relatively high. Their lowest scoring emotion is sad,
which was equally confused by raters as either angry or happy. Pre-intervention, both HF-ASD groups had lower judged accuracy of emotion productions. Raters tended to misjudge angry as happy and vice versa for the HF-ASD computer group. A similar confusion is present for those from the RMHA group, but raters were also particularly inaccurate in judging sad-sounding speech samples from these children, and often mistook the intended emotion for happy. Non-parametric analyses, due to a ceiling effect from TD controls, were conducted to compare rater accuracy between TD and HF-ASD groups, as well as to compare between pre- and post-intervention rater accuracy within the HF-ASD groups. No significant differences were revealed from the analyses.
Table 7.
Percentage of time raters judged the intended emotion of the participants correctly, and when they confused it with the other two alternative emotions.

<table>
<thead>
<tr>
<th>Intended Emotion</th>
<th>Rated Emotion</th>
<th>TD Controls</th>
<th>HF-ASD Computer group</th>
<th>HF-ASD RMHA group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Rated Emotion</td>
<td>Rated Emotion (Pre-intervention)</td>
<td>Rated Emotion (Post-intervention)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Angry</td>
<td>Happy</td>
<td>Sad</td>
</tr>
<tr>
<td>TD Controls</td>
<td></td>
<td>Angry</td>
<td>97.56</td>
<td>1.83</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Happy</td>
<td>1.83</td>
<td>95.12</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sad</td>
<td>0.61</td>
<td>3.05</td>
</tr>
<tr>
<td>HF-ASD Computer group</td>
<td></td>
<td>Angry</td>
<td>90.32</td>
<td>6.09</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Happy</td>
<td>7.17</td>
<td>92.11</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sad</td>
<td>2.51</td>
<td>1.79</td>
</tr>
<tr>
<td>HF-ASD RMHA group</td>
<td></td>
<td>Angry</td>
<td>94.38</td>
<td>3.21</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Happy</td>
<td>4.82</td>
<td>95.58</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sad</td>
<td>0.80</td>
<td>1.20</td>
</tr>
</tbody>
</table>
Correlations between acoustic parameters and raters’ judgements

Table 8 highlights significant Pearson correlation results between acoustic parameters and the accuracy of raters’ judgements, for each emotion within each participant group. For individual children and emotions, the accuracy score was obtained by summing results across the four assessment sessions. In TD children, there are no significant relationships between acoustic measures and rater accuracy for angry and happy. However, the subjective rating of sad-sounding speech samples was associated with differences in pitch range and mean intensity. For HF-ASD children, a wider range of acoustic measures were correlated with the perception of angry, happy, and sad emotions in the children’s speech samples. There were relatively few correlations for the TD group, which may reflect the reduced variation in each measure for the TD children who were more consistent and more accurate in their productions.
Significant correlations between acoustic parameters and accuracy of raters’ judgements for each emotion, calculated separately for TD and HF-ASD groups.

<table>
<thead>
<tr>
<th>Emotion</th>
<th>Acoustic Parameter</th>
<th>TD Children</th>
<th>HF-ASD Children</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>r</td>
<td>p</td>
<td>r</td>
</tr>
<tr>
<td>Angry</td>
<td></td>
<td></td>
<td>Mean F0 (Hz)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Range (semitones)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SD (semitones)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Mean intensity (dB)</td>
</tr>
<tr>
<td>Happy</td>
<td>Mean F0 (Hz)</td>
<td>.64</td>
<td>.024</td>
</tr>
<tr>
<td></td>
<td>Range (Hz)</td>
<td>.72</td>
<td>.019</td>
</tr>
<tr>
<td></td>
<td>SD (Hz)</td>
<td>.63</td>
<td>.049</td>
</tr>
<tr>
<td></td>
<td>Range (semitones)</td>
<td>.64</td>
<td>.048</td>
</tr>
<tr>
<td>Sad</td>
<td>Range (Hz)</td>
<td>.64</td>
<td>.018</td>
</tr>
<tr>
<td></td>
<td>SD (Hz)</td>
<td>.59</td>
<td>.036</td>
</tr>
<tr>
<td></td>
<td>Range (semitones)</td>
<td>.79</td>
<td>.008</td>
</tr>
<tr>
<td></td>
<td>Mean intensity (dB)</td>
<td>.80</td>
<td>.002</td>
</tr>
<tr>
<td></td>
<td>Mean F0 (Hz)</td>
<td>.63</td>
<td>.030</td>
</tr>
<tr>
<td></td>
<td>Maximum F0 (Hz)</td>
<td>.81</td>
<td>.005</td>
</tr>
<tr>
<td></td>
<td>SD (semitones)</td>
<td>.61</td>
<td>.037</td>
</tr>
<tr>
<td></td>
<td>Intensity range (dB)</td>
<td>.71</td>
<td>.050</td>
</tr>
<tr>
<td></td>
<td>Minimum intensity (dB)</td>
<td>.61</td>
<td>.035</td>
</tr>
<tr>
<td></td>
<td>Maximum intensity (dB)</td>
<td>.60</td>
<td>.038</td>
</tr>
<tr>
<td></td>
<td>SD intensity (dB)</td>
<td>.62</td>
<td>.030</td>
</tr>
</tbody>
</table>

Mean fundamental frequency (F0), along with various measures of pitch range were most sensitive to emotion differences, and were associated with accuracy of subjective judgements, at least in this study’s HF-ASD sample. The judgement of sad seemed to rely heavily on variations in loudness.
**Evaluation of RMHA trial in the classroom**

Lastly, paired samples t-tests were conducted to investigate changes in the HF-ASD children’s and teacher’s report of listening experience and behaviours in the classroom after the 3-week RMHA trial at school. Figure 4 shows the pre- and post-intervention average group scores for the student version of the LIFE questionnaire. There was a significant reduction in overall listening difficulty ($M_{pre}=2.35$, $SD_{pre}=0.65$; $M_{post}=1.55$, $SD_{post}=0.43$; $t_{(10)}=3.51$, $p=.006$). One specific factor that was significantly improved was problems with listening in noise ($M_{pre}=2.97$, $SD_{pre}=0.98$; $M_{post}=1.67$, $SD_{post}=0.63$; $t_{(10)}=4.09$, $p=.002$).

![Figure 4. Pre- and post- RMHA trial at school LIFE questionnaire scores, self-reported levels of listening difficulty by HF-ASD children. The dashed lines represent mean scores for each factor of the questionnaire calculated from normative New Zealand data collected by Purdy and colleagues in 2011. Listening in noise = 2.14±0.83; Listening in quiet = 1.23±0.46; Focused listening = 1.60±0.59; Overall = 1.73±0.51; * represents $p\leq.05$ statistical significance.](image)
Figure 5 shows the pre- and post-intervention average group scores for the teacher versions of the LIFE questionnaire. Assessment of children’s listening behaviours was done with the revised-LIFE questionnaire from the USA (LIFE-R; Anderson et al., 2011), whereas the UK version (LIFE-UK; Canning, 1999) was used to provide a subjective efficacy score from the teacher evaluating the success of the RMHA trial for their particular student. Results from a paired samples t-test showed that there was a significant improvement in listening behaviours in the classroom ($M_{pre}=41.64$, $SD_{pre}=11.13$; $M_{post}=54.55$, $SD_{post}=11.19$; $t(10)=-4.08$, $p=.002$). The average improvement score for the amplification device trial was 13.7, which fell into the category of “successful intervention” based on the scoring guidelines of the questionnaire (Canning, 1999).

Figure 5. Pre- and post- RMHA trial at school LIFE questionnaire scores, reflecting the evaluation by the teacher of listening behaviours in the classroom, supporting the positive outcomes of the RMHA trial.
**Discussion**

Overall the results showed positive prospective effects of a computer-based training intervention on emotion processing in children with HFASD. There were no large-scale benefits apparent based on the addition of RMHA use during training and at school, however some small gains were evident. Prosody and facial emotion perception were specifically targeted, and changes hypothesised for this skill were achieved. Although prosody production was not addressed in the training, there were also a few small effects from both acoustic analyses and raters’ judgements. Future studies involving a larger participant group, and a longer and more intense trial period are needed to verify these findings.

**Emotion and prosody perception**

Repeated measures results showed that both the HF-ASD computer and RMHA groups improved in their abilities to identify facial expressions, as well as to match emotions from the face to affective speech. Pre-intervention, both groups scored lower than their TD peers. Children with HF-ASD in the RMHA group were significantly poorer, whereas the HF-ASD computer group had lower mean scores that did not reach statistical significance. This may reflect individual differences between the children with HF-ASD. Due to the heterogeneous nature of ASD as a condition, as well as not randomly assigning the HF-ASD participants to different intervention groups due to limited numbers and incremental recruitment success, it was difficult to control for variations between the groups.

The intervention was able to lift the performance of HF-ASD participants to surpass that of the TD controls. Due to the cross-sectional nature of the comparisons between participants with HF-ASD and the TD children, post-intervention improvements are more likely to be attributable to the training, rather than reflect generalisable effects. As mentioned, TD children were not assessed repeatedly because of potential ceiling effects, as well as
prioritising participant engagement to voluntarily remain in the study. Having had tested the TD children four times, it is likely that they would have reached, or come close to, perfect performances on the ACS social perception tasks and would still surpass the abilities of participants with HF-ASD. However, the result of interest from this study is that with exposure training to facial expressions and affective prosody through computer-based tasks, children with HF-ASD aided and unaided by RMHAs have the potential to match the abilities of their untrained TD peers. Future work could also expand the collection of normative data from TD children so that raw scores from the ACS test can be standardised for ages younger than 16 years, the current cut off point for available standardised data. All participants were able to complete the ACS tasks with some success, indicating the suitability of the measures for younger ages.

Comparable results have been reported in another training study that involved matching-to-sample tasks across modalities (visual and auditory) (Matsuda & Yamamoto, 2013). Despite the use of a relatively simple computer activity in the current study, instead of more engaging interventions like animated videos (Golan et al., 2010) or video games (Ploog et al., 2009), improvements were also evident. This suggests there may be potential benefits of implementing technology into therapy sessions, as the computer training was supported by a researcher in the current study. These findings do not suggest that therapy involving real-person interactions should be replaced by a computer game, but they do suggest that computer training may be a useful adjunct to conventional therapy. A few published reports in the literature cite the positive effects of using computer games to support therapeutic programs with children. Case study examples from Gardner in 1991 showed improvements in children’s psychotherapy outcomes as they used computer games in conjunction with other methods such as relationship building, story-telling, and expression through art. Another large scale qualitative survey questioned therapists and their patients about the applicability
and efficacy of a particular computer game used with children participating in cognitive
behavioural therapy (Brezinka, 2014). The survey revealed that the game enhanced the
motivation of the children to participate in therapy sessions, and strengthened the
relationships between clinician and patient.

Although there were no significantly different improvement effects for the HF-ASD
RMHA group compared to the computer group, it is interesting to note that there were
differences in the Social Perception Prosody score at the follow up Session 4 which occurred
approximately two weeks after the end of the intervention. This result suggests that the
retention of the performance gain after intervention exhibited by the participants in the
RMHA group was better than retention in the computer group, but further research with a
longer follow up period would be needed to confirm this trend. At the final assessment scores
from the computer group reduced and did not differ from the TD controls. Scores from the
RMHA group remained at a higher level than those from the TD group at Session 4. Other
auditory training studies with (Sharma et al., 2012) and without (Henshaw & Fergusson,
2013) RMHAs report various results, from the full retention of improvements, to post-
intervention changes only for specific measures of language processing and for specific
intervention types (Sharma et al., 2012). It would be useful to extend the follow-up period in
future studies to determine whether the improvement attributed to the computer training and
RMHA trial are sustained.

Further work in this area would benefit from modifying intervention dosage to see if
the size of the effect would change. This work could contribute towards identification of an
optimum dosage, which is currently not well defined because of the many influences of
client, clinician, condition severity, and health service and policies variables (Baker, 2012),
when examining intervention outcomes for children with ASD. In future studies, it would
also be beneficial to include an additional HF-ASD group that just wore the RMHAs without
doing computer emotion training, to pinpoint whether observed improvements in prosody perception are attributable to the computer-based tasks, the RMHAs, or the combination of both.

**Prosody production: Acoustics and raters’ judgements**

The lack of systematic differences in speech acoustics across assessment sessions and between participant groups could be attributed to highly variable vocalisations from children with HF-ASD. Other studies have reported exaggerated F0, pitch range, and pitch contours from the recorded speech of individuals on the spectrum (Green & Tobin, 2009; Sharda et al., 2010). In this study, participants from the HF-ASD computer group consistently displayed heightened variations in pitch and intensity, not specific to any particular emotion. The individuals from this group also consistently vocalised at a much higher intensity compared to HF-ASD children from the RMHA group, regardless of whether the voice recordings were collected at pre- or post-intervention assessment sessions. This difference is difficult to explain as there was no systematic difference in recruitment to the two intervention groups, but could reflect the relatively small sample size in each group.

Interestingly, HF-ASD children from the RMHA group and their TD peers only differed significantly in their speech acoustics at Session 4 post-intervention. All three significant acoustic factors were based on measures of intensity for happy-sounding vocalisations. The improved signal-to-noise ratio (Thibodeau, 2010; Thibodeau, 2014) from RMHA transmittance serves to increase the saliency of speech. A potential effect of wearing RMHAs is that children with ASD may become more aware of affective prosody produced by others, as well as themselves, which could have resulted in better self-monitoring and self-correcting of their speech. Similar effects have been documented in the video self-modelling literature, where repeated exposure to seeing oneself successfully performing a specific skill
(video footage is edited, with separate clips put together to form what looks like a successful behavioural sequence) leads to the actual mastery of the task (Smith et al., 2014).

Raters’ subjective judgements also did not reveal systematic differences across sessions and between groups. Intervention effects for both rater accuracy and degree of perceived emotional intensity were minimal. However, the confusion matrix (Table 7) highlighted that accuracy of emotional speech improved post-intervention for both the computer and RMHA groups. The accurate judgement of sad-sounding speech samples from the computer group surpassed that of the TD controls, although accuracy for angry and happy remained below that of the controls. This suggests that children with ASD are better at sounding sad than happy or angry. For the RMHA group, post-intervention raters’ judgements increased in accuracy and were only slightly below that of their TD peers across all emotions. The raters were blinded to the group and session and intended emotion when rating the speech samples. High accuracy in raters’ judgements overall may be attributed to the fact that all three raters were students from the Speech Science department, and naïve listeners may have been less reliable and less accurate in perceiving the different emotions. Although none of the raters had prior experience with the rating task, all of them received an introductory briefing and training. Future extensions of this study would benefit from having raters with no related background in the area of Speech Science or Psychology to obtain the ratings that reflect the perceptions of the general public rather than listeners trained in the perception of disordered speech.

Correlations between acoustics and raters’ judgements

Observed relationships between acoustic parameters and rater accuracy are predominantly from the results of HF-ASD children. Measures of mean F0, pitch range in Hz and semitones, and standard deviation of pitch in Hz and semitones showed significant
correlations with subjective accuracy across all emotions – angry, happy, and sad. These results corroborate other reports in the literature that anger and happiness are characterised by an increase in mean F0, F0 range, and vocal intensity (Banse & Scherer, 1996; Murray & Arnott, 1993; Yildirim et al., 2004). Sadness has been characterised by a slower rate of speech and prolonged inter-articulation silence, however in this study’s case, the correct identification of sad-sounding speech samples was driven by similar pitch parameters, as well as quite strongly by intensity factors. This might reflect the differences in speech material across studies.

There is dichotomy in the literature with regards to how emotions should be classified or characterised, with one body of research focussing on objective acoustics, and the other on subjective ratings using the dimensions of valence, arousal, and intensity (Belin et al., 2008; Belyk & Brown, 2014; Johnstone & Scherer, 2000). There is need for more studies that involve the cross examination of acoustics and listener judgements to reach a consensus on emotion, in order for other areas like synthetic speech simulation (Murray & Arnott, 1993) and speech therapy for individuals with ASD involving social communication targets to be more consistent in their approach to affective prosody.

**Trial of RMHA systems in the classroom**

Overall, the results from both the students’ and teachers’ report of trialling RMHAs at school for three weeks was positive. Both students and teachers reported a significant attenuation of listening difficulties, especially in the presence of background noise. There was a trend for difficulties with focused listening to diminish based on the self-report questionnaire; future research with a larger sample and greater statistical power is needed to verify this trend. Pre-intervention teachers’ LIFE-R scores averaged at 41.6, indicating that the teachers felt that the students often displayed listening challenges in the class. The post-
intervention average score of 54.5 meant that upon re-evaluation, teachers rated the children as only occasionally having listening challenges. The successful trial of RMHAs in the classroom mirrors results reported by Schafer and colleagues (2013) and Rance and colleagues (2014) who documented significant enhancements in speech recognition in noise, on-task behaviours in the classroom, teacher-rated listening behaviours, and ease of communication. Teachers anecdotally commented to the researcher via email that the RMHAs seemed to have had noticeable effects on students’ self-esteem and demeanour at school. Particularly for students with ASD, RMHAs seem to contribute profound benefits in reducing anxiety, aiding comprehension, and facilitating on-task behaviour. Further extensions to this study would benefit from the inclusion of auditory assessment tasks akin to those used for the diagnosis of APD, to strengthen and validate the evident benefits of RMHAs for children with ASD.

**Limitations and future directions**

A limitation of this study was the absence of a HF-ASD group that only trialled the RMHAs without receiving computer-based emotion training. Future research in this area would benefit from including such a group, so that observed effects could be more clearly attributed to either the computer training, or the RMHAs, or indeed a combination of both interventions. As mentioned above, the study design could also be improved by administering the outcome measures to the TD group at four time points, mimicking the study design administered to the children with HF-ASD, except without the intervention. This would allow for the tracking of natural learning effects as a result of repeating the task, and would again contribute to the evaluation of the “real” effects of the social perception training.

The numbers of participants in each group were too small to facilitate highly controlled selectivity and randomised allocation to different interventions, and thus limiting
the interpretation of the results that arose from the study. However, future extensions of this work would benefit from more detailed data on each participant’s IQ, and confirmation of ASD diagnosis via multiple sources, to better match the characteristics of the participant groups.

Another limitation of this study was the elicitation of potential power-of-suggestion type effect due to the leading nature of the some of the questions in the RMHA demonstration video for the children with HF-ASD. With the intention of reassuring the children, most of whom were initially anxious about the trial, the researcher did not control for the effects of the materials being used for the familiarisation procedure for the RMHAs. Changing the video for future studies would be preferable over introducing a RMHA turned-off placebo participant group, as denying potential treatment to a participant with HF-ASD would be unethical.

In line with the method of recording, analysing, and rating prosody production developed by Ross and colleagues (1997), the participants were only asked to produce emotive expressions. When the study results revealed rather large variations in these utterances, and only a fair agreement between raters with regards to emotional intensity, it could be assumed that future study designs will need to cater for the wide range of individual differences that arise as part of the heterogeneous nature of ASD. One solution would be to include a “neutral” prosody production expression, so that there is a ‘normal’ benchmark from which all subsequent emotive expressions could be evaluated.

All in all, the results of this study provide a promising foundation for further research into prosody and emotion processing training in children with ASD, and support additional trials of personal hearing amplification devices. The computer training developed for this research was associated with positive outcomes for affective prosody and hence has
implications for future development of programs for use in therapy sessions, at school, and at home to help individuals with ASD improve their social communication abilities.
Chapter 5. Cortical auditory evoked potentials and Mismatch responses
Cortical auditory evoked potentials (CAEPs) and Mismatch negativity (MMN)

Apart from investigating deficits in prosody perception, and the effects of training interventions and auditory enhancement, from a behavioural perspective – it is also important to factor in one of the most common frameworks the condition of ASD – which stems from a neurodevelopmental standpoint. These series of studies aimed to supplement the behavioural data with neurophysiological data to demonstrate the likely presence of a neurological basis to emotion perception and auditory processing difficulties observed for individuals with ASD. The aim is to also contribute to the growing field of work that strives to find an objective method to evaluate and potentially diagnose ASD. Neurophysiological approaches to measuring responses to auditory stimuli (in particular, affective speech, which individuals with ASD have difficulty perceiving) may present as a useful tool to accompany or validate subjective diagnostic decisions based on behavioural observations and clinical assessments.

CAEP components: Acquisition, interpretation, and limitations

Evoked, or event-related neural activity can be measured as fluctuations in electrical potentials associated with neural activity. In human studies, this neural activity is usually recorded via electrodes placed noninvasively on the scalp, using an electroencephalographic (EEG) method, which measures continuous changes in voltage from clusters of activated neurons from various regions in the underlying cerebral cortex (Scherg & von Cramon, 1986). Cortical electrode placement is prescribed by standardised maps, like the international 10-20 system (Jasper, 1957) that are based on average anatomical measurements of cortical areas extrapolated to the surface of the scalp.

Human evoked potentials (EPs) elicited by incoming auditory information can be recorded at three levels within the central auditory system – from the brainstem immediately after stimulus presentation; then from thalamocortical projections and the primary auditory
cortex in a middle-latency timeframe; and lastly from higher cortical areas of sensory processing in a longer-latency timeframe (Picton, Hillyard, Kraus & Galambos, 1974; Picton & Stuss, 1980; Näätänen, 1992). Whereas EPs are generally associated with the automatic responses that occur in the early to mid-latencies after delivery of an auditory stimulus, the term event-related potentials (ERPs) refers to the electrophysiological responses that occur later in the auditory pathway that are dependent on stimulus presentation, stimulus delivery manipulation, and subsequent sensory, motor, or cognitive events (Näätänen, 1992; Näätänen & Picton, 1987; Picton et al., 1974; Picton & Stuss, 1980). ERPs are commonly recorded using either an obligatory stimulus paradigm with repeated presentation of the same stimulus, or a discriminative oddball paradigm where both frequent (standard) and infrequent (deviant) stimuli are presented.

There are four main components of the auditory ERP identifiable by their polarity (positive or negative potential) and their occurrence in the sequence (Näätänen & Picton, 1987). P1 (first positive peak) is usually observable at around 50 ms post stimulus presentation. N1 (first negative trough) follows at approximately 100 ms. P2 occurs on average at 175 ms, followed by N2 observed at around 250-300 ms post stimulus presentation (Näätänen, 1992; Picton et al., 1974). N1 and P2, the most obvious deflections in the auditory ERP, were first documented as responses to sounds in the waking human brain by Pauline Davis in 1939, and subsequently have been studied intensively in terms of acoustic experimental manipulation and psychophysiological processes (Näätänen, 1992; Näätänen & Picton, 1987; Picton et al., 1974; Scherg, Vajsar & Picton, 1989). It is well documented that auditory cortical responses are sensitive to stimulus factors such as changes in pitch, intensity, stimulus presentation rate, stimulus habituation, stimulus prediction, as well as participant factors such as alertness, stages of sleep, drug-induced effects and attention (Näätänen, 1992).
Studies investigating participants in different age groups have revealed maturation-related differences and differing rates of maturation for the presence, size, and timing of the major auditory ERP components (Ponton, Eggermont, Khosla, Kwong & Don, 2002). During adulthood, latencies for auditory ERP components increase with age, measured between 20 and 88 years (N1 in parietal regions, P2 in frontal regions, N2 across the scalp) (Anderer, Semlitsch & Saletu, 1996). In newborns aged less than seven days old, Wunderlich, Cone-Wesson and Shepherd (2006) reported that the dominant observable components were P2 and N2 across frontal, central, and parietal electrode sites, with significantly larger responses to speech tokens (short word with consonant-vowel-consonant structure) compared to tone bursts. However, another report shows successful recording of an N2 component only after six months of age, and a P2 emerging later between 8 and 30 months (Shafer, Yu & Wagner, 2014). Dissimilarities in findings compared to Wunderlich et al. (2006) could be attributed to the different stimulus, which in this case consisted of a single 250 ms long vowel presentation. Variations in acoustic salience between articulated speech sounds have been shown to influence the elicitation of CAEPs with regards to component onset latency and amplitude (Ostroff, Martin & Boothroyd, 1998).

In infancy (Cone & Whitaker, 2013), and progressing to toddlerhood (Wunderlich et al., 2006), there is evidence of the emergence of a substantial P1 component, which has a peak amplitude that surpasses those recorded at both the newborn and adult stage (Wunderlich et al., 2006). P1 has been easily detected at frontal electrode sites starting from just three months of age, and is shown to increase in amplitude linearly with age progression up to 8 years (Shafer et al., 2014). Similar to P2 and N2, P1 is also dependent on the acoustic properties of the stimulus. For example, stimulus presentation level is an important consideration as there is some evidence that P1 is only elicited in infants when presentation
levels exceed 50 dB SPL (Cone & Whitaker, 2013), although other studies have recorded responses at lower levels (Purdy, Sharma, Munro & Morgan, 2013).

The last auditory ERP component to arise in young children is the N1, which has been recorded as early as 14 months of age (Shafer et al., 2014), but is most commonly observed at 3-4 years from fronto-central cortical sites (Ponton et al., 2002; Wunderlich et al., 2006). In general, however, studies have demonstrated that N1 does not present as a reliable and independent component until adulthood, unless the experiment consisted of substantially lengthened stimuli (Čeponienė, Rinne & Näätänen, 2002) or rapid rates of stimulus presentation (Sussman, Steinschneider, Gumenyuk, Grushko & Lawson, 2008), in which case N1 could be detectable in late childhood and adolescence.

ERP latencies decrease with age across childhood (Ponton et al., 2002; Sussman et al., 2008; Wunderlich et al., 2006), with an average difference of 100-150 ms between potentials recorded in infants compared to adults (Cone & Whitaker, 2013). The pattern of peak amplitude change is quite complex. There are reports of P1 and N2 amplitudes decreasing, and N1-P2 components increasing (Anderer et al., 1996; Wunderlich et al., 2006); though the rates of these changes differ, with N2 progressing much faster (50% per year) compared to the other three components (11-17% per year) (Ponton et al., 2002). In addition, the scalp distribution of the P1 and P2 components migrate from frontal to fronto-central regions, and central to fronto-central regions, respectively, as the brain matures. In contrast N1 is consistently measured as maximal at fronto-central cortical sites, and N2 at frontal sites only (Sussman et al., 2008).

It has been suggested that maturation-related variations in the ERPs reflect the neurodevelopment of the processing of auditory stimuli - progressing from bottom-up sensory perception, to top-down selectivity and attention orienting to the incoming
information (Čeponienė et al., 2002). Indeed, the latest ERP component to emerge – N1 – although readily recorded in passive listening paradigms, has also been attributed to the engagement of selective attention towards target sounds showing an increase in amplitude with attention (Coch, Sanders & Neville, 2005; Hillyard, Hink, Schwent & Picton, 1973).

Early studies of scalp-recorded late, transient ERP components recorded in response to tone bursts suggested that they were generated from the supratemporal plane across both hemispheres (Vaughan & Ritter, 1970). The involvement of the primary auditory cortices from both temporal lobes is supported by the spatio-temporal dipole source model developed by Scherg and colleagues. This approach compares a waveform representation (tangential and radial vectors) of a time-locked, spatially restricted, electrically active cortical region, with scalp recordings, to estimate the source of the generated potential (Scherg & von Cramon, 1986; Scherg et al., 1989). Radial sources have been attributed to mid-latency responses, whereas tangential sources have been linked to the later, larger components of P1, N1, and P2 (Ponton et al., 2002). Other studies indicate that, in adults, auditory N1-P2 components arise from multiple superior temporal and frontal sources, which converge at the fronto-central region of the scalp’s vertex to provide recordable waveforms with maximum amplitudes (Näätänen & Picton, 1987; Picton et al., 1974).

Due to the distance between the source locations and the scalp measurements of auditory ERPs, the EEG data acquired from scalp recordings can be affected by a range of artifacts (Näätänen, 1992). Artifacts may be especially problematic in uncooperative participants like young children, and those with neurological or behavioural problems. Excessive tension in head and neck muscles, as well as large muscle reflexes and movements produce noisy electrical activity that contaminates and elevates or depresses ERPs depending on the nature of the artifact (Näätänen, 1992). Ocular artifacts are caused by changes in potential between a positively charged cornea and a negatively charged retina (eye
movements), as well as changes in electrical charge as the eyelid slides across the eyeball (eye blinks) (Croft & Barry, 2000; Jervis, Nichols, Allen, Hudson & Johnson, 1985). Maintaining eye fixation during ERP recordings has been utilised as a common compensatory strategy, however some may argue that it imposes an additional load on the cognitive processes that are being employed, because of its counterintuitive nature in restricting the movement of the eye towards attention-grabbing visual or auditory stimuli (during attentional tasks) (Croft & Barry, 2000; Näättänen, 1992). Another common practice is to correct for ocular artifacts offline during the data analysis stage (Croft & Barry, 2000). Methods include averaging across multiple artifacts present in the continuous EEG recording to determine a threshold that encompasses most of the data without excessively rejecting too many epochs (Croft & Barry, 2000; Semlitsch, Anderer, Schuster & Presslich, 1986). Another approach involves correlating electrooculographic (EOG) measurements of eye movements (recorded as calibration trials) with the EEG data, to produce an adequate correction factor that can be applied to obtain eyeblink-adjusted ERPs (Croft & Barry, 2000; Jervis et al., 1985).

The Mismatch Negativity/Response (MMN/MMR) effect

An electrophysiological response specific to the detection of a change in incoming auditory information was first described as “a later negative shift superimposed on potentials elicited by the former stimuli” (Näättänen, Gaillard & Montysalo, 1978). This effect, named the mismatch negativity (MMN), is present in both unattended and attended conditions (Näättänen et al., 1978; Näättänen, Teder, Alho & Lavikainen, 1992; Sussman, Chen, Sussman-Fort & Dinces, 2014), and is considered to solely reflect sensory change detection between stimuli presented in sequence (Csépe, 1995; Näättänen, 1992; Sams, Paavilainen, Alho & Näättänen, 1985). It differs from an N1 enhancement, which is evident when the listener selectively attends to the source of the incoming sound while discriminating between targets and non-targets (Hillyard et al., 1973). MMNs have been observed in experimental
designs that require ignoring of the stimuli, actively engaging in perceiving stimulus change (Erlbeck, Kübler, Kotchoubey & Veser, 2014; Novak, Ritter, Vaughan & Wiznitzer, 1990), or participating in simultaneous non-related tasks such as reading or visual discrimination (Sams, Alho & Näätänen, 1984; Sams et al., 1985).

Early studies discovered the MMN by subtracting ERPs to numerous repeated “standard” stimuli, from ERPs to rarely and randomly presented “deviant” stimuli, though mismatches are not necessarily restricted to this configuration. Sams and colleagues (1984) demonstrated the presence of MMNs for a deviant preceded by a standard, and vice versa, as well as for two repeated deviants (although the amplitude was smaller). Ruusuvirta, Huotilainen and Näätänen (2008), however, proposed that the ratio of standard to deviant stimuli should be at least 3:2 in order to elicit a reliable mismatch. The effect typically occurs at the latency of an N2 auditory ERP component in adult listeners – with ERP waveforms to standard and deviant stimuli generally starting to deviate at around 100 ms, and the deviation lasting until approximately 250 ms post stimulus onset (Näätänen, 1992; Näätänen & Alho, 1995). The recorded response has been described as containing two distinct parts – the first negative wave represents the pre-attentive MMN, whereas the second negativity N2b represents the progression from subconscious to conscious perception of changes in sound and, according to early MMN researchers, only emerges in attended experimental conditions, not ignored ones (Näätänen, 1992; Näätänen, Paavilainen, Rinne & Alho, 2007; Novak et al., 1990; Sams et al., 1985). There is some evidence from more recent studies, however, that a late MMN can be recorded in a passive listening task. For example, a late discriminative negativity (LDN) that has been reported to occur at approximately 400 ms post stimulus presentation in adults, and between 400-460 ms in children (Korpilahti, Lang & Aaltonen, 1995), and is thought to reflect the processing of more complex auditory stimuli, especially
with regards to language and speech processing (Cheour, Korpilahti, Martynova & Lang, 2001; Halliday et al., 2014; Korpilahti, Krause, Holopainen & Lang, 2001.

The literature suggests that the auditory MMN predominantly arises from the auditory cortex in the supratemporal region (Alho, 1995; Näätänen & Alho, 1995; Sams et al., 1985), which is consistent with the putative cerebral generators of the auditory ERPs (Näätänen & Picton, 1987; Picton et al., 1974; Ponton et al., 2002; Vaughan & Ritter, 1970). Patients with damage to their temporal and parietal lobes demonstrate impairments in auditory discrimination and have diminished MMNs (Alain, Woods & Knight, 1998). Magnetically recorded ERPs suggest that the MMN sources depend on stimulus manipulation, emerging in different regions of the supratemporal cortex (Alho, 1995). ERPs recorded using conventional EEG (Scherg et al., 1989) suggest additional sources as the response progresses through the lateral temporal cortex and into the frontal areas (Näätänen & Alho, 1995). Different MMN time courses could potentially reflect the progression of cerebral activation, with pre-attentive change detection processes manifesting earlier in the supratemporal region, and the conscious perception and attentional redirection to novel acoustic information driven by activation of frontal regions (Alho, 1995; Näätänen & Alho, 1995).

The effects of different stimuli and attention

Studies exploring manipulations of pitch (Alho, Woods & Algazi, 1994; Näätänen, Teder, Alho & Lavikainen, 1992; Novak et al., 1990; Paavilainen et al., 1991; Tiitinen, May, Reinikainen & Näätänen, 1994; Winkler et al., 1995), intensity (Näätänen, Paavilainen, Alho, Reinikainen & Sams, 1989; Paavilainen et al., 1991), and temporal (Kujala, Kallio, Tervaniemi & Näätänen, 2001; Paavilainen et al., 1991) features, have all reported larger ERPs (more negative) recorded in response to deviant compared to standard stimuli, thus producing a MMN. For pitch-related MMNs, larger amplitudes and reduced latencies are
positively correlated with increasing frequency differences between deviant and standard stimuli (Novak et al., 1990), although it has been shown that only deviant frequencies above a particular threshold (8 Hz higher than standard) were sufficient to elicit a mismatch (Sams, Paavilainen, Alho & Näätänen, 1985). For intensity-related MMNs, amplitudes increase and latencies decrease when the deviant stimulus grows weaker in loudness compared to the standard (Näätänen et al., 1989). Kujala and colleagues (2001) also showed that the differential evocation of MMNs can be linked to the temporal manipulation of intervals between tone-pair stimuli. MMNs were elicited in response to deviant tone pairs that were separated by 20 ms and 60 ms, compared to the standard which had 120 ms separation. However, no responses were recorded for deviant tone pairs that were separated by 100 ms, with only 20 ms difference from the standard. Increasing MMN amplitudes have also been correlated with larger angles of deviance from zero degrees azimuth (Paavilainen, Karlsson, Reinikainen & Näätänen, 1989).

MMNs are evoked by speech changes, for example when the speech syllables /da/ and /ga/ are contrasted (Kraus, McGee, Sharma, Carrell & Nicol, 1992; A. Sharma, Kraus, McGee, Carrell & Nicol, 1992; M. Sharma, Purdy, Newall, Wheldall, Beaman & Dillon, 2004). Mismatch responses to phoneme changes have been recorded with (A Sharma et al., 1992) and without (Aaltonen, Niemi, Nyrke & Tuhkanen, 1987) control of acoustic parameters. There have been numerous studies reporting mismatches in response to formant changes (Kraus et al., 1992; Kraus, McGee, Carrell & Sharma, 1995), different syllables (Näätänen, 2001), and grammar and semantics (Näätänen et al., 2007). A review by Csépe (1995) proposed that MMN components vary greatly in amplitude and latency depending on whether the auditory contrasts are frequency, vowel, or articulation based.
MMNs are largest and earliest when stimulus presentations are separated by shorter inter-stimulus intervals. The dominant theory in the literature proposes that the auditory MMN is an indication of when a transient sensory memory trace, established by repeated exposure to the same stimulus, is violated by sudden changes in the sensory features (Bartha-Doering, Deuster, Giordano, Zehnhoff-Dinnesen & Dobel, 2015; Näätänen et al., 1989; Winkler, Reinikainen & Näätänen, 1993). Sensory memory refers to the automatic encoding of stimuli received through the senses and into the brain, but not encoded in long term memory in the hippocampus. This allows for real time comparison and integration of auditory stimuli with the continuous stream of incoming information (Alain et al., 1998). Winkler and colleagues (1993) demonstrated, both behaviourally and electrophysiologically, that the fleeting memory trace could be eliminated if a masking stimulus was presented shortly after the standard stimulus, thus producing no mismatch effect to the following deviant stimulus. The presence of MMNs regardless of stimulus type (for standard and deviant) also suggests that the neural mechanisms involved are capable of storing multiple auditory features and cross-referencing between several memory traces to detect acoustic irregularities (Sams et al., 1984). It has been suggested that this reflects an innate response to one’s acoustic environment which was crucial for survival when humans had to respond quickly to a change in the auditory environment, reflecting highly evolved neuronal adaptations in the human brain (Malmierca, Sanchez-Vives, Escera & Bendixen, 2014; Tiitinen, May, Reinikainen & Näätänen, 1994). The pre-attentive response to auditory change reflected by MMN may be linked to higher cognitive processes like anticipation, reasoning, and learning (Näätänen et al., 2007).

The majority of published evidence documents the substantial influence of attention on the mismatch effect (Sussman et al., 2014). MMN amplitudes are larger for tasks that require the participant to actively attend to and discriminate the acoustic stimuli (passively
and actively), in comparison with when they were tasked to ignore the stimuli (Alho et al.,
1994; Erlbeck et al., 2014; Novak et al., 1990; Sams et al., 1985). However, data has also
been recorded showing no difference in MMN magnitude between successfully detected
stimulus change, and conditions where the listeners were unsure about whether they had
heard the change (unpublished data reported in Näätänen, 1992). There is also report of a
later occurring MMN if another condition is introduced, for example a visual attention task
presented concurrently with auditory stimuli (Alho et al., 1994).

**The effects of age and individual differences**

In concordance with the emergence of the obligatory auditory N2 component at the
beginning of one’s lifespan (Wunderlich et al., 2006), the mismatch effect has also been
observed in ERPs belonging to newborns, infants, and young children (Alho, Sainio,
Sajaniemi, Reinikainen & Näätänen, 1990; Cheour, Leppänen & Kraus, 2000; Csépe, 1995;
Dehaene-Lambertz & Dehaene, 1994; Morr, Shafer, Kreuzer & Kurtzberg, 2002).

Deviant pitch changes in tones have resulted in the elicitation of a slow negative
component resembling an adult MMN in sleeping (Alho et al., 1990; Leppänen et al., 2004;
Martynova, Kirjavainen & Cheour, 2003) and awake (Čeponienė et al., 2002) newborns.
Similarly, temporal and frontal change detection responses have been recorded in infants
after the presentation of a phonetically different novel stimulus (Dehaene-Lambertz &
Dehaene, 1994). However, the mismatch effect only seemed to be reliably evoked across all
participants aged between 3 and 44 months for deviants that differed substantially from the
standard – mismatch was absent in newborns, and only tentatively identified in infants, if the
deviant stimulus only varied from the standard by 200 Hz. The mismatch effect was present
in all age groups once the difference between the two stimuli was increased to 1000 Hz (Morr
et al., 2002). The average MMN latency which peaks at around 250 ms in children (Csépe,
1995; Gomot, Giard, Roux, Barthélémy & Bruneau, 2000; Morr et al., 2002; Shafer et al., 2002), decreases with age (Gomot et al., 2000). This latency change occurs at both frontal and central cortical sites at an approximate rate of 1 ms per month in response to pitch changes in non-speech stimuli (Morr et al., 2002; Shafer, Morr, Kreuzer & Kurtzberg, 2000), until MMN latency plateaus at the average adult latency of about 150 ms (Näätänen, 1992) around 8-10 years of age (Cheour et al., 2000). MMN amplitude is smaller in infants compared to young children (Cheour et al., 2000), and significantly smaller mismatches were reportedly elicited in adults compared to children aged 5-7 years (Gomot et al., 2000; Shafer et al., 2000). The maturation studies referred to above adopted a cross-sectional design, in that participants of various ages were placed into independent groups and compared. So although there are significant differences in MMN morphology between age groups, the maturational progression of the response has not been established using a longitudinal experimental design that follows the same individuals across the years.

Some studies have investigated differences in CAEP and MMN responses between hemispheres. Hemispheric differences in auditory CAEPs were reported by Bellis, Nicol and Kraus (2000) who demonstrated that the obligatory evoked P1-N1 components showed larger left temporal involvement in children aged 8-11 years and young adults aged 20-25 years, and became symmetrical across the hemispheres in adults past the age of 55 years. Despite this difference in scalp distribution which could reflect changes in auditory brain connectivity with age (Sala-Llonch, Bartrés-Faz & Junqué, 2015), MMN morphology did not differ between the three age groups. Because of the cross-sectional design of the experiment, the reported MMN changes did not represent the continuous maturation and then age-related changes in the neural response, and the study did not examine what happens during middle age. It has been observed that MMNs recorded from the right hemisphere in adult participants are consistently larger across multiple types of stimuli, regardless of which ear received the
sounds (Paavilainen, Alho, Reinikainen, Sams & Näätänen, 1991), suggesting some asymmetry of MMN sources from the two hemispheres. These studies have generally not considered participant age and probably largely reflect responses recorded from young adult volunteers.

Both negative and positive mismatch responses (MMR) have been reported for infants. The positive MMR is measurable from frontal and central cortical regions for both non-speech, and speech stimuli (Kraus et al., 1992; Shafer, Yu & Datta, 2010), although similar to the MMN the MMR is also subject to stimulus factors, such as the degree of perceptual phonetic difference (Kraus et al., 1993). Maturation studies have documented a reduction of the positive MMR amplitude with age in childhood (Shafer et al., 2010), with MMR shown to be present in children aged 5-10 years, but not in adults aged 20-30 years (Gomot et al., 2000). Another study that manipulated the presentation rates of non-speech stimuli reported a large positive long latency auditory evoked potential at approximately 400 ms in infants and children up to age four (Choudhury & Benasich, 2011). This is one of few studies that followed a longitudinal design, and the authors were also able to report that the morphology of MMRs in response to changing stimulus presentation rates at infancy predicted subsequent language abilities of the children at 3-4 years of age, which potentially reflected individual differences in temporal processing.

Positive and negative MMRs have been observed in studies investigating change detection of lexical tones in languages other than English (Chandrasekaran, Krishnan & Gandour, 2007; Chandrasekaran, Krishnan & Gandour, 2009; Cheng et al., 2013). Larger deviances between stimulus pairs evoke more prominent mismatches; however Cheng and colleagues (2013) also reported slightly different response patterns that showed the emergence of a positive MMR around 300 ms in new-borns, followed by the addition of a negativity around 150 ms at 6 months of age. Aside from stimulus-related manipulations,
variations in mismatch responses within the same age group could be attributed to individual auditory and linguistic experiences. As demonstrated by Chandrasekaran and colleagues (2007; 2009) native speakers of Mandarin Chinese had MMNs with larger amplitudes and shorter latencies in response to contrasting Mandarin Chinese lexical tones, compared to participants who were native English speakers. The amplitudes of the MMN responses from the Chinese participants were also affected by the degree of similarity between the two contrasting stimuli, and the results suggested that exposure to one's native language resulted in an in-group advantage with faster and more fine-tuned auditory change-detection. Since the discovery of positive mismatch, studies have taken to reporting MMR for both positive and negative responses, instead of just MMN.

**MMRs in clinical populations**

MMR components, emerging at the pre-attentive stages of sensory processing and being present across the lifespan, are not dependent on the cooperation or participation of the individuals being tested. This allows for MMRs (and obligatory cortical responses) to be used as a tool to evaluate difficult to test populations, like infants, young children, and those with behavioural difficulties and neurological impairment. The objective analysis of MMRs provides insight into central auditory dysfunction in different clinical populations such as children with language impairment (Bartha-Doering et al., 2015; Bishop, 2007; Csépe, 1995; Krishnamurti, 2000; Näätänen, 1992; Näätänen, Sussman, Salisbury & Shafer, 2014).

MMRs and other discriminative evoked potentials (including P300 and acoustic change complex (Brett, Tremblay & Korczak, 2008) have been used to investigate central auditory functioning in individuals with a peripheral impairment, sensorineural hearing loss (Kraus et al., 1995). Compared to participants with normal hearing, those with hearing impairment have reduced sensitivity to pitch changes in a continuous acoustic stimulus, and
poorer pitch discrimination is associated with smaller evoked potential amplitudes (Kuruvilla-Mathew, Purdy, Welch, Pontoppidan & Rønne, 2015). Hearing impaired individuals who display poor auditory behavioural discrimination abilities, as well as those who self-report negative experiences with assistive hearing devices, such as cochlear implants, have also shown an absence of MMN (Kraus et al., 1995). Ortmann and colleagues (2013) found that within a group of cochlear implant users, there are large discrepancies in each individual’s development of speech and language and in MMN results. Good phoneme discrimination was attributed to increased MMN activity in frontal cortical regions and auditory sensory memory capacity. Those who demonstrated poor speech perception showed low MMN activity coupled with delayed activation of the left temporal auditory cortex.

In addition to people with hearing loss due to injury to or abnormal development of their peripheral hearing structures, individuals with Auditory Processing Disorder (APD) also demonstrate different CAEP and MMR responses compared to typically developing counterparts (Krishnamurti, 2000; M. Sharma, Purdy & Kelly, 2014). In normal hearing individuals, Kozou and colleagues (2005) documented the effect of different types of noise on central auditory processing. Reduced mismatch amplitudes were recorded in responses to both speech and non-speech stimulus presentations interspersed with either babble or industrial noise; while smaller MMN amplitudes and increased latencies were observed when speech stimuli were presented together with traffic noise. The behavioural discrimination of speech was more affected than that of non-speech stimuli in the presence of noise. Reduced amplitudes and increased latencies have also been reported for N250 (also referred to as N2) components elicited in children with and without APD while processing auditory stimuli in the presence of noise (M. Sharma et al., 2014). Counterintuitively, Sharma et al. observed that the largest difference in the magnitude of the N250 response between children with APD and their typically developing peers occurred in a quiet testing environment. This suggests
that the link between behavioural and neurophysiological measures of speech perception is not straightforward. P1 amplitudes were reduced in APD regardless of stimulus condition (quiet versus noise), which lends support to the hypothesis that deficits in auditory processing stem from compromised cortical representation and encoding of incoming auditory information very early in the process.

Central auditory deficits underlie some language and learning difficulties, and consequently atypical CAEP and MMR patterns have been reported for children with these difficulties (Bishop, 2007). Longitudinal research by Choudhury and Benasich (2011) revealed that infants who were at risk of developing specific language impairment (SLI) based on their family history showed significantly reduced MMRs, compared to peers from the control group, across all points of assessment from 6-36 months of age. Smaller and later obligatory N2 components have been reported for infants with a later diagnosis of SLI; responses recorded at 6-9 months of age correlated with poorer language and cognitive scores measured in the same individuals at 3-4 years of age – demonstrating the potential of using neurophysiological methods to identify and predict developmental outcomes for at-risk children (Shafer & Sussman, 2011).

Children diagnosed with learning disorder, attention deficit disorder, or both, have been shown to perform worse than their neurotypical peers at behaviourally detecting “just-noticeable” differences between two consonant-vowel syllables that varied minutely in their spectral and temporal parameters (Kraus et al., 1996). However, not all the children with learning problems were impaired to the same extent, and Kraus and colleagues (1996) reported that good perceivers of stimulus change had larger MMN responses, both in terms of mismatch magnitude and duration, whereas MMNs were close to absent for those who performed poorly at detecting stimulus contrasts. Kemner and colleagues (1996) documented
smaller MMNs from children with attention deficit hyperactivity disorder, for both passive and active experimental paradigms. In addition, the children had smaller P3 components when required to actively target and respond to deviant auditory stimuli. The results suggest that the children had underlying pre-attentive as well as attentional impairments.

Children with dyslexia (Halliday, Barry, Hardiman & Bishop, 2014) and/or reading disorder (M. Sharma et al., 2006) also display auditory processing deficits both behaviourally and neurophysiologically. Sharma and colleagues (2006) reported that children with reading disorder found it significantly more demanding to perform well on a battery of auditory processing tasks, compared to their typically developing peers and children who had received previous intervention for their reading disorder - failing at least one of the behavioural assessments of auditory processing. The children with reading delay also had substantially reduced MMN magnitude (presented as MMN area) from those with reading disorder, in response to speech stimuli only, and not for alternative variations of tone or chord sounds. Contrasting results were obtained by Halliday and colleagues (2014) who proposed that the outward manifestation of auditory processing difficulties may not be reflected in the immediate stages of change detection and discriminations they found no differences in early mismatch responses between children with and without dyslexia. Instead, reduced mismatch amplitudes for a later discriminative negativity component were reported in those with dyslexia, in response to smaller, less obvious stimulus deviances.

Advances in using neurophysiological methods to investigate pre-attentive processes in clinical populations have resulted in a multitude of studies using various stimuli, cortical recording sites, and participant groups. Suppressed MMR amplitudes have been reported in alcoholics, patients with Alzheimer’s disease (Bartha-Doering et al., 2015), individuals with schizophrenia, bipolar disorder, epilepsy, and aphasia as a result of stroke (Csépe, 1995;
Näätänen et al., 2014). Kileny, Boerst and Zwolan (1997) demonstrated that the major CAEP components, as well as MMRs, were present and measurable in children with cochlear implants. Response latencies varied as expected under different stimulus manipulation conditions, and correlated highly with behavioural tasks of speech recognition, which highlights the potential of using electrophysiological methods to evaluate the effectiveness of assistive hearing devices.

Potential limitations and possible solutions

Despite promising findings for a range of clinical populations, the MMR is subject to a wide range of intra- and inter-individual variations (Lang et al., 1995), and while group neurophysiological data may be robust, there remains significant doubt as to whether MMRs are suitable for individual evaluation of auditory discrimination, especially for clinical application (Kujala, Tervaniemi & Schröger, 2007; Kurtzberg, Vaughan, Kreuzer & Fliegler, 1995). It is near impossible to monitor and experimentally control for intra-individual factors such as drowsiness, alertness, and true passive and active reactions towards the changing stimuli (Lang et al., 1995), all of which have been documented to influence the elicitation and magnitude of CAEP components and mismatch effects (Näätänen, 1992). Substantial shifts in amplitudes and latencies of later evoked potentials, for example the N2, have also been shown in investigations into test-retest effects (M. Sharma et al., 2006). The reliability of such a malleable outcome measure is therefore questionable, especially if the brain is still developing (Kurtzberg et al., 1995), and experiencing more rapid rates of AEP morphology compared to one that has already matured.

Large inter-individual variability could predominantly be attributed to poor signal-to-noise ratios (SNR) generated by a combination of external and internal factors that occurred during a particular block of stimulus presentation. The fact that MMRs are derived from two
separate responses means that the SNR is potentially worsened due to combined SNRs from both standard and deviant waveforms (Kurtzberg et al., 1995). SNR variations could be calculated offline and used to readjust the data prior to analysis and the identification of evoked components (Elberling & Don, 1984). Other possible ways to ameliorate the issue includes adding extra blocks where the standard and deviant stimuli are reversed, so that both are equally accounted for; or including additional blocks until the SNR for the lesser represented deviant stimulus is considered passable (Kujala et al., 2007). However, neither option would be favoured when young participants are involved, or when evaluating clinical populations (Kurtzberg et al., 1995), due to the additional time required. In response to concerns raised with regards to doing neurophysiological research with difficult to test individuals, Näätänen, Pakarinen, Rinne & Takegata (2004) proposed an “optimum paradigm” where several deviant stimuli were embedded into the same block against one standard stimulus, demonstrating that it was possible to test for change detection and obtain MMRs for different stimulus combinations in the shortest period of time. Nevertheless, the heterogeneity of neuro-atypical populations (Näätänen et al., 2014) still restricts the clinical applications of MMRs - although mismatch trends are present, they remain inconsistent and inconclusive for the time being (Bishop, 2007; Lang et al., 1995) and consequently MMRs largely remain a research rather than a clinical tool.

CAEPs and MMRs in ASD

Similar to other neurodevelopmental disorders, individuals with ASD display atypical auditory processing behaviour. In light of the emerging overlap between APD and ASD (discussed in Chapter 2), a significantly growing body of work has seen CAEP research being conducted with individuals with ASD (reviews: Bomba & Pang, 2004; Groen, Zwiers, van

Early work by Novick, Vaughan, Kurtzberg and Simson (1980) reported no differences in response magnitudes between children with and without autism for initial P60 and N100 obligatory evoked potential components. Instead they found reduced P200 and P300 components in response to their various experimental conditions, consisting of continuous, pitch changing, and within-sequence deleted auditory stimuli. It was initially proposed that auditory processing deficits in individuals with autism may be reflected in impairments at higher levels of cognitive functioning, involving the encoding and storing of sensory information, instead of impairments at the earliest stages of stimulus detection and perception. Another early study found reduced P3 responses to deviant and novel phonemes, but failed to observe atypical MMNs from children with ASD (Kemner, Verbaten, Cuperus, Camfferman & van Engeland, 1995). Subsequent studies have reported significantly smaller and later secondary negative responses from bilateral temporal regions (N1c/T-complex) in children with autism while they processed a constant stream of auditory tones bursts (Bruneau, Roux, Adrien & Barthélémy, 1999); as well as a left hemisphere deficit while they processed stimuli of increasing intensities (absence of positive correlation with response amplitude), and significant links between N1c amplitude and the verbal and non-verbal communicative abilities of these children with autism (Bruneau, Bonnet-Brilhault, Gomot, Adrien & Barthélémy, 2003).

More recent work has confirmed that autism-related deficits manifested equally, if not predominantly, at the primary stages of auditory processing, affecting CAEPs within the 80-200 ms latency period. The consequences of this early dysfunction are thought to contribute to deficits in later cognitive-based stages of stimulus processing (Bomba & Pang, 2004; Groen et al., 2008). This has given rise to the more extensive investigation of MMR patterns
in the ASD population - using these pre-attentive electrophysiological responses to evaluate potential impairments in the underlying processes governing fundamental detection of stimulus change in one’s immediate auditory environment. Findings in this area have remained inconclusive however. Depending on various factors - from experimental paradigm to individual participant differences - those with ASD have revealed both hyper- or hypo-activation of the auditory pathway compared to typically developing peers (Bomba & Pang, 2004; Kujala et al., 2013), with different studies showing either increased or reduced MMR magnitudes in ASD.

Studies using tone-based stimuli that differed in frequency have found larger MMN amplitudes towards deviant sounds, and enhanced P3a responses towards novel sounds, in children and adolescents with ASD and mental retardation (Ferri et al., 2003). Another study reported significantly shorter mismatch latencies, coupled with large P3a amplitudes, which correlated with an increased intolerance to change - a factor derived from a behavioural evaluation of autism symptoms (Gomot et al., 2011). Kujala and colleagues (2007) found enhanced MMNs in response to a multitude of varying sound features - in particular larger amplitudes elicited by deviant gaps in the sequence or shorter stimulus lengths, and faster responses elicited by changes in stimulus frequency in children with ASD. The link between auditory evoked responses and behaviour is complex, but such results suggest that an autistic individual’s tendency to overreact to incoming sensory information may be reflected in their CAEPs. However, the opposite effect has also been observed - in one case, magnetoencephalography (MEG) revealed no identifiable mismatch field (combined strength, latency, and location of the response) in the group with ASD (Tecchio et al., 2003). In another study, children with ASD exhibited substantially smaller MMNs compared to controls during the passive experimental condition, when they were not required to pay attention to tonal stimuli deviating in pitch. Conversely, when they were asked to attend to
the tones, the magnitude of the mismatch responses did not differ between the two groups, suggesting that individuals with ASD may have deficits in their automatic auditory processing, but may engage in compensatory mechanisms through the effortful use of attention to successfully discriminate between changing sounds in their surroundings (Dunn, Gomes & Gravel, 2008).

Although neurophysiological abnormalities evident in children with ASD have been reported to be more pronounced in responses towards tones compared to syllables (Jansson-Verkasalo et al., 2003); studies using semi-synthetic and real speech sounds have also reveal substantial effects. More recently, compared to typically developing peers, children with ASD have been shown to display larger MMNs when detecting changes in syllable intensity, but smaller MMNs for discriminating changing syllable frequencies and durations (Kujala et al., 2010). Delayed mismatch field responses recorded using MEG were also present for phonemic differences, which correlated significantly with the expression of more severe autism symptoms (Kasai et al., 2005). The mix of both hyper- and hypo-responses again highlights the heterogeneity of abilities and neural responses found in this clinical population.

ASD-related irregularities seem to be more evident in studies that compare the processing of speech with non-speech stimuli. Kuhl, Coffey-Corina, Padden and Dawson (2005) demonstrated that children with ASD, through an auditory preference spontaneous head-turn test, tended to be drawn towards non-speech sounds instead of infant-directed speech. Compared to their typically developing peers, these children with ASF also did not show a significant MMN in response to phonetic changes. A series of studies by Lepistö and colleagues (Lepistö et al., 2006; Lepistö, Nieminen-von Wendt, von Wendt, Näätänen & Kujala, 2007; Lepistö et al., 2008) delved further into this line of investigation. CAEPs were obtained using an oddball paradigm that incorporated speech-based pitch, duration, and phonetic changes, and corresponding deviances in sequences of non-speech stimuli. Children
with ASD exhibited poorer discriminative performance and reduced MMN amplitudes in response to different stimulus durations, yet showed increased MMNs towards pitch and phoneme changes. In addition, children with ASD had diminished attentional P3a components for speech-based pitch and phoneme deviances, but this effect was not mirrored in their non-speech equivalents (Lepistö et al., 2006). Adults with ASD also showed reduced P3a magnitudes for speech stimuli, but unlike children they demonstrated enhanced P3a components in response to changes in non-speech sounds, indicating the possible effects of maturation and learned compensatory strategies. Furthermore, adults with ASD had larger MMNs towards deviating stimulus pitch (consistent with children), and durations (different from children) (Lepistö et al., 2007). Lepistö and colleagues (2008) additionally documented that children with ASD had enhanced MMNs towards changing pitch and phonemes, as long as the stimuli in each sequence remained from within the same category. Once deviations were crossed between categories, for example phonemes varying in pitch, the authors reported that the advantageous effects were attenuated. These findings support the observation that many with ASD may display good discrimination abilities, but are less able to simultaneously process too many varying features of sensory information, highlighting the complexity of auditory processing in this population.

Research looking into ASD-related emotion change detection is relatively sparse, most likely due to the ambiguity surrounding the acoustic parameters of emotion (discussed in Chapter 1). Nevertheless, a few studies have reported significant findings. Kujala, Lepistö, Nieminen-von Wendt, Näätänen and Näätänen (2005) found that a group of adults with ASD were less likely than controls to have a measurable MMN in response to changes in emotional undertones delivered through a spoken word. In addition, any mismatch effects elicited were significantly smaller in amplitude and delayed compared to an age-matched typically developing control group. However, there is a potential misattribution of the source
of the MMNs, because the word presented in a neutral emotion as the standard, and either in a commanding, scornful, or sad undertone as deviants, all also differed in stimulus duration and hence the findings could also reflect responses to changes in stimulus duration. Korpilahti and colleagues (2007) also conducted an oddball experiment with a word spoken in either tender (standard) or angry (deviant) tones of voice. In this case, stimulus duration, both overall as well as for each phoneme’s length, was the same for both emotions. The authors reported both early and late MMNs. For the early mismatch, children with ASD exhibited greater amplitudes and longer latencies compared to controls. For the late MMN, the children with ASD had responses of a similar magnitude to controls, but which peaked earlier. More recently, Fan and Cheng (2014) used syllables and non-vocal sounds produced in neutral (standard), happy, and angry (deviants) emotional tones to evoke MMRs from adults with and without ASD. All the stimuli were matched for length and intensity. The authors reported that individuals with ASD did not have any MMNs in response to deviating emotional syllables, and showed significantly diminished MMNs in response to deviating emotional non-vocal sounds. Additionally, it was observed that the amplitude of the mismatch towards angry sounding deviant tones was negatively correlated with the number of autistic traits displayed by those in the clinical group. These findings implicate a fundamental deficit in acoustically discriminating between varying affective prosodic cues irrespective of the type of stimulus.

Some have stressed that individuals with varying degrees of ASD severity, especially those additionally impacted by intellectual disability, might skew the neurophysiological data, and could potentially reflect wider cognitive deficits instead of difficulties with fundamental sensory processing (Seri et al., 2007). The inconclusiveness of the existing literature around this field of study could reflect in part variations across participants and emphasises the need for further consistent investigations into how atypical
neurophysiological patterns reflect an underlying auditory processing deficit in children with ASD (Bomba & Pang, 2004; Groen et al., 2008; Kujala et al., 2013). It has also been proposed that the use of MEG may provide more comprehensive insight, including both spatial and temporal information, into the sources of auditory processing dysfunction in ASD (Roberts et al., 2008).

Studies 2 and 3: Background methodology

The second and third studies in this series of work used neurophysiological methods to record CAEPs in response to changes in affective prosody. By using an oddball paradigm, we hoped to reveal MMRs driven by the perception of changing emotional speech stimuli, comparing typically developing adults, typically developing children, and children with HF-ASD. We also investigated the effect of a combined intervention (computer-based emotion training and RMHAs) on the CAEPs recorded from the children with HF-ASD, by measuring their responses before and after the 3-week trial.

The making and selection of the speech stimuli

Four speech stimuli – monosyllabic utterances “ba” in angry, happy, sad, and neutral tones of voice – were used in an oddball paradigm to elicit CAEPs from the participants. These came from natural speech samples that were taken from Ross and colleagues’ (1997) battery of prosody production tasks. The original samples consisted of a string of seven “ba” utterances, spoken by a single male actor in neutral, angry, happy, sad, surprised, and disinterested tones. The researcher chose the ones corresponding to angry, happy, sad, and neutral, and digitally cut and edited them using Adobe Audition software (version CS6) into individual monosyllables – each 200 ms in length, and on average 15 dBV in root mean square intensity. A total of 28 separate samples were derived from the originals, 7 “ba’s” from the four original affective utterances. The researcher then recruited 14 neuro-typical
young adults with normal hearing to blindly rate the 28 monosyllables, with the task of identifying which of the four emotions they thought was being conveyed. The final four speech stimuli were selected, one for each emotion, because they received a 14/14 rater consensus and were all judged as conveying the intended emotion.

**Speech stimuli presentation settings**

The CAEP recordings took place in a sound treated two-room environment. Prior to the experiment, measures were taken to adjust the intensity of the speech stimuli in relation to the testing environment. A sound level meter (Brüel & Kjaer, type 2215, Denmark) on an adjustable tripod was placed on a leather reclining chair at 50 cm in height, and 150 cm away from a Turbosound IMPACT 50 loudspeaker positioned centrally facing the chair. This recliner was where the participants sat during the experiment, and remained in a fixed position for the duration of the study. The stimuli were presented through the Gentask software on NeuroScan STIM2, and was incrementally adjusted until the sound pressure level (SPL, linear weighting) was 70 dB SPL +/- 0.9 dB. The loudspeaker remained in this fixed position for all the CAEP recordings, ensuring that the four speech stimuli were presented to all participants at the same intensity level. Following a trial run, the researcher noticed some audible inter-stimulus interference, as there were some software-to-hardware delays between the stimulus presentation computer and the loudspeaker; 40 ms of silence (20 ms either end of the utterance) was added to each of the speech stimuli, which removed the noise, but retained the 200 ms of voiced time; this silence was accounted for in the post-processing of response.
Chapter 6. Exploring change detection in affective prosody in adults through cortical auditory evoked potentials and mismatch responses

This chapter is written with the intention to submit it as a manuscript for publication.
Introduction

Detecting changes in a stream of incoming auditory stimuli at the earliest, pre-attentive stages of sensory information processing has primarily been investigated through the recording of CAEPs, and the subsequent calculation of a MMN effect (Csépe, 1995; Näätänen, 1992; Sams et al., 1985). An MMN is typically derived by subtracting evoked potentials elicited in response to a numerously repeated “standard” stimulus, from that elicited in response to a less frequently and randomly presented “deviant” stimulus (Näätänen, 1992). However, it has been demonstrated that standard-deviant, or deviant-deviant-standard experimental configurations also produce mismatch effects, albeit with a smaller magnitude (Sams et al., 1984). It has been suggested that this reflects an innate response to one’s acoustic environment which was crucial for survival when humans had to respond quickly to a change in the auditory environment, reflecting highly evolved neuronal adaptations in the human brain (Malmierca et al., 2014; Tiitinen et al., 1994). The pre-attentive response to auditory change reflected by MMN may be linked to higher cognitive processes like anticipation, reasoning, and learning (Näätänen et al., 2007).

Initially, it was most popular for MMN studies to explore variations in neural responses resulting from the manipulation of specific acoustic parameters, such as pitch (Alho et al., 1994; Näätänen et al., 1992; Novak et al., 1990; Paavilainen et al., 1991; Tiitinen et al., 1994; Winkler et al., 1995); intensity (Näätänen et al., 1989; Paavilainen et al., 1991); and duration (Kujala et al., 2001; Paavilainen et al., 1991) of non-speech acoustic stimuli. Most of these studies reliably reported significantly larger responses to deviant compared to standard stimuli, resulting in substantial a mismatch effect, usually observed as a negative trough between 100 and 250 ms post presentation of the deviant stimulus (Näätänen, 1992; Näätänen & Alho, 1995).
Changes in more complex stimuli such as speech, including phoneme changes (Aaltonen et al., 1987), formant changes (Kraus et al., 1992; Kraus et al., 1995), different syllables (Näätänen, 2001), and grammar and semantics (Näätänen et al., 2007) also produce reliable MMN. It has been proposed (Csépe, 1995) that speech-related MMNs vary considerably more in amplitude and latency, compared to non-speech, due to the significantly greater multitude of parameters that can be manipulated – from basic acoustic factors, to phonemic and articulation, up to semantic interpretation. Using speech as stimuli also introduces language related confounds. Studies show an in-group advantage in native speakers, manifesting behaviourally as faster change detection responses with the ability to discriminate between the smallest lexical contrasts, and neurophysiologically as MMNs with larger amplitudes and shorter latencies compared to non-native speakers (Chandrasekaran et al., 2007; Chandrasekaran et al., 2009). In children, mismatch responses to manipulating speech-related factors show not just negative MMNs, but also positive mismatches (Cheng et al., 2013; Kraus et al., 1992; Shafer et al., 2010), peaking either once or twice at 100 and 300 ms post stimulus presentation. The term mismatch response (MMR) has been increasingly used in the literature to describe both positive and negative responses.

Despite many studies investigating how the brain makes pre-attentive discriminations between auditory changes in both speech and non-speech sounds, few studies have examined the impact of manipulating stimulus emotion or affect. Researchers are yet to reach a consensus regarding how to define emotion. Some authors have proposed that different configurations of acoustic properties underlie emotional expression – reporting that speech expressing anger and happiness generally has increased mean F0, F0 range, and vocal intensity; sadness is characterised by slower rate of speech and longer inter-articulation silence; and the perception of fear in a voice is attributed to increased mean F0 and articulation rates (Banse & Scherer, 1996; Murray & Arnott, 1993; Yildirim et al., 2004).
Others have proposed that expression of affect cannot be categorised in a simple way into separate emotions, and instead characterise affect in three dimensions that each vary across a spectrum, valence, arousal, and intensity (Belyk & Brown, 2014; Banse & Scherer, 1996; Johnstone & Scherer, 2000). This makes it hard to implement consistently in studies, and is likely to be part of the reason why the literature in this area remains relatively sparse (McCann & Peppé, 2003).

Few emotion discrimination experiments have investigated healthy populations undiagnosed with any neurological or developmental impairments. Cheng, Lee, Chen, Wang and Decety’s (2012) work with neonates showed that auditory evoked responses elicited by fearful utterances were larger in negative amplitude at a latency of approximately 400 ms, and those elicited by happy utterances were larger in positive amplitude at around 500 ms, with these significant differences recorded from frontal electrode sites F3 and F4. Additionally, the authors demonstrated that the newborns had significant MMNs when angry and happy speech syllables were contrasted, but not for non-vocal equivalent sounds. Thönnessen and colleagues’ (2010) recorded significant MMNs in adults at latencies less than 200 ms in response to contrasts between neutral as a standard disyllabic pseudo-word, and angry, happy, and sad as deviants. Similar results can be seen in clinical studies, where adults have been used as control participants. For example, MMNs occurring at approximately 200 ms post stimulus presentation, derived from angry-deviant-neutral-standard speech syllables and non-vocal equivalent sounds, were larger than responses derived from the happy-deviant-neutral-standard experimental condition (Fan & Cheng, 2014). There have also been reports of adult participants having early-MMNs at latencies of 140-190 ms, and late-MMNs at 300-650 ms, in response to deviant changes from a neutral standard stimulus (Korpilahti et al., 2007; Kujala et al., 2005).
The literature indicates that MMN to emotional contrasts can be recorded in adults, but the MMN latency region differs somewhat across studies. The purpose of this study is to contribute to the sparse body of research addressing the early sensory processes involved in detecting changes in emotion in speech in neurologically healthy adults. By primarily focusing on adults, who’s CAEPs have matured and stabilised, the findings from this study would justifiably be a baseline measure from which many other groups can be compared - from children with and without various neurological, developmental, or learning impairments, to other adults with similar difficulties. Using novel, natural speech monosyllables with different underlying emotional tones, this feasibility study aims to explore and describe the obligatory evoked potentials and MMRs derived from the differential perception of these stimuli. It was hypothesised that there would be minimal differences in response magnitude between emotions at the early obligatory CAEP components. However, as auditory processing progresses, variations in response and mismatch elicitation between the four emotions are expected.

Methods

Participants

Twenty adults (10 males, 10 females, $M_{age}$=27 years, $SD_{age}$=3.84, Range$_{age}$=22-36 years) participated. They were mostly associates of the researcher, or were recruited via word of mouth. Eight were NZ Europeans, 8 Asians, 2 Eastern Europeans, 1 Latin American, and 1 Samoan. This was a reasonably close reflection of the ethnic diversity of New Zealand (Census: Statistics New Zealand, 2013), although Asians were over represented, and no Maori were recruited, due to the researcher’s personal network of acquaintances. Nine of the participants did not have English as their first language, but all have been exposed to English since at least secondary school age, if not younger. IQ was not measured, but were assumed to be comparatively equal between individuals because all participants had tertiary education.
qualifications. All participants also self-reported normal hearing and no neurological or psychiatric condition.

**Cortical auditory evoked potential recordings**

**Stimuli and sequences**

The speech stimuli were from existing recordings of monosyllables produced by a male speaker in angry, happy, sad, and neutral emotional tones of voice (Ross et al., 1991). Details of the sampling and editing method and the selection of the four speech stimuli are described in Chapter 5.

All the speech stimuli were 200 ms in length. The time waveforms in Figure 6 show that the speech stimuli were, on average, matched for root mean square intensity but differed in their temporal characteristics.

An oddball paradigm was used to investigate auditory affect discrimination using MMRs. The stimuli were programmed into sequence blocks using the NeuroScan STIM2 Gentask software. Each block consisted of 216 speech stimuli in a ratio of 70 standard stimuli (156 stimuli) to 10:10:10 deviant stimuli (20 per emotion). Seven blocks were sequenced to have ‘neutral’ as the standard, and ‘angry’, ‘happy’, and ‘sad’ as the deviants. Three blocks were sequenced to each have either ‘angry’, ‘happy’, or ‘sad’ as standard, and the remaining emotions as deviants. Table 9 summarises the number of stimuli per emotion per presentation (standard/deviant).

Every sequence began with 20 standard stimuli, followed by a pseudorandom presentation of standard and deviant stimuli. The order of emotional deviants inserted into the sequence was randomised, but was adjusted so that at 2 or 3 standards occurred between each deviant. There was a 640 ms inter-stimulus interval (ISI).
Figure 6. Electrical time waveforms (Adobe Audition CS6) depicting the monosyllabic /ba/ stimulus presented in four emotions. dBV represents relative sound intensity in decibel-voltages, plotted against time in milliseconds.

Table 9.
Total number of stimulus presentations per emotion, in their two different forms (as standard and as deviant).

<table>
<thead>
<tr>
<th>Emotion</th>
<th>Number of stimuli as Standard</th>
<th>Number of stimuli as Deviant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neutral</td>
<td>1152</td>
<td>60</td>
</tr>
<tr>
<td>Angry</td>
<td>156</td>
<td>180</td>
</tr>
<tr>
<td>Happy</td>
<td>156</td>
<td>180</td>
</tr>
<tr>
<td>Sad</td>
<td>156</td>
<td>180</td>
</tr>
</tbody>
</table>
Experimental setup

The CAEP recordings took place in a sound treated two-room setup, with a leather reclining chair for the participants to sit in. Using the Gentask software on NeuroScan STIM2, the stimuli sequences were presented at 70 dB SPL via an Australian Monitor Synergy SY400 power amplifier and Sabine Graphi-Q GRQ-3102 equaliser, connected to a Turbosound IMPACT 50 loudspeaker. An additional half-inch polarised condenser free-field microphone, connected to a Bruel and Kjaer measuring amplifier and oscilloscope, was used to calibrate and externally monitor the sound levels of the stimuli in the enclosed testing environment.

The loudspeaker was positioned directly in front of the participant seated on the recliner, approximately 150 cm away (dependent on individual size). Situated behind the loudspeaker was a television on a stand. Participants were instructed to watch a movie of their choice with the audio turned off and the subtitles on. This was to ensure their alertness, while their passive neural responses to the stimuli were being recorded. Participants were also asked to minimise their blinking and body movements during the experiment.

The CAEPs were recorded using the NeuroScan Inc. Evoked Potential System (version 4.5) with a SynAmps 2 amplifier. The surface of the participant’s scalp was cleaned using NuPrep EEG and ECG skin prep gel, and eight 10 mm gold electrodes were placed on Cz, Fz, F3, F4, A1, and A2 locations, with a ground electrode on the forehead, and an eye blink electrode above the right eye. The electrode on the right mastoid (A2) served as the reference electrode. During offline processing, the left and right mastoid electrodes were linked, and Cz, Fz, F3, and F4 were subsequently re-referenced. The electrodes were filled with Quik-Gel™ conductive gel (Neuromedical supplies) and secured onto the scalp with
Transpore surgical tape. Electrode impedance was kept at or below 5 kΩ. Recordings were made with a sampling rate of 500 Hz and a bandpass filter setting of 0.1-100 Hz.

**Data processing**

Further offline processing was done using the Edit software from NeuroScan Inc. Continuous recording files were epoched from -100 ms pre-stimulus to 850 ms post-stimulus, followed by a baseline correction. Any responses exceeding ±150 uV were rejected as artifacts. A minimum of 20 blinks were required to estimate an average blink, and to avoid double detection the ocular artifact rejection procedure discarded anything that occurred <400 ms before a previous one. The data was filtered through a 30 Hz low-pass filter (12 dB/octave slope, zero phase shift). Separate average files were generated for each participant, for each of the four emotions in their two presentation forms (standard/deviant). Grand average waveforms were plotted from -100 ms to 850 ms to encompass pre-stimulus responses, the 200 ms stimulus length, and post-stimulus responses for the duration of the 640 ms ISI, without overlapping with the subsequent stimulus.

**Data analysis**

The data was analysed in four steps to address the following research questions:

1) Are there significant differences in peak amplitudes and latencies between the four emotions – neutral, angry, happy, and sad – for each identified obligatory CAEP component? Does the site of electrode placement have an effect?

The researcher firstly created grand average waveforms for each electrode site by combining all four emotions in their standard form. Single-sample t-tests were conducted on these grand average waveforms at each millisecond. Contiguous time windows were
highlighted where the waveforms deviated at a significance level of $p \leq 0.001$ from 0 uV across all four electrodes. The midpoint within each of these deviant time periods was located.

The researcher then identified 6 time windows of interest by measuring ±20 ms either side of the midpoint. These consisted of a P1 component reflecting the onset of the /b/ consonant (P1c: 60-100 ms), an N1 consonant component (N1c: 90-130 ms), a P1 component reflecting the onset of the /a/ vowel (P1v: 140-180 ms), and its N1 equivalent (N1v: 180-220 ms), followed by a P2 component (260-300 ms). Studies looking into ERP patterns in response to spectro-temporal changes during consonant-vowel transitions have reported differential amplitudes and latencies in the initial P1 and N1 components (Čeponienė, Torki, Alku, Koyama & Townsend, 2008; Sandmann et al., 2007). An N400 component (450-550 ms), thought to represent cognitive appraisal of the semantic (Kutas & Hillyard, 1984), and possibly emotional content of the stimulus, was also investigated. The negative component had a visibly longer time window of deviance from 0 uV, and so the time frame was set at a wider range. Peak amplitudes and latencies were picked for each of the 20 adult participants within the above time frames. A within-subject repeated measures ANOVA was conducted for each component - across four levels of emotions, four electrode sites, and with age as a covariant, followed by Bonferroni-corrected post-hoc analyses.

2) Within an emotion-deviant-neutral-standard oddball stimulus block: are there significant differences in MMR magnitudes between angry, happy, and sad difference waveforms? Is there an effect of electrode location?

Firstly, the researcher created a combined difference waveform by subtracting an averaged wave consisting of angry, happy, and sad deviant responses, from the standard neutral wave. Combined difference waveforms were generated for each electrode placement site (Cz, Fz, F4, F3). The same method of finding contiguous time windows that deviated significantly from 0 uV for all electrodes was implemented. Once again, MMR latency
windows were derived by measuring ±20 ms either side of the midpoint of these time frames. MMR magnitudes were calculated by taking an average of the amplitude within each of these 40 ms latency windows. A total of four MMRs were identified, bound by the following latencies: 130-170 ms (negative MMR), 210-250 ms (positive MMR), 260-300 ms (negative MMR), and 380-420 ms (negative MMR).

These MMR latency windows were then applied to the three separate difference waveforms for each ‘emotion as deviant’ wave compared to the ‘neutral as standard’ wave. MMR averaged amplitudes were calculated for each participant. A within-subject repeated measures ANOVA was conducted for each MMR - across three levels of emotional deviance, four electrode sites, and with age as a covariate, followed by Bonferroni-corrected post-hoc analyses.

3) Within a neutral-deviant-emotion-standard oddball stimulus block: are there significant differences in MMR magnitudes between angry, happy, and sad difference waveforms? Is there an effect of electrode location?

Using the same procedure as the previous analysis, except in reverse with neutral as a deviant, and an average of the three other emotions presented as standard stimuli, the researcher identified three MMRs bound by the following latencies: 230-270 ms (negative MMR), 310-350 ms (positive MMR), and 500-850 ms (negative). This last response could be regarded as a late discriminative negativity (LDN). The LDN is a relatively unstudied component that has been reported to occur at approximately 400 ms post stimulus presentation in adults, and between 400-460 ms in children (Korpilahti et al., 1995), and is thought to reflect the processing of more complex auditory stimuli, especially with regards to language and speech processing (Cheour et al., 2001; Halliday et al., 2014; Korpilahti et al., 2001). The additional delay observed in this study’s MMR latency results might be attributed to the complex nature of the speech stimulus with underlying emotion.
These three MMR latency windows were then applied to the three separate difference waveforms that illustrate mismatch effects for the ‘neutral as deviant’ wave compared to each ‘emotion as deviant’ wave. MMR averaged amplitudes were calculated for each participant. A within-subject repeated measures ANOVA was conducted at each MMR - across three levels of emotional deviance, four electrode sites, and with age as a covariate, followed by Bonferroni-corrected post-hoc analyses.

4) What do the deviance effects for each emotion look like?

Firstly, the researcher derived difference waveforms for each of the four emotions by subtracting the data for when they were presented as deviants from that collected when they were presented as standards. Each emotion consisted of four difference waveforms, one from each electrode site. The researcher also collapsed across electrode sites and generated an averaged difference waveform for each emotion.

Single-sample t-tests were conducted at each millisecond to determine when the waveform amplitudes significantly deviated from 0 uV at p<.001. Time windows were identified by highlighting periods of contiguous deviation. Average amplitudes were calculated within each time window. Within-subject repeated measures ANOVAs were conducted at each time window to compare MMR magnitudes between the four electrode sites, as well as across emotions for the averaged waveform. No significant differences were found across electrodes, hence results are reported for the difference waveform collapsed across electrodes.

Results

1) CAEP components from all emotions presented as standard stimuli

The first step of the analysis procedure aimed to validate the stimuli, which have not been used in any previous studies. The stimuli were matched for average intensity and
duration, although the time waveforms differed from each other, due to the emotional content and natural source of the monosyllable and the uncontrolled movements of the speaker’s vocal chords. The protocol meant that there were large discrepancies in the number of times each stimulus was presented (see Table 9), with fewer presentations of the emotional compared to the deviant stimuli.

A within–subjects repeated measures ANOVA, for the early P1-N1 components, revealed mostly no significant differences between emotions. The only exception was an emotion effect on latency of the P1c component ($F_{(3, 24)}=3.16, p=.043$), which appeared to be between neutral ($M=88.24, SD=11.8$) and angry ($M=78.24, SD=12.57$), however post-hoc analyses showed no significant difference ($p=.076$). There were significant electrode effects on amplitude at P1c ($F_{(3, 24)}=3.44, p=.033$), and latency at N1c ($F_{(3, 24)}=5.33, p=.007$), with a significant difference between Fz ($M=2.31, SD=1.64$) and F4 ($M=2.013, SD=1.61; p=.038$). There were minimal between-emotion differences in the early cortical peaks suggesting that the acoustic differences across the stimuli had a negligible effect.

For the P2 component, an emotion–electrode interaction effect was evident ($F_{(9, 63)}=2.74, p=.034$), with response amplitudes for neutral ($M=3.73, SD=1.95$) exceeding those for sad ($M=2.90, SD=1.98$) at all four electrode sites ($p=.034$). There was a trend for an emotion effect on P2 latency for ($F_{(3, 24)}=2.37, p=.083$), with significant post-hoc differences for sad ($M=325.18, SD=28.45$) that peaked later than neutral ($M=275.13, SD=10.31$), Angry ($M=262.24, SD=16.79$), and Happy ($M=272.54, SD=39.65$; all $p<.001$).

Emotion effects observed for the N400 component potentially reflect secondary higher level processing of the emotional content of the stimuli. For N400 amplitude ($F_{(3, 45)}=3.71, p=.018$), angry ($M=-2.81, SD=1.17$), and happy ($M=-3.34, SD=1.61$) both exceed sad ($M=-1.97, SD=1.51$) at $p=.001$ and $p=.005$, respectively. For N400 latency ($F_{(3, 45)}=3.08$,}
once again angry ($M=485.45, SD=24.16$), and happy ($M=468.66, SD=34.84$) both occur earlier than sad ($M=502.75, SD=26.47$) at $p=.003$ and $p=.012$, respectively.

Figure 7 illustrates overlapping grand average CAEP waveforms for each emotion, from each of the four electrode sites.
Figure 7. Grand average CAEP waveforms generated for each emotion, for the four electrode sites (Cz, Fz, F4, F3).
2) MMRs in ‘Emotion as deviant’ – ‘Neutral as standard’ waveforms

The second step of the analysis focused on emotion and electrode effects at MMRs identified from ‘emotion as deviant - neutral as standard’ difference waveforms. Figure 8 illustrates grand average difference waveforms averaged across emotions (in black), together with separated grand average difference waves for each emotion (in colour). The four MMRs were derived from, and can be seen in the averaged waveform.

Repeated measures ANOVA revealed that electrode differences were only present at MMR2, latency window 210-250 ms \( (F(3,51)=4.00, p=.012) \), with mismatch magnitudes significantly larger at Fz \( (M=-2.19, SD=0.99) \) than F4 \( (M=-1.96, SD=0.94; p=.004) \) across emotions.

Although no overall emotion effect was significant, however individual pairwise comparisons showed that angry and happy both elicited MMRs of reduced magnitudes compared to sad. Within the latency window of MMR1, angry \( (M=-1.79, SD=1.31) \) and happy \( (M=-1.78, SD=1.45) \) were significantly smaller than sad \( (M=-3.34, SD=1.77) \), at \( p=.002 \) and \( p=.003 \), respectively. Within MMR4, angry \( (M=-1.68, SD=1.67) \) was significantly smaller than sad \( (M=-3.15, SD=2.04; p=.046) \).

At MMR3, there was an overall effect of emotion on response magnitude \( (F(2, 34)=3.66, p=.036) \). However, there were no significant post-hoc comparisons. This could potentially be attributed to the additional observation that age also significantly influenced \( (F(2, 34)=4.09, p=.026) \) the variance seen between emotions in a systematic way.
Figure 8. Grand average difference waves averaged across emotions (in black), and underneath, separated grand average difference waves for each emotion (in colour). Presented for the four electrode sites (Cz, Fz, F4, F3). Dev – deviant, Std – standard.
3) MMRs in ‘Neutral as deviant’ – ‘Emotions as standard’ waveforms

The third step of the analysis procedure focused on investigating potential emotion and electrode effects at MMRs identified from ‘Neutral as deviant - emotion as standard’ difference waveforms. Figure 9 illustrates the difference wave averaged across emotions (in black), together with the separated difference waves for each emotion (in colour). The two MMRs and potential LDN were derived from, and can be seen in the averaged waveform.

Results from the within-subject repeated measures ANOVA are presented in Table 10. Significant emotion-electrode interaction effects were observed for MMR1 and the LDN. For the positive MMR2, the ANOVA showed an overall effect of emotion on response magnitude but there were no significant post-hoc comparisons. The effect of age as a covariate was significant ($ F_{(2, 36)}=6.18, p=.005$), even though the age range of participants was relatively narrow (22 to 36 years).
Table 10.
Repeated measures ANOVA results comparing MMR magnitudes between emotions at identified regions of mismatch. df (e) – degrees of freedom (error), M – mean, SD – standard deviation.

<table>
<thead>
<tr>
<th>MMR</th>
<th>F (df, dfe)</th>
<th>p-value</th>
<th>Electrode</th>
<th>Emotion 1 (M, SD)</th>
<th>Emotion 2 (M, SD)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>MMR1</td>
<td>4.07 (6, 108)</td>
<td>.017</td>
<td>Cz</td>
<td>Angry (-3.36, 2.13)</td>
<td>Sad (-2.11, 1.20)</td>
<td>.004</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Fz</td>
<td>Angry (-3.57, 2.31)</td>
<td>Sad (-2.12, 1.43)</td>
<td>.005</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>F4</td>
<td>Angry (-3.45, 2.13)</td>
<td>Sad (-1.80, 1.42)</td>
<td>.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>F3</td>
<td>Angry (-3.61, 2.17)</td>
<td>Sad (-1.97, 1.59)</td>
<td>.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Fz</td>
<td>Happy (-3.31, 2.17)</td>
<td>Sad (-2.11, 1.20)</td>
<td>.015</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>F4</td>
<td>Happy (-3.31, 2.41)</td>
<td>Sad (-2.12, 1.43)</td>
<td>.034</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>F3</td>
<td>Happy (-3.12, 1.71)</td>
<td>Sad (-1.80, 1.42)</td>
<td>.005</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>F3</td>
<td>Happy (-3.26, 1.81)</td>
<td>Sad (-1.97, 1.59)</td>
<td>.013</td>
</tr>
<tr>
<td>MMR2</td>
<td>4.94 (2, 36)</td>
<td>.013</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>(310-350ms)</td>
<td></td>
<td></td>
<td>LDN</td>
<td>Happy (-1.83, 1.09)</td>
<td>Sad (-2.26, 1.56)</td>
<td>.037</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Fz</td>
<td>Happy (-1.77, 1.68)</td>
<td>Sad (-2.40, 2.02)</td>
<td>.020</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>F4</td>
<td>Happy (-1.96, 1.60)</td>
<td>Sad (-2.78, 1.97)</td>
<td>.010</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>F3</td>
<td>Happy (-2.41, 2.39)</td>
<td>Sad (-2.72, 2.35)</td>
<td>.524</td>
</tr>
<tr>
<td>LDN</td>
<td>2.87 (6, 108)</td>
<td>.012</td>
<td>Cz</td>
<td>Happy (-1.83, 1.09)</td>
<td>Sad (-2.26, 1.56)</td>
<td>.037</td>
</tr>
<tr>
<td>(500-850ms)</td>
<td></td>
<td></td>
<td>Fz</td>
<td>Happy (-1.77, 1.68)</td>
<td>Sad (-2.40, 2.02)</td>
<td>.020</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>F4</td>
<td>Happy (-1.96, 1.60)</td>
<td>Sad (-2.78, 1.97)</td>
<td>.010</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>F3</td>
<td>Happy (-2.41, 2.39)</td>
<td>Sad (-2.72, 2.35)</td>
<td>.524</td>
</tr>
</tbody>
</table>
Figure 9. Grand average difference waves averaged across emotions (in black, top), and grand average difference waves for each emotion (in colour, below). Results are shown for the four electrode sites (Cz, Fz, F4, F3). Dev – deviant, Std – standard.
4) MMRs in ‘Deviant - Standard’ waveforms for each of the four emotions

When comparing the usual mismatch configuration (deviant emotions and standard neutral), with the flipped mismatch configuration (deviant Neutral and standard emotions) - substantial differences were observed for the size and timing of MMRs suggesting that each emotion may elicit their own specific pattern of deviance. Figure 10 and Table 11 portray the deviant-standard difference waveforms for each emotion (averaged across electrode sites), with significant regions of mismatch highlighted, and latency ranges and average amplitudes reported alongside the waveform.
Table 11. 
Latency range and average amplitudes of MMRs for each emotion.

<table>
<thead>
<tr>
<th>Emotion</th>
<th>Latency window</th>
<th>Average amplitude</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neutral</td>
<td>191-299ms</td>
<td>-2.78uV</td>
</tr>
<tr>
<td></td>
<td>513-650ms</td>
<td>-1.60uV</td>
</tr>
<tr>
<td>Angry</td>
<td>100-134ms</td>
<td>1.45uV</td>
</tr>
<tr>
<td></td>
<td>163-189ms</td>
<td>-0.91uV</td>
</tr>
<tr>
<td></td>
<td>233-498ms</td>
<td>0.96uV</td>
</tr>
<tr>
<td>Happy</td>
<td>57-119ms</td>
<td>0.94uV</td>
</tr>
<tr>
<td></td>
<td>141-183ms</td>
<td>-0.80uV</td>
</tr>
<tr>
<td></td>
<td>255-302ms</td>
<td>-0.71uV</td>
</tr>
<tr>
<td></td>
<td>380-492ms</td>
<td>-0.76uV</td>
</tr>
<tr>
<td></td>
<td>543-793ms</td>
<td>-1.09uV</td>
</tr>
<tr>
<td>Sad</td>
<td>116-209ms</td>
<td>-2.42uV</td>
</tr>
<tr>
<td></td>
<td>234-273ms</td>
<td>-1.31uV</td>
</tr>
<tr>
<td></td>
<td>320-478ms</td>
<td>-2.38uV</td>
</tr>
</tbody>
</table>

Figure 10. Deviant-standard difference waveforms, averaged across electrode sites, for each emotion. The grey shaded areas indicate latency windows.
Discussion

Overall this study demonstrated that the natural speech stimuli were able to elicit measurable obligatory CAEP components and significant mismatch effects from adult participants. The four emotions presented as standard stimuli produced CAEP waveforms containing the same six components (P1c, N1c, P1v, N1v, P2, N400) indicating that the differences in the stimulus acoustic characteristics did not affect the obligatory response to the stimuli. The lack of significant variations in evoked potential amplitude between emotions for P1c, N1c, P1v, and N1v, suggests that the stimuli were successfully matched in average intensity and duration. The subsequent emotion effects observed for the P2 component suggest that the auditory evaluation of emotional expression may partially, if not entirely, rely on processing variations in pitch, since P2 is reported to be sensitive to pitch variation in auditory stimuli (Istok, 2013; Wunderlich & Cone-Wesson, 2001).

Emotion effects observed for N400 support the hypothesis that the processing of affective cues may share similar neural mechanisms to those associated with semantics (Kutas & Hillyard, 1984). In their work investigating the visual recognition and comprehension of emotional words, Kanske, Plitschka and Kotz (2011) demonstrated that changes in semantic or emotional cues both modulated N400 amplitudes. This is consistent with Holt, Lynn and Kuperberg (2009) who reported that both positive and negative words elicited significantly larger N400 magnitudes, to the same extent, compared to neutral words. For more complex stimulus paradigms involving attention to the stimuli there are varying reports of the effects of emotion on P2 magnitude (Kanske et al., 2011), as well as on a late positive response between 500 and 700ms biased towards negative words (Holt et al., 2009). Studies comparing emotional prosody and semantics using auditory stimuli have revealed greater N400 amplitudes in response to sounds with both prosodic and semantic information,
regardless of whether the experiment involved attention-to-target tasks (Wambacq & Jerger, 2004) or incongruence detection (Kotz & Paulmann, 2007). Changes in affective prosodic cues alone, separated from semantics, tend to increase positive evoked potentials more than negative ones (Kotz & Paulmann, 2007; Wambacq & Jerger, 2004). Thus, complex effects of stimulus emotion and pitch (with and without attention) have been reported in several studies of CAEPs in adult participants, and could underlie the positive MMRs observed in the current study.

The monosyllabic stimuli did evoke mismatch effects that mirror what has been reported in the literature. For the regular mismatch configuration (Emotions as deviant – Neutral as standard), MMRs 1, 3, and 4 resemble “early” and “late” mismatch components that have been similarly observed in other studies that also manipulated emotion (Korpilahti et al., 2007; Kujala et al., 2005). MMRs 3 and 4 were in close proximity to each other and hence may reflect the same underlying MMR component. However, the statistical significance testing did define these as two separate regions of contiguous deviance that were consistently observed in most, if not all, of the 20 participants. Other mismatch studies that used monosyllables like /da/ and /ga/ as stimuli (Kraus et al., 1992; Kraus et al., 1995; A. Sharma et al., 1992; M. Sharma et al., 2004), but without emotional content, have not reported these two separate later-occurring MMR components. Thus, these late MMR components could be associated with emotion change detection. Due to the sparse literature in this area, it is not possible to validate this hypothesis, and further study is needed.

Although elicited on average 100-200 ms later than MMRs from the regular mismatch configuration, the MMRs obtained from the flipped configuration have a similar negative-positive-negative MMR sequence, with more prominent differences in amplitudes between the emotions. Focusing on the first MMR from both regular and flipped mismatch configurations, there are difference when comparing how the MMRs to the three emotions
differ from each other. In the regular configuration, the sad stimulus elicited the largest negative amplitude; whereas in the flipped configuration, MMNs to angry and happy exceeded sad. Thönnessen and colleagues (2010) found that happy deviated significantly more from neutral compared to sad, but not compared to angry, which seems to corroborate what was found in the regular mismatch configuration in the current study. Fan and Cheng (2014) reported significant MMR differences between angry and happy. Overall, these findings indicate differences in MMRs across emotions but the nature of the differences vary depending on the subtraction method and the specifics of each study.

The final step of the analysis procedure explored the deviance effects for each of the four emotions. Results revealed that neutral-as-deviant resulted in the traditionally recorded early and late MMNs. However, angry, happy, and sad all showed varying regions of mismatch with little or no consistency across stimuli. One reason for the lack of consistent MMRs, especially for happy could be due to the relatively small MMR amplitudes in the adult participants and poor SNRs resulting from the method of deriving mismatch effects by subtracting two separate waveforms (Kurtzberg et al., 1995). An “optimum” experimental design has been described, aimed at improving SNR (Näätänen et al., 2004; Thönnessen et al., 2010) by having a better balance of deviant to standard stimuli (50:50, but with varying deviants within the 50 in each run), reducing the need for many test blocks in order to record sufficient responses to the deviant. In the current study, within each block, there were always three deviant stimuli to one standard, with a ratio of 70:10:10:10 for the standard and three deviants. In future work involving adult participants who have relatively small MMR amplitudes to emotional speech stimuli it may be better to use the “optimum” 50:50 ratio instead of the 70:30 ratio used in the current study. Future work would also benefit from a simplified oddball paradigm, with a decreased number of emotions being compared within each sequence, and an increased number of separate block trials. This would create a more
even distribution of trials per each emotion and could potentially better reflect differences in perception of an emotion in its standard vs. deviant form.

Further work needs to be undertaken to determine exactly what “change” is being detected when different emotions are presented. The preliminary nature of this study means that it was not possible to tease out objective acoustic factors from subjective emotional valence-arousal factors. The study did not control for external confounds such as individual reactions towards the emotional stimuli, based on personal experiences, cultural background (Chandrasekaran et al., 2007; Chandrasekaran et al., 2009), and mood (Chwilla, Virgillito & Vissers, 2011) that could significantly alter CAEPs to emotional stimuli. Differences in CAEP recordings could be better validated against an extra measure of psychological wellness, or their reliability could be strengthened by re-testing the adult participants more than one time, which was not implemented in this study due to pragmatic difficulties, and is acknowledged as a limitation to the interpretation of the results.

No systematic electrode differences were obtained, in contrast to other studies that report right hemisphere dominance for processing emotional prosody (Borod, 1992; Borod et al., 2002; Heilman, Bowers, Valenstein & Watson, 1986; Ross, 1981; 1988; 2000). Hemisphere differences may have been more apparent if the analysis had focused on the comparison of results for left (F3) versus right (F4) hemisphere electrodes. Schirmer and Kotz (2006) proposed that the processing of vocal emotion may involve multiple stages that differ in latency and neural mechanisms involved, starting with bilateral processing of sensory information, and progressing to more specialised cognitive and emotional engagement. This is consistent with the complex neural responses recorded in the current study to emotional stimuli and the differing pattern of MMR responses in early versus late latency regions. The results indicate that the mismatch paradigm can be used to determine differential processing of emotional speech stimuli, but further investigation is needed to
fully understand the complex effects of stimulus acoustics, emotion and subtraction paradigm on the results.
Chapter 7. Exploring change detection in affective prosody through cortical auditory evoked potentials and mismatch responses: Investigating differences between neurotypical controls and individuals with high-functioning Autism Spectrum Disorder, and evaluating the effects of emotion perception training and auditory amplification

This chapter is written with the intention to submit it as a manuscript for publication.
Introduction

Auditory processing difficulties in children with ASD and other neurodevelopmental conditions have been investigated both behaviourally (reviews: Haesen et al., 2011; O’Connor, 2012) and using evoked potentials. A relatively recent area of research has focused on investigating the overlap in auditory perceptual difficulties experienced by those with ASD and those with APD. For both groups, there is little or no evidence for physiological disadvantages in measures of more peripheral auditory function such as brain stem responses, otoacoustic emissions, and acoustic reflexes (Tharpe et al., 2006). Both groups however show behavioural auditory processing difficulties in the presence of background noise (Alcántara et al., 2004, Foxe et al., 2013), or when asked to recall speech stimuli that were presented in competition with each other (Brooks & Ploog, 2013; Carpenter et al., 2014), suggesting an underlying auditory filtering deficit as a result of poor or atypical sensory processing (Tomcheck & Dunn, 2007). APD-associated deficits, such as understanding and following verbal instructions and messages, localising sound, and maintaining focus (Dawes et al., 2008; de Wit et al., 2016), have been shown to significantly contribute to academic underachievement and the aggravation of anxiety, hypersensitivity, hyperactivity, and oppositional behavioural issues in children with ASD (Ashburner et al., 2008).

The recording of CAEP data shows that while neural responses for speech recognition from TD children only showed signs of degradation at high levels of background noise, responses from children with ASD already showed reduced amplitudes and increased latencies in the presence of minimal noise (Russo et al., 2009). Individuals with ASD also do not display typical increases in the size of their neural responses to increases in auditory stimulus intensity (Bruneau et al., 2003). The manifestation of ASD-related auditory neural response abnormalities in children as young as 5 years of age suggests early developmental
deficits in auditory processing, which potentially underlies the behavioural difficulties observed (Bomba & Pang, 2004; Groen et al., 2008).

The calculation of a MMR from CAEPs recorded in an oddball experimental paradigm is regarded as one of the main methods to evaluate the pre-attentive detection of changes in incoming auditory stimuli (Csépe, 1995; Näätänen, 1992; Sams et al., 1985). The MMR is typically derived by subtracting evoked potentials elicited in response to a numerously repeated “standard” stimulus, from evoked potentials elicited in response to a less frequently and randomly presented “deviant” stimulus (Näätänen, 1992). The brain’s ability to store information about multiple auditory features and react automatically to sudden changes in the acoustic environment has been attributed to highly evolved neuronal adaptations which was crucial for survival when humans had to respond quickly to a change in the auditory environment, reflecting highly evolved neuronal adaptations in the human brain (Malmierca et al., 2014; Tiitinen et al., 1994). The pre-attentive response to auditory change reflected by MMRs may be linked to higher cognitive processes like anticipation, reasoning, and learning (Näätänen et al., 2007). Figure 11, reproduced with permission from M. Sharma and colleagues (2006), depicts a typical negative MMR elicited from TD children in response to standard and deviant speech stimuli.
Figure 11. Column A shows grand average responses to standard (thick line) and deviant (thin line) /ga/ stimulus. Column B shows the difference response (deviant-standard) with a dotted line. The vertical line in Column B shows where the negative MMR offset latency was identified at 250 ms. Reproduced with permission from M. Sharma and colleagues (2006).

Although most MMRs are typically recorded in a passive listening oddball paradigm (Dunn et al., 2008), both children and adults with ASD often exhibit conflicting MMR results. Amplitudes and timing of MMRs vary considerably, depending on various factors like experimental paradigm and participant differences (Bomba & Pang, 2004; Kujala et al., 2013). Compared to the manipulation of tones, speech stimuli are also more likely to evoke varied MMR morphology (Csépe, 1995), due to the wider multitude of parameters that can be manipulated – from basic acoustic factors (Kraus et al., 1992; Kraus et al., 1995), to differences in phonemes (Aaltonen et al., 1987), articulation (Näätänen, 2001), and semantic interpretation (Näätänen et al., 2007).

In children with ASD, large negative MMRs have been recorded to deviant tones or gaps in a sequence (Ferri et al. 2003), shortened stimulus lengths (Kujala et al., 2007), changes in speech (syllabic) intensity (Kujala et al., 2010), and phonemes (Lepistö et al., 2008). Smaller amplitude MMRs are evident when detecting changes in the pitch and duration of speech-based stimuli (Kujala et al., 2010; Lepistö et al., 2006). MMRs recorded from children with HF-ASD have been shown to decrease in response to an increase in stimulus pitch (Gomot et al., 2011). The authors additionally correlated this effect with behavioural measures of an increased intolerance to changes in the auditory environment.
MMRs with later latencies have also been recorded in response to vowel changes between /a/ and /o/ (Kasai et al., 2005), and in contrast to the previous study (Gomot et al., 2011), Kasai et al. reported an association between the delayed latency effect and greater autism symptom severity.

Only a handful of studies have attempted to address the evaluation of pre-attentive emotion change detection in ASD. The lack of consensus (Peppé, 2009) over whether definitions should be categorically-based on discrete emotions (Banse & Scherer, 1996; Murray & Arnott, 1993; Yildirim et al., 2004), or dimensionally-based on valence, arousal, and intensity (Belyk & Brown, 2014; Johnstone & Scherer, 2000), may account for the reluctance to venture into manipulating emotion in CAEP experiments. As is the case for other stimuli, findings are ambiguous. Individuals with ASD are reported to either not have any MMR to deviating emotional stimuli (Fan & Cheng, 2014), or to have smaller and delayed MMRs (Kujala et al., 2005). There are also reports of significant “dual” mismatch responses from children with ASD, with an early large and long MMR, followed by a smaller, later MMR that is comparable to TD peers (Korpilahti et al., 2007).

Diminished emotion perception (Golan et al., 2006; Golan et al., 2007; Rutherford et al., 2002) and atypical auditory processing (Haesen et al., 2011; O’Connor, 2012) are prominent features of children with ASD. Encouragingly, studies have reported positive results from active interventions with individuals in this population (Golan et al., 2010; Matsuda & Yamamoto, 2013; Tse et al., 2007). Engaging children with ASD in computer-based activities has led to the successful training of increased attention towards sentence content and prosodic intonation (Ploog et al., 2009), as well as pro social skills such as collaboration, turn-taking, understanding emotions, and sharing interests (Hourcade et al., 2012). There is little consistency in experimental designs between intervention studies. Reports have been based on four case studies (Matsuda & Yamamoto, 2013), as well as
sample groups of 18-26 individuals (Golan et al., 2010; Hourcade et al., 2012; Tse et al., 2007). Some researchers were stringent in confirming the children’s ASD diagnosis above a certain evaluation cut off point (Golan et al., 2010), while others adopted an all-inclusive approach and had ASD participants ranging from very low- to very high-functioning (Hourcade et al. 2012). Some studies measured one group of children with ASD across multiple time points (Hourcade et al., 2012), whereas others included a TD control group, as well as an ASD no-intervention control group (Golan et al., 2010).

Children with ASD are also responsive to auditory speech-in-noise training (Irwin et al., 2014), and studies have reported significant enhancements in speech recognition in noise, on-task behaviours in the classroom, and teacher-rated listening behaviours (Schafer et al., 2013), as well as increased ease of communication, and lessened perceived effects of background noise (Rance et al., 2014), after trialling personal sound amplification systems. Personal amplification systems, also known as remote microphone hearing aids (RMHAs) provide better access to important acoustic signals, such as the teacher’s voice in the classroom, by placing a microphone on the speaker that transmits remotely to small receivers worn on the child’s ears. RMHAs significantly improve the signal to noise ratio (SNR) and have been shown to enhance the neural representation of speech (Hornickel, Zecker, Bradlow & Kraus, 2012).

There is substantial evidence for auditory training effects on CAEPs and MMRs, although most work has been done with normal-hearing, typically developing adults. With this population, Tremblay, Kraus, McGee, Ponton and Otis (2001) were able to demonstrate an increase in obligatory N1-P2 response amplitudes, in conjunction with improved behavioural performance, by training the discrimination of two speech stimuli that differed in voice onset time. Auditory training seems to differentially affect earlier versus later CAEP components, as Tremblay and Kraus (2002) found increased magnitudes for earlier
components with a right hemisphere bias, whereas increased magnitudes for later components are more bilaterally distributed. Improved reaction times on stimuli discrimination tasks as a result of training has also been correlated with larger P2 response components (Tong, Melara & Rao, 2009). Kraus and colleagues (1995) reported the significant elicitation, and continued growth in amplitude, of a negative MMR after training normal hearing adults to discriminate between two differing phonemic speech stimuli, whereas no detectable MMR was elicited in response to this specific combination prior to training.

Fewer studies have looked at training effects on CAEPs in children or in people with language or auditory processing difficulties or ASD. For individuals with language learning difficulties, Russo, Nicol, Zecker, Hayes and Kraus (2005) demonstrated improved neural synchrony post-training, that enabled more efficient encoding of complex sounds. Changes in CAEP amplitude after language and discrimination training have also been reported for children with APD (M. Sharma, Purdy & Kelly, 2014). There is need for more work that addresses the functional impact of auditory training in clinical populations such as children with ASD, and evaluates training effects on neural responses as a potential reflection of the plasticity of the central auditory system.

This study’s primary aim was to use natural speech stimuli produced with angry, happy, sad, and neutral emotion in an oddball paradigm to investigate differences in mismatch responses in children with HF-ASD compared typically developing controls. MMRs were compared to recordings from control group children and adults to examine the effects of maturation and emotion on response. Peak amplitudes and latencies of the obligatory cortical components P1, N1, P2, and the N400 response were also investigated. MMRs were derived by subtracting standard from deviant responses, and average mismatch amplitudes were determined for the latency ranges with significant MMRs. A secondary aim was to evaluate the effects of a 3-week training intervention for the children with HF-ASD.
The training consisted of computer-based emotion perception training activities combined with a trial of RMHAs. The impact of the training on neurophysiological responses was investigated.

Adults were compared to TD children controls, TD children were compared to HF-ASD children before intervention, and HF-ASD children’s results before and after intervention were compared. It was anticipated that contrasting control adults with children would illustrate maturational variations in latencies and amplitudes of obligatory CAEP components (Ponton et al., 2002; Wunderlich et al., 2006), as well as MMRs (Gomot et al., 2000; Kraus et al., 1992; Shafer et al., 2000).

It was hypothesised that CAEPs and MMRs would also differ between TD children and children with HF-ASD based on results from other studies demonstrating impaired behavioural (Peppé et al., 2007) and neurophysiological (Fan & Cheng, 2014; Kujala et al., 2005) responses to affective stimuli in children with ASD. Reduced amplitude MMR for deviant emotions, and reduced differences between emotions were anticipated for children with HF-ASD.

It was hypothesised that CAEPs and MMRs of children with HF-ASD would show training-related changes and that the computer training in emotion perception, combined with exposure to an enriched acoustic environment through the use of the RMHA system would enhance the neurophysiological distinction between deviant emotions (angry, happy, and sad) compared to the standard (neutral) stimulus.
Methods

Participants

Out of the 12 children with HF-ASD who participated in the RMHA trial, two did not assent to the evoked potential recording sessions, therefore results are only available for 10 of the 12 children with HF-ASD (8 males, 2 females, $M_{age}=10.3$ years, $SD_{age}=2.21$). There were 14 TD children who participated as TD controls (4 males, 10 females, $M_{age}=9.43$ years, $SD_{age}=1.87$). In addition, the 20 adults (10 males, 10 females, $M_{age}=27$ years, $SD_{age}=3.84$) from the previous study (Chapter 6) were also included for comparison in the analysis. The child participants all passed a distortion product otoacoustic emissions (DPOAE) screen, performed using the Grason-Stadler GSI Audioscreener (version 3.21), which screens for cochlear and middle ear pathology that may lead to physiologically impaired hearing. The passing criteria was to obtain a SNR above 6 dB across five frequency levels, from 2 – 6 kHz. Mean SNR measures from the right and left ears for both participant groups are reported in Table 12. The adult participants self-reported no past or current history of hearing concerns.

Table 12. DPOAE results (signal-to-noise ratios, SNR) measured in decibels (dB) from TD and HF-ASD children. DPOAEs were collected from the left and right ears of each participant, across five frequencies ranging from 2-6 kHz.

<table>
<thead>
<tr>
<th>Frequency (kHz)</th>
<th>TD SNR (dB)</th>
<th>HF-ASD SNR (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Left ear</td>
<td>Right ear</td>
</tr>
<tr>
<td>2</td>
<td>16.60</td>
<td>15.70</td>
</tr>
<tr>
<td>3</td>
<td>23.22</td>
<td>19.25</td>
</tr>
<tr>
<td>4</td>
<td>16.94</td>
<td>15.09</td>
</tr>
<tr>
<td>5</td>
<td>12.85</td>
<td>23.39</td>
</tr>
<tr>
<td>6</td>
<td>25.88</td>
<td>22.68</td>
</tr>
</tbody>
</table>
CAEPs were recorded once for each of the TD children and adult control participants. The children with HF-ASD, were tested twice, once before and once after a 3-week training program which consisted of computer-based emotion perception training, coupled with a trial of a RMHA system in school. The RMHA system was also worn when undertaking computer-training tasks which were completed at home. The computer-based emotion training included tasks which targeted both facial expressions and affective prosody.

Cortical auditory evoked potential recordings

Stimuli and sequences

The speech stimuli were sampled from existing recordings of monosyllables produced by a male speaker in angry, happy, neutral, and sad emotional tones of voice (Ross et al., 1991). Details of the sampling and editing method and the selection of the four speech stimuli are described in Chapter 5. All the speech stimuli (/ba/) were 200 ms in length. The time waveforms in Figure 12 show that the speech stimuli were, on average, matched for root mean square intensity and duration but differed in their temporal characteristics.

An oddball paradigm was used to investigate auditory affective prosodic discrimination using the MMR. The stimuli were programmed into sequence blocks using the NeuroScan STIM2 Gentask software. Each block consisted of 216 speech stimuli in a ratio of 70 standard stimuli (156 stimuli) to 10:10:10 deviant stimuli (20 per emotion). In total, there were 1092 Neutral stimuli, and 140 stimuli for each of the emotions.

Every sequence began with 20 standard stimuli, followed by a pseudorandom presentation of standard and deviant stimuli. The order of emotional deviants inserted into the sequence was randomised, but was adjusted so that at least 2 or 3 standards occurred between each deviant. There was a 640 ms inter-stimulus interval (ISI).
Figure 12. Electrical time waveforms (Adobe Audition CS6) depicting the monosyllabic stimulus /ba/ stimulus presented in four emotions. dBV represents relative sound intensity in decibel-voltages, plotted against time in milliseconds

**Experimental setup**

The CAEP recordings took place in a sound treated two-room setup, with a leather reclining chair for the participants to sit in. Using the Gentask software on NeuroScan STIM2, the stimuli sequences were presented at 70 dB SPL via an Australian Monitor Synergy SY400 power amplifier and Sabine Graphi-Q GRQ-3102 equaliser, connected to a Turbosound IMPACT 50 loudspeaker. A half-inch polarised condenser free-field microphone, connected to a Bruel and Kjaer measuring amplifier and oscilloscope, was used to calibrate and externally monitor the sound levels of the stimuli in the enclosed testing environment.

The loudspeaker was positioned directly in front of the participant seated on the recliner, approximately 150 cm away (dependent on individual size). Situated behind the
A loudspeaker was a television on a stand. Participants were instructed to watch a movie of their choice with the audio turned off and the subtitles on. This was to ensure their alertness and cooperation, while their passive evoked responses to the stimuli were being recorded. Participants were also asked to minimise their blinking and body movements during the experiment.

The CAEPs were recorded using the NeuroScan Inc. Evoked Potential System (version 4.5) with a SynAmps 2 amplifier. The surface of the participant’s scalp was cleaned using NuPrep EEG and ECG skin prep gel, and eight 10mm gold electrodes were placed on Cz, Fz, F3, F4, A1, and A2 locations, with a ground electrode on the forehead, and an eye blink electrode above the right eye. The electrode on the right mastoid (A2) served as the reference electrode. During offline processing, the left and right mastoid electrodes were linked, and Cz, Fz, F3, and F4 were subsequently re-referenced. The electrodes were filled with Quik-Gel™ conductive gel (Neuromedical supplies) and secured onto the scalp with Transpore surgical tape. Electrode impedance was kept at or below 5kΩ. Recordings were made with a sampling rate of 500 Hz and a bandpass filter setting of 0.1-100 Hz.

Data processing

Further offline processing was done using the Edit software from NeuroScan Inc. Continuous recording files were epoched from -100 ms pre-stimulus to 850 ms post-stimulus, followed by baseline correction. Any responses exceeding ±150 uV were rejected as artifacts. A minimum of 20 blinks were required to estimate an average blink, and to avoid double detection the ocular artifact rejection procedure discarded anything that occurred <400 ms before a previous one. The data was filtered using a low-pass at 30 Hz (12 dB/octave slope, zero phase shift). Separate average files were generated for each participant for each of the four emotions.
On average, about 15% of responses were rejected for the TD adult participants, and 20% of responses were rejected from the TD children participants, as a result of ocular and other noise artifacts. For the children with HF-ASD, an average of about 40% of responses were rejected. Their over-sensitivity to sounds, easy irritableness towards the electrodes placed on their scalp and skin, and other behavioural issues, contributes to comparatively noisy ERP recordings. Although accepted as a natural consequence of working with clinical populations, this factor is acknowledged as a limitation of the study, and future protocol parameters would benefit from further adjustments to cater for high noise conditions.

All grand average waveforms were plotted from -100 ms to 850 ms to encompass pre-stimulus responses, the 200 ms stimulus length, and post-stimulus responses for the duration of the 640 ms ISI, without overlapping with the subsequent stimulus.

Data analysis

Independent samples t-tests were conducted to compare results for adults with TD children, and to compare TD with HF-ASD children. Paired samples t-tests were conducted to compare the recordings of HF-ASD children before and after the intervention. Each comparison was undertaken as follows to answer the three questions:

1) Are there significant differences in peak amplitudes and latencies between groups for each identified obligatory CAEP component?

Firstly, the researcher generated two grand average waveforms for each participant group. One was the evoked potential in response to the neutral-sounding monosyllable presented as a standard stimulus. The other comprised an average of the angry, happy, and sad deviant responses.

Next, the researcher used the grand average waveforms as reference waveforms to pick peak amplitudes and latencies from each participant. As reported in the previous chapter,
adult control participants exhibited six peaks that corresponded to a P1 component reflecting the onset of the /b/ consonant (P1c), an N1 consonant component (N1c), a P1 component reflecting the onset of the /a/ vowel (P1v), and its N1 equivalent (N1v), a P2 component, and an N400 component. However, evoked responses from the two groups of children failed to display comparable components for P1c, N1c, and P2. Instead their waveforms were characterised by a single positive peak, and slight negative trough, and a significant N400. Consequently, between-group comparisons were only computed for the P1v, N1v, and N400 components, at each electrode site (Cz, Fz, F4, F3), which were consistently detected in all participants.

2) Are there significant differences in the magnitude, polarity (negative versus positive mismatch), and the timing of MMRs between groups?

For each participant group, the researcher first subtracted the grand averaged standard waveform (neutral) from the grand averaged deviant waveform (combination of angry, happy, and sad). This was done for recordings from each electrode site.

Single-sample t-tests were conducted on these difference waveforms at each millisecond. Contiguous time periods were highlighted where the waveform deviated at a significance level of $p \leq 0.001$ from 0 uV across all electrodes. These formed MMR latency windows, which are listed in Table 16 in a timeline format in the Results section, with the times rounded up to the nearest multiple of ten. Mismatches consisted of both positive and negative deviances, and ranged from 40 ms windows to longer periods, which were considered as late discriminative negativities (LDN). LDN components are thought to reflect the processing of more complex auditory stimuli, especially with regards to language and speech processing (Cheour et al., 2001; Halliday et al., 2014; Korpilahti et al., 2001). Long deviance periods such as these were clearly identified in both groups of children in this study.
The latency windows were subsequently used as references for the picking of individual MMR peak latencies for the grand averaged difference wave from each participant. MMR magnitudes were calculated by taking an average of the amplitudes included within ±20 ms either side of the peak latency, or by ±50 ms either side of an approximate midpoint for the LDNs (Luck, 2005). Repeated measures ANOVAs were conducted on average amplitude and peak latencies within each MMR latency region to explore differences between recordings from each electrode site (Cz, Fz, F4, and F3). No significant results were found, and consequently results were collapsed across electrodes before commencing the between-group analyses.

3) Are there significant differences in the magnitude and timing of MMRs between angry, happy, and sad? How do these differ between participant groups?

The MMR latency windows identified above were also used as reference points for the three separate difference waveforms generated by subtracting the ‘neutral as standard’ waveform from each ‘emotion as deviant’ waveform. Once again, peak latencies were picked, and MMR averaged amplitudes were calculated for each participant.

Within-subject repeated measures ANOVAs were conducted for each MMR - with three levels of emotional deviance collapsed across four electrode sites. This was followed by Bonferroni-corrected post-hoc analyses. The presence of significant emotion effects was compared between groups for the different regions of mismatch.
Results

1) CAEP components from ‘neutral as standard’ and ‘emotions as deviants’ waveforms

Adult controls vs. TD children controls

Peak amplitudes and latencies for P1, N1, and N400 components were compared between adult and TD children controls. The hypothesised maturational differences in evoked responses were evident for both deviant and standard stimuli, shown statistically by the independent t-tests (Tables 13 and 14) and in the comparison of grand average waveforms (Figure 13). With the exception of N1 amplitude, differences between adult and child latencies and amplitudes were evident across the four electrode locations. In general, P1 and N400 were larger and earlier and N1 was larger and later in the TD children compared to adults.
Table 13.
Independent samples t-test results showing significant differences between typically developing adults and children in CAEP response component amplitudes and latencies, measured for the neutral-standard stimulus condition.

<table>
<thead>
<tr>
<th>CAEP Component (Amplitude/Latency)</th>
<th>Electrode Site</th>
<th>Adult Controls</th>
<th>TD Children Controls</th>
<th>t(32)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>P1 - Amplitude</td>
<td>Cz</td>
<td>2.37</td>
<td>1.19</td>
<td>5.11</td>
<td>3.38</td>
</tr>
<tr>
<td></td>
<td>Fz</td>
<td>2.88</td>
<td>1.55</td>
<td>6.16</td>
<td>3.32</td>
</tr>
<tr>
<td></td>
<td>F4</td>
<td>2.69</td>
<td>1.22</td>
<td>6.44</td>
<td>2.79</td>
</tr>
<tr>
<td></td>
<td>F3</td>
<td>2.70</td>
<td>1.32</td>
<td>6.04</td>
<td>3.21</td>
</tr>
<tr>
<td>N1 - Amplitude</td>
<td>F4</td>
<td>-2.39</td>
<td>2.30</td>
<td>-4.72</td>
<td>3.00</td>
</tr>
<tr>
<td>N1 - Latency</td>
<td>Cz</td>
<td>187.45</td>
<td>9.06</td>
<td>219.00</td>
<td>5.46</td>
</tr>
<tr>
<td></td>
<td>Fz</td>
<td>185.10</td>
<td>8.14</td>
<td>217.09</td>
<td>6.43</td>
</tr>
<tr>
<td></td>
<td>F4</td>
<td>184.05</td>
<td>10.45</td>
<td>218.75</td>
<td>7.98</td>
</tr>
<tr>
<td></td>
<td>F3</td>
<td>182.45</td>
<td>13.16</td>
<td>218.55</td>
<td>8.10</td>
</tr>
<tr>
<td>N400 - Amplitude</td>
<td>Cz</td>
<td>-3.24</td>
<td>1.38</td>
<td>-8.58</td>
<td>3.68</td>
</tr>
<tr>
<td></td>
<td>Fz</td>
<td>-3.94</td>
<td>1.35</td>
<td>-10.06</td>
<td>4.25</td>
</tr>
<tr>
<td></td>
<td>F4</td>
<td>-3.25</td>
<td>0.95</td>
<td>-9.69</td>
<td>4.35</td>
</tr>
<tr>
<td></td>
<td>F3</td>
<td>-3.61</td>
<td>1.27</td>
<td>-10.14</td>
<td>3.96</td>
</tr>
<tr>
<td>N400 - Latency</td>
<td>Cz</td>
<td>450.05</td>
<td>20.04</td>
<td>431.14</td>
<td>16.04</td>
</tr>
<tr>
<td></td>
<td>Fz</td>
<td>455.60</td>
<td>17.51</td>
<td>435.57</td>
<td>13.97</td>
</tr>
<tr>
<td></td>
<td>F4</td>
<td>452.55</td>
<td>21.32</td>
<td>431.00</td>
<td>16.03</td>
</tr>
<tr>
<td></td>
<td>F3</td>
<td>458.25</td>
<td>17.39</td>
<td>431.93</td>
<td>12.26</td>
</tr>
</tbody>
</table>

166
Table 14.
Independent samples t-test results showing significant differences between typically developing adults and children in CAEP response component amplitudes and latencies, measured for the emotional-deviant stimulus conditions (combined across angry, happy, and sad CAEP recordings).

<table>
<thead>
<tr>
<th>CAEP Component (Amplitude/Latency)</th>
<th>Electrode Site</th>
<th>Adult Controls</th>
<th>TD Children Controls</th>
<th>( t(32) )</th>
<th>( p )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>( M )</td>
<td>( SD )</td>
<td>( M )</td>
<td>( SD )</td>
</tr>
<tr>
<td>P1 - Amplitude</td>
<td>Cz</td>
<td>2.22</td>
<td>1.85</td>
<td>4.69</td>
<td>2.28</td>
</tr>
<tr>
<td></td>
<td>Fz</td>
<td>2.96</td>
<td>2.17</td>
<td>5.72</td>
<td>2.67</td>
</tr>
<tr>
<td></td>
<td>F4</td>
<td>2.59</td>
<td>2.06</td>
<td>5.91</td>
<td>2.60</td>
</tr>
<tr>
<td></td>
<td>F3</td>
<td>2.81</td>
<td>2.10</td>
<td>5.62</td>
<td>2.48</td>
</tr>
<tr>
<td>P1 - Latency</td>
<td>Cz</td>
<td>147.35</td>
<td>18.63</td>
<td>122.36</td>
<td>9.80</td>
</tr>
<tr>
<td></td>
<td>Fz</td>
<td>152.90</td>
<td>12.86</td>
<td>125.64</td>
<td>7.02</td>
</tr>
<tr>
<td></td>
<td>F4</td>
<td>149.65</td>
<td>14.05</td>
<td>126.00</td>
<td>8.55</td>
</tr>
<tr>
<td></td>
<td>F3</td>
<td>155.35</td>
<td>15.20</td>
<td>125.43</td>
<td>7.01</td>
</tr>
<tr>
<td>N1 - Amplitude</td>
<td>Cz</td>
<td>-1.04</td>
<td>2.21</td>
<td>-2.66</td>
<td>1.79</td>
</tr>
<tr>
<td></td>
<td>Fz</td>
<td>-0.91</td>
<td>2.38</td>
<td>-3.42</td>
<td>2.53</td>
</tr>
<tr>
<td></td>
<td>F4</td>
<td>-0.89</td>
<td>2.28</td>
<td>-3.97</td>
<td>2.43</td>
</tr>
<tr>
<td></td>
<td>F3</td>
<td>-0.87</td>
<td>2.35</td>
<td>-3.34</td>
<td>2.64</td>
</tr>
<tr>
<td>N1 - Latency</td>
<td>Cz</td>
<td>198.10</td>
<td>14.56</td>
<td>225.57</td>
<td>8.07</td>
</tr>
<tr>
<td></td>
<td>Fz</td>
<td>203.40</td>
<td>11.99</td>
<td>224.14</td>
<td>6.67</td>
</tr>
<tr>
<td></td>
<td>F4</td>
<td>198.45</td>
<td>12.53</td>
<td>223.36</td>
<td>8.65</td>
</tr>
<tr>
<td></td>
<td>F3</td>
<td>201.90</td>
<td>10.93</td>
<td>220.93</td>
<td>6.06</td>
</tr>
<tr>
<td>N400 - Amplitude</td>
<td>Cz</td>
<td>-2.41</td>
<td>1.25</td>
<td>-7.06</td>
<td>3.49</td>
</tr>
<tr>
<td></td>
<td>Fz</td>
<td>-3.19</td>
<td>1.25</td>
<td>-8.83</td>
<td>3.89</td>
</tr>
<tr>
<td></td>
<td>F4</td>
<td>-2.80</td>
<td>1.00</td>
<td>-8.62</td>
<td>3.90</td>
</tr>
<tr>
<td></td>
<td>F3</td>
<td>-2.83</td>
<td>1.27</td>
<td>-8.22</td>
<td>3.99</td>
</tr>
<tr>
<td>N400 - Latency</td>
<td>Cz</td>
<td>483.85</td>
<td>27.23</td>
<td>418.86</td>
<td>16.97</td>
</tr>
<tr>
<td></td>
<td>Fz</td>
<td>482.80</td>
<td>21.56</td>
<td>421.71</td>
<td>21.16</td>
</tr>
<tr>
<td></td>
<td>F4</td>
<td>486.15</td>
<td>19.07</td>
<td>415.36</td>
<td>20.78</td>
</tr>
<tr>
<td></td>
<td>F3</td>
<td>486.40</td>
<td>20.56</td>
<td>418.14</td>
<td>13.57</td>
</tr>
</tbody>
</table>
Figure 13. CAEP waveforms from typically developing adults and children, for both neutral-as-standard and emotions-as-deviant stimulus presentation conditions. Response potentials are measured in microvolts (μV) and time in milliseconds (ms), across four electrode sites (Cz, Fz, F4, and F3).
**TD Children controls vs. HF-ASD Children (Pre-intervention)**

Peak amplitudes and latencies for P1, N1, and N400 components were compared using independent t-tests between TD children and HF-ASD children before they received the intervention. The waveforms were generally in line with each other, which supports the view that children with ASD seem to process auditory stimuli in a similar way to their TD peers, and the difference lies in the strength and rapidness of their neural responses. The groups showed equal variance based on the non-significant results from Levene’s tests. Figure 14 shows grand average CAEP waveforms for each group elicited in response to neutral as standard, and the combination of the three emotions as deviants.

For the neutral as standard condition, TD children ($M=125.64, SD=7.02$) had a significantly earlier P1 component at Fz ($t_{(22)}=-2.72, p=.012$) compared to HF-ASD children ($M=136.50, SD=12.47$). The N1 component was also significantly earlier for TD controls at both Fz ($t_{(14)}=-2.83, p=.013; M_{(TD)}=217.09, SD_{(TD)}=6.43); M_{(ASD)}=230.00, SD_{(ASD)}=12.10$) and F3 ($t_{(14)}=-2.52, p=.024; M_{(TD)}=218.55, SD_{(TD)}=8.10); M_{(ASD)}=230.00, SD_{(ASD)}=9.17$) electrode sites. In addition, the N1 responses elicited in TD children were significantly larger compared to HF-ASD children at both Cz ($t_{(18)}=-4.14, p=.001; M_{(TD)}=-2.66, SD_{(TD)}=1.79); M_{(ASD)}=0.70, SD_{(ASD)}=1.29$) and Fz ($t_{(19)}=-2.11, p=.048; M_{(TD)}=-3.42, SD_{(TD)}=2.53); M_{(ASD)}=-1.14, SD_{(ASD)}=1.84$) electrode sites. Lastly, the only significant difference in the N400 component was the earlier latency for TD children ($M=415.36, SD=20.78$) compared to children with HF-ASD ($M=435.60, SD=19.88$), at F4 ($t_{(22)}=-2.40, p=.026$). The F4 electrode site was also sensitive to N1 amplitude differences between TD and HF-ASD groups for the neutral as standard stimulus.

For the combined emotions as deviant condition, there were no significant differences in the size and timing of the P1 component between groups. However, N1 responses at F3
were significantly earlier ($t(19)=-2.12, p=.047$) for TD children ($M=220.93, SD=6.06$) compared to the children with HF-ASD ($M=228.57, SD=10.61$). In addition, TD children consistently displayed larger and later N400 components. With regards to N400 amplitude, significant differences between the two groups were recorded at Cz ($t(22)=-2.26, p=.034$; $M_{(TD)}=-8.58, SD_{(TD)}=3.68$; $M_{(ASD)}=-5.51, SD_{(ASD)}=2.60$); Fz ($t(22)=-2.79, p=.011$; $M_{(TD)}=-10.06, SD_{(TD)}=4.25$; $M_{(ASD)}=-5.95, SD_{(ASD)}=2.16$); F4 ($t(22)=-3.09, p=.005$; $M_{(TD)}=-9.69, SD_{(TD)}=4.35$; $M_{(ASD)}=-5.03, SD_{(ASD)}=2.28$), and F3 ($t(22)=-3.17, p=.004$; $M_{(TD)}=-10.14, SD_{(TD)}=3.96$; $M_{(ASD)}=-5.50, SD_{(ASD)}=2.80$). For N400 latency, significant differences were observed at Cz ($t(22)=2.12, p=.046$; $M_{(TD)}=431.14, SD_{(TD)}=16.04$; $M_{(ASD)}=415.80, SD_{(ASD)}=19.41$); Fz ($t(22)=2.57, p=.017$; $M_{(TD)}=435.57, SD_{(TD)}=13.97$; $M_{(ASD)}=421.00, SD_{(ASD)}=13.26$), and F3 ($t(22)=2.07, p=.050$; $M_{(TD)}=431.93, SD_{(TD)}=12.26$; $M_{(ASD)}=418.20, SD_{(ASD)}=20.19$).
Figure 14. CAEP waveforms from TD children and HF-ASD children (recorded before intervention), for both neutral-as-standard and emotions-as-deviant stimulus presentation conditions. Response potentials are measured in microvolts (uV) and time in milliseconds (ms), across four electrode sites (Cz, Fz, F4, and F3).
**HF-ASD Children – Pre and Post intervention**

Paired samples t-tests were conducted to compare the peak amplitudes and latencies for P1, N1, and N400 components between recordings obtained before and after the intervention in children with HF-ASD. There were no significant changes in the size and timing of components elicited in response to neutral-sound monosyllables presented as a standard stimulus. However, large differences can be seen, especially for N1 and N400, in responses towards deviant stimuli. Results of these changes are reported statistically in Table 15, and are illustrated in Figure 15.

Table 15. 
*Paired samples t-test results showing significant differences between HF-ASD children before and after they received the intervention in CAEP response component amplitudes and latencies, measured in the emotional-sounding deviant stimuli condition (combined across Angry, Happy, and Sad CAEP recordings).*

<table>
<thead>
<tr>
<th>CAEP Component (Amplitude/Latency)</th>
<th>Electrode Site</th>
<th>CAEP Component Pre intervention</th>
<th>CAEP Component Post intervention</th>
<th>t(18)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td><strong>M</strong></td>
<td><strong>SD</strong></td>
<td><strong>M</strong></td>
<td><strong>SD</strong></td>
</tr>
<tr>
<td>N1 - Amplitude</td>
<td>Cz</td>
<td>-2.56</td>
<td>2.03</td>
<td>-4.10</td>
<td>3.04</td>
</tr>
<tr>
<td></td>
<td>Fz</td>
<td>-2.54</td>
<td>2.46</td>
<td>-5.36</td>
<td>3.06</td>
</tr>
<tr>
<td></td>
<td>F4</td>
<td>-3.88</td>
<td>3.96</td>
<td>-6.36</td>
<td>3.15</td>
</tr>
<tr>
<td></td>
<td>F3</td>
<td>-4.50</td>
<td>4.78</td>
<td>-6.43</td>
<td>3.17</td>
</tr>
<tr>
<td>N400 - Amplitude</td>
<td>Cz</td>
<td>-5.51</td>
<td>2.60</td>
<td>-10.89</td>
<td>2.47</td>
</tr>
<tr>
<td></td>
<td>Fz</td>
<td>-5.95</td>
<td>2.16</td>
<td>-12.50</td>
<td>3.49</td>
</tr>
<tr>
<td></td>
<td>F4</td>
<td>-5.03</td>
<td>2.28</td>
<td>-11.27</td>
<td>3.91</td>
</tr>
<tr>
<td></td>
<td>F3</td>
<td>-5.50</td>
<td>2.80</td>
<td>-13.42</td>
<td>3.33</td>
</tr>
<tr>
<td>N400 - Latency</td>
<td>Cz</td>
<td>419.60</td>
<td>10.36</td>
<td>460.58</td>
<td>11.50</td>
</tr>
<tr>
<td></td>
<td>Fz</td>
<td>421.00</td>
<td>13.26</td>
<td>463.10</td>
<td>12.30</td>
</tr>
<tr>
<td></td>
<td>F4</td>
<td>427.50</td>
<td>14.01</td>
<td>465.10</td>
<td>9.89</td>
</tr>
<tr>
<td></td>
<td>F3</td>
<td>418.20</td>
<td>20.19</td>
<td>468.70</td>
<td>11.60</td>
</tr>
</tbody>
</table>
Figure 15. CAEP waveforms from HF-ASD children recorded before and after intervention, for both neutral-as-standard and emotions-as-deviant stimulus presentation conditions. Response potentials are measured in microvolts (uV) and time in milliseconds (ms), across four electrode sites (Cz, Fz, F4, and F3).
2) **Magnitude, polarity, and timing of MMRs between participant groups**

Single sample t-tests conducted at each millisecond revealed that each participant group demonstrated different patterns of mismatch effects in response to deviating emotional auditory stimuli. Due to the novel application of the stimuli used for the MMR recordings, the next step in the analysis was to describe the variations in MMR magnitude, polarity, and timing between adults, TD children, and children with HF-ASD before and after the intervention. Figure 16 illustrates the combined difference waveforms, collapsed across electrodes, for each participant group. The adult difference waveform (black line) shows the expected early mismatch negativity at about 150 ms. Larger MMNs are evident in the difference waveforms for the child participants. Table 16 lists the latency windows that defined each of the MMRs in a timeline format. These are numbered as regions of significant MMR within each group. Results from the between group comparisons of average amplitudes and latencies are reported in Table 17.

It can be observed that ERP traces do not return to baseline before the elicitation of the new stimulus. This is more evident for the children, and especially for those with HF-ASD (see figures 13, 15 and 16). This could be attributed to possible contingent negative variation, where the participant elicits unconscious anticipation of incoming stimuli, however without having epoched the response to include a longer baseline (200 ms for a full alpha cycle), this cannot be verified. To remedy this, future extensions of this work would benefit from introducing jitter into the inter-stimulus intervals between presentations.
Figure 16. Group average difference waveforms (neutral-as-standard minus emotions-as-deviants), combined across emotions and electrode sites, resulting in a single grand average difference waveform for each group – typically developing adults, children, and HF-ASD children before and after intervention. Response amplitudes are measured in microvolts (uV) and time in milliseconds (ms).
Table 16. Timeline format presentation of mismatch latency windows for each participant group, derived by conducting single t-tests at each time point of group average emotion-combined difference waveforms, and identifying contiguous periods of significant deviance from 0 uV.

<table>
<thead>
<tr>
<th>Participant Group</th>
<th>MMR1</th>
<th>MMR2</th>
<th>MMR3</th>
<th>MMR4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adult Controls</td>
<td>130-170 ms (Negative)</td>
<td>210-250 ms (Positive)</td>
<td>260-300 ms (Negative)</td>
<td>380-420 ms (Negative)</td>
</tr>
<tr>
<td>TD Children Controls</td>
<td>70-110 ms (Positive)</td>
<td>150-200 ms (Negative)</td>
<td>270-310 ms (Negative)</td>
<td>550-650 ms (LDN)</td>
</tr>
<tr>
<td>HF-ASD Pre-intervention</td>
<td>210-310 ms (Negative)</td>
<td></td>
<td></td>
<td>540-840 ms (LDN)</td>
</tr>
<tr>
<td>HF-ASD Post-intervention</td>
<td>180-220 ms (Negative)</td>
<td>340-380 ms (Positive)</td>
<td></td>
<td>490-790 ms (LDN)</td>
</tr>
</tbody>
</table>
Table 17. Independent samples t-test, and paired samples t-test results showing significant differences in MMR average amplitude and peak latency for three between group comparisons. Average amplitude is measured in microvolts (μV) and latency in milliseconds (ms). The numbering of the MMR is based on Table 16.

### Adult Controls vs. TD Children Controls

<table>
<thead>
<tr>
<th>MMR from Adults</th>
<th>MMR from TD Children</th>
<th>Amplitude/Latency</th>
<th>t(32)</th>
<th>p</th>
<th>Adults M</th>
<th>SD</th>
<th>TD Children M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>Latency</td>
<td>-4.89</td>
<td>&lt;.001</td>
<td>146.75</td>
<td>35.44</td>
<td>99.57</td>
<td>6.94</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>Amplitude</td>
<td>4.11</td>
<td>&lt;.001</td>
<td>1.45</td>
<td>0.99</td>
<td>3.35</td>
<td>1.71</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>Latency</td>
<td>-2.82</td>
<td>.008</td>
<td>218.00</td>
<td>52.48</td>
<td>177.71</td>
<td>11.17</td>
</tr>
</tbody>
</table>
It was hypothesised that there would be maturational differences in the elicitation of MMRs between adults and children, but due to minimal literature around mismatch effects in response to affective prosodic discrimination, no concrete predictions were made. Both TD adults and children exhibit four MMRs at different times after stimulus presentation. The polarities of the first two MMRs are reversed between adults and children, and aside from the first MMR which was not significantly different in average amplitude, the following three MMRs all showed that TD children had consistently larger responses.

It was hypothesised that MMRs would vary minimally in timing between TD and HF-ASD children, because of their comparable ages. Based on earlier MMR studies of children with ASD (e.g. Kujala et al., 2010), greater differences in MMR amplitudes were expected. However, both timing and amplitude of the mismatch effect differed between the groups. The waveforms of HF-ASD children were characterised by two large negativities (referred to as MMR1 and MMR2 in Table 16). The first one was significantly different in size and timing to MMRs 2 and 3 from TD peers, occurring somewhat in between the two. In addition, the LDN from the children with HF-ASD was smaller and later compared to the TD children.

Lastly, it was hypothesised that MMRs elicited by the HF-ASD children after they have participated in the intervention would exhibit enhanced amplitudes in response to changes in emotion. This was not evident in the results. Instead, comparisons of latencies for the first and last MMRs showed that post-intervention, the children’s responses occurred significantly earlier, suggesting more rapid or automatic change detection. Post-intervention recordings also revealed that the HF-ASD children now exhibit a large positive MMR at around 360 ms that is distinctly different in timing compared to the large negativity from the pre-intervention waveform.
3) **Emotion effects within each MMR between participant groups**

Significant differences were found between the deviant emotions angry, happy, and sad within each MMR latency region (Table 18 and Figure 17). Average amplitudes from each emotion have been converted into their absolute value (removing the negative sign). Due to the absence of significant electrode effects from the within-subject repeated measures ANOVAs, waveforms in Figure 17 represent difference waves collapsed across electrodes.

In both adult and children control groups, sad speech consistently evoked larger negative mismatch responses in the first stages of processing compared to angry and happy. In Figure 17, sad also seems to evoke a larger and earlier MMR in the children with HF-ASD before intervention. However, there is a considerable amount of variation in amplitude and latency between individual participants, and collectively the group does not show a statistically significant difference in amplitude. Post-intervention, the children with HF-ASD do display a significant separation of MMRs attributed to the different emotions, which suggests strong responses as well as reduced group variance.
Table 18. Within-subject repeated measures ANOVA results showing significant differences in amplitude and latency between each emotional deviance (angry, happy, and sad) within identified MMR regions from each participant group. Average amplitude is measured in microvolts (μV) and latency in milliseconds (ms). Bonferroni adjusted $p<.017$ was applied to post hoc result interpretations.

<table>
<thead>
<tr>
<th>Participant Group</th>
<th>MMR (Amplitude/Latency)</th>
<th>F (df, dfe)</th>
<th>$p$-value</th>
<th>Post-hoc Comparisons</th>
<th>Emotion 1 (M, SD)</th>
<th>Emotion 2 (M, SD)</th>
<th>$p$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adult Controls</td>
<td>MMR1 (Amplitude)</td>
<td>9.47 (2,57)</td>
<td>&lt;.001</td>
<td>Angry</td>
<td>(1.78, 1.21)</td>
<td>Sad</td>
<td>.006</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Happy</td>
<td>(1.69, 1.39)</td>
<td>Sad</td>
<td>.003</td>
</tr>
<tr>
<td></td>
<td>MMR2 (Latency)</td>
<td>3.72 (2,57)</td>
<td>.033</td>
<td>Happy</td>
<td>(205.34, 49.14)</td>
<td>Sad</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>TD Children</td>
<td>MMR1 (Latency)</td>
<td>8.64 (2, 39)</td>
<td>.001</td>
<td>Happy</td>
<td>(106.36, 7.63)</td>
<td>Sad</td>
<td>.006</td>
</tr>
<tr>
<td></td>
<td>MMR3 (Amplitude)</td>
<td>2.35 (2, 39)</td>
<td>.038</td>
<td>Angry</td>
<td>(5.06, 3.24)</td>
<td>Sad</td>
<td>.009</td>
</tr>
<tr>
<td>HF-ASD (Pre)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>HF-ASD (Post)</td>
<td>MMR1 (Amplitude)</td>
<td>28.71 (2, 27)</td>
<td>&lt;.001</td>
<td>Angry</td>
<td>(7.14, 2.88)</td>
<td>Happy</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Happy</td>
<td>(3.85, 2.01)</td>
<td>Sad</td>
<td>.009</td>
</tr>
<tr>
<td></td>
<td>MMR3 (Amplitude)</td>
<td>22.43 (2, 27)</td>
<td>.001</td>
<td>Angry</td>
<td>(9.20, 4.96)</td>
<td>Happy</td>
<td>.006</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Angry</td>
<td>(9.20, 4.96)</td>
<td>Sad</td>
<td>.001</td>
</tr>
</tbody>
</table>
Figure 17. Grand average difference waveforms for each emotional deviance, collapsed across electrodes, presented for each participant group. Response potentials are measured in microvolts (uV) and time in milliseconds (ms).
**Discussion**

This study demonstrated that the natural speech monosyllables with four different emotions (with neutral emotion as the standard and happy, sad and angry as deviants) evoked measurable obligatory CAEP components and significant mismatch effects from adults and children with and without HF-ASD. The statistical differences between groups suggest that such stimuli, implemented in oddball paradigm, have the potential to be a sufficiently sensitive indicator of emotion processing in children with ASD, a factor that contributes to social communication disorder in ASD. A combination of computer-based emotion perception training and a RMHA trial for children with HF-ASD was associated with altered MMR latency and amplitude and hence appeared to change the way they perceived emotionally loaded auditory stimuli, as reflected by significant differences in mismatch response effects.

**Between group differences for CAEPs**

CAEPs were compared between TD adults and children. The recorded waveforms closely resemble those published in maturation studies documenting the development of auditory evoked responses (such as in Ponton et al., 2002). As documented, cortical responses from younger individuals tend to be dominated by P2-N2 components around 200-300 ms post stimulus presentation across frontal, central, and parietal electrode sites (Wunderlich et al., 2006). The subsequent emergence of P1 and N1 between 100-200 ms has been reported to be highly dependent on the acoustic properties of the auditory stimulus (Cone & Whitaker, 2013). The tracking of CAEP maturation has by and large stemmed from studies using pure tones and simple acoustic or linguistic manipulations of speech stimuli.

Studies have reported differential amplitudes and latencies of P1-N1 components that correspond to spectro-temporal changes during consonant-vowel transitions in complex...
speech stimuli (Čeponienė et al., 2008; Sandmann et al., 2007). The presentation of complex speech stimuli differing in ‘emotion’ in the current study introduced changes to neural responses even in our TD sample, compared to studies that have used simpler stimuli. While the waveforms (Figure 13) of the TD adults and children had the expected structure – with adult participants clearly exhibiting P1-N1-P2 components, and the children’s waveforms consisting of the expected large positive response, followed by a large negative trough – the timing of responses is later than typically reported. For adults, peak latencies are approximately 50-80 ms later than what is usually reported in the literature. Similar shifts in latency have been observed in a study involving pitch, emotional tone, and instrumental timbre perception (Goydke, Altenmüller, Möller & Münte, 2004), where P1-N1-P2 across frontal, central, and parietal cortical regions on average about 80 ms later in response to emotion and timbre stimuli, compared to acoustic pitch. In general, the children’s P1 response to emotional deviant stimuli was both earlier and larger than the adults’ P1. This is consistent with reports of P1 decreasing in amplitude with age (Anderer et al., 1996; Wunderlich et al., 2006). P1 latency did not differ between TD adults and children for the neutral standard stimulus, but did differ for the emotional deviant stimuli, which suggests that the observed delay in the children’s response for other stimuli could be emotion-driven.

CAEPs were compared between TD children and HF-ASD children before intervention. For the neutral-standard stimulus, diminished amplitudes for the children with HF-ASD were only evident for the N1 component. Children with HF-ASD had significantly later P1, N1 and N400 responses which could reflect a fundamental temporal deficit in sound processing. Differences were not consistently significant across electrodes, however, which could reflect the relatively small sample and variability in the data responses to the emotionally deviant stimuli. No robust group differences for the early P1-N1 components were shown. A group difference at one electrode site that showed a significantly earlier N1
response from TD children may be a false positive, or there may be differences at other electrode sites that were not evident due to the small sample size.

Lastly, CAEPs of children with HF-ASD were compared between recordings made before and after intervention. Changes in response magnitude and timing were only evident towards emotional stimuli, and not the neutral standard. This could reflect a computer-training-specific effect as the training only included emotional speech stimuli. Significant pre-post intervention differences in N1 amplitude and N400 amplitude and latency were evident across all electrode sites. Amplitudes of both components were larger, and N400 was later, in these children’s waveforms after the intervention.

It is relevant to point out that all participant groups displayed a large N400 response that is not usually reported for other types of auditory stimuli. It is interesting that this late negative component was the most sensitive to between group differences, as well as to the effects of the intervention. This is consistent with the proposal in the literature that the N400 is modulated by semantic and emotional processing (Kutas & Hillyard, 1984). Studies investigating the relationship between semantics and affect in written words (Holt et al., 2009; Kanske et al., 2011), as well as prosodic cues (Kotz & Paulmann, 2007; Wambacq & Jerger, 2004) have reported increases in N400 amplitude as a function of both factors. N400 has been associated with the cognitive appraisal of the semantic and emotional content of the auditory stimulus (Kutas & Hillyard, 1984). The finding of group differences supports the use of emotional speech stimuli to examine CAEPs in children with HF-ASD; in particular the N400 component may be useful in identifying emotion processing dysfunction, as well as evaluating the effects of emotion-based auditory training.

Results from this study lend support to the view that emotion is primarily modulated in speech by changes in pitch contours and temporal resolution (Banse & Scherer, 1996;
Murray & Arnott, 1993; Yildirim et al., 2004). That is not to say that other acoustic parameters like harmonic-to-noise ratio, and formants, would not reveal more in-depth differences. More sophisticated analyses of natural speech stimuli would better inform future studies on what parameters to control for or manipulate. It would also be useful in future studies to determine the sensitivity of the N400 and other CAEP peaks, based on degrees of emotional valence and arousal, to investigate if there are neurological correlations between acoustics and dimensional ratings of emotion (Belyk & Brown, 2014; Johnstone & Scherer, 2000).

**Between group differences in emotion change detection, exhibited by MMRs**

Auditory change detection studies reported in the literature most commonly involve non-speech sounds and simple speech stimuli. These studies report that a prominent MMN is usually detectable in children at around 250 ms post-stimulus presentation (Csépe, 1995; Morr et al., 2002; Shafer et al., 2002). This response decreases with age (Gomot et al., 2000), plateauing at the average adult latency of 150ms (Näätänen, 1992). Consistent with other studies, the TD children consistently exhibited larger magnitude MMRs compared to adults (Gomot et al., 2000; Kraus et al., 1992; Shafer et al., 2000).

The MMRs recorded in the current study included negativities observed at the expected latencies for TD adults (labelled MMR1) and children (labelled MMR2) (Table 16). However, presumably due to the nature of the stimuli, additional MMRs were also evident after subtracting the CAEPs to neutral-standard monosyllables from those to emotional deviant monosyllables. The presence of positive MMRs in adults is interesting as these have only been previously documented in children as a result of manipulating speech-related factors (Cheng et al., 2013; Kraus et al., 1992; Shafer et al., 2010). This may reflect the complex nature of emotion perception, and lack of previous reports may reflect the sparse
literature around CAEPs to emotional speech stimuli. Future work in this area would benefit from a cross-sectional or longitudinal investigation of maturational effects on CAEPs evoked by a range of stimuli with different emotional content, valence and arousal strength.

The two TD sample groups, each displayed four distinct regions of mismatch with varying magnitudes and latencies, which may reflect complex processing of the natural emotional speech stimuli. In contrast, the children with HF-ASD (before intervention) displayed a simpler mismatch pattern that consisted of two MMR regions. Their first negative MMR is consistent with commonly reported MMN in children evoked primarily by non-speech stimuli (Csépe, 1995; Morr et al., 2002; Shafer et al., 2002). This suggests that the children with HF-ASD may have processed the stimuli as simple sounds, as the MMR did not reflect the spectral and temporal complexity of the speech stimuli. Evidence of a later negative MMR indicates that, like their TD peers, the HF-ASD children did detect differences between the emotional deviants and the neutral standard. However, the significantly diminished amplitude and increased latency suggests that the children with HF-ASD were less effective and efficient in discriminating the emotions (Fan & Cheng, 2014; Kujala et al., 2005).

Post intervention, the HF-ASD children had a negative MMR at around 200 ms, which was similar to the MMR recorded in this latency region from TD children and adults. The significant reduction in latency for this MMN, as well as the late MMR post-intervention, could be attributed either to the auditory training on the computer, or to the clearer speech signals received through the RMHAs, or a combination. Koerner, Zhang, Nelson, Wang and Zou (2016) found a decrease of amplitude and increase in MMN latency to consonant and vowel changes in speech stimuli with an increase in background noise. The opposite effect was observed post-intervention in the results of this study, suggesting that
children with HF-ASD became better able to discriminate between the spectro-temporal features and emotional content of the complex speech sounds post-intervention. The effects of the intervention on CAEP and MMRs in the HF-ASD group should be interpreted with caution however as a control HF-ASD group was not included to check for test-retest effects.

The significant emergence of a large positive MMR could also be attributed to the effects of emotion perception training resulting in increased sensitivity towards pitch. Studies have reported significantly enhanced P2 (which could be contributing to positive polarity MMRs) after auditory discrimination training targeting a wide range of factors, from voice onset times (Tremblay et al., 2001; Tremblay & Kraus, 2002), to phoneme changes (Kraus et al., 1995), and pitch (Tong et al., 2009). It has been proposed that the auditory evoked P2 may be considered as an indicative biomarker of learning and plasticity (Tremblay, Ross, Inoue, McClannahan & Collet, 2014), however this notion has also been refuted by reports of similar P2 enhancements with or without discrimination training (Sheehan, McArthur & Bishop, 2005). A study design including attended-to stimuli might better elucidate effects of training on P2.

**Distinguishing emotions: between participant groups, and the effects of intervention**

TD adults and children showed larger mismatch amplitudes in response to sad as the emotional deviant compared to happy and angry. The children revealed a significantly earlier first MMR in response to sad deviant stimuli. Consistent between adult and child TD groups was the lack of difference in MMR size and timing for angry and happy emotional deviants. These findings suggest that emotion discrimination may primarily be pitch-driven, as anger and happiness are characterised by increased mean pitch, pitch range, and vocal intensity, whereas these acoustic parameters are usually reduced for sadness, coupled with a slower rate...
of speech and longer inter-articulation silences (Banse & Scherer, 1996; Murray & Arnott, 1993; Yildirim et al., 2004).

In contrast, children with HF-ASD did not show any differences between emotions in their MMRs pre-intervention. The grand average waveforms suggest that sad is separable from angry and happy, but due to large variations between individuals and a small sample size, this effect is not significant. This is consistent with results from behavioural (Peppé et al., 2007) and neurophysiological (Fan & Cheng, 2014; Kujala et al., 2005) studies that show impaired emotion discrimination in children with ASD. After receiving the emotion training intervention, the MMRs of the children with HF-ASD look substantially different. Focusing on the two early and late MMRs, there is evidence for the amplitude of responses to angry being clearly distinguishable from happy, and happy differing significantly from sad. A possible reason behind MMRs towards angry being especially enhanced in the children with HF-ASD could be the dominance of reactions towards threatening stimuli, driven by heightened anxiety, evident in the visual attention and facial discrimination literature (Cisler, Bacon & Williams, 2009).

It is also interesting to note that variations in responses to emotional deviance occur only during the early MMRs. Although both adults and children exhibited negative mismatches at around 400 ms, as well as late discriminative negativities at around 600 ms, the responses to the three emotions did not differ. This suggests that discrimination between emotions for neurotypical individuals is more pre-attentive, and does not engage later-occurring higher cognitive processes. This, however, was not the case for those with HF-ASD. Significant differences between emotions for both the early and late negative MMRs suggests both pre-attentive and conscious appraisal of emotional differences. This suggests that the emotion perception training, did not ‘fix’ the atypical processes of the children with HF-ASD, and instead they were still using atypical processes to differentiate the emotional
stimuli post-intervention. Thus, training may enable people with ASD to utilise compensatory methods that may aid social communication but not ‘normalise’ underlying auditory processing. Future research that includes other measures of auditory processing and assesses social communication would help to elucidate this.

**Limitations and future directions**

One limitation of this study was that both TD adults and children only had one set of CAEP recordings. In the effort to attract as many volunteers as possible, the researcher observed that people were more receptive towards a one-off participation instead of committing to multiple visits. However, it would have been more relevant to look at CAEP data from two time points, between TD participants who received no emotion training and RMHA intervention, and participants with HF-ASD who did. This would help differentiate the true effects of the intervention, from natural changes in CAEPs as a result of listening to the auditory stimuli twice.

Another concern is that assigning electrode reference to the mastoid sites would elicit inflated responses over fronto-cortical sites because of interference. Whilst it is typical to have a nose-reference for mismatch studies, in this case, the researcher foresaw the risk of the children with HF-ASD not being able to tolerate an electrode on their nose. Their high anxiety levels, coupled with tactile hyper-sensitivity which is common in ASD, would have meant probable non-compliance and high participant drop out of the study. A review of the methodology of nine studies (ranging from years 1995 to 2011) who have conducted ERP recordings looking at MMN with children with ASD, revealed that only 2 used nose references. The other studies used the right mastoid (2), linked ear lobes (2), Cz, (1), and FCz (2), and it is suspected that these authors may have also encountered pragmatic difficulties with using nose references. A potential solution to electrode interference would be to re-
reference, in an offline analysis, to FCz, which is the furthest point away from the auditory cortex available in this study.

Another limitation is that there were substantially more Neutral stimuli than emotive stimuli in the overall number of trials, and may have influenced component morphology. Future extensions of this work would benefit from a more simplified block design, with the comparison of fewer emotions within a single sequence (e.g. Neutral vs. Angry, Happy vs. Sad), and the addition of more blocks of trials to cover all the combinations. This would ensure a more even distribution of standard and deviant versions of each emotion.

This study used brief consonant-vowel natural speech emotional stimuli and a neurophysiological method to examine emotional speech processing in children with HF-ASD. The paradigm and stimuli were sensitive to effects of maturation, emotion, ASD and training. Emotion perception dysfunction may underlie social communication difficulties in children with HF-ASD. Determining the relative contribution of auditory discrimination versus emotional perceptual difficulties in the findings for children with HF-ASD requires future work to examine effects of changing stimulus acoustic properties and emotional dimensions. These stimulus manipulations could also help to elucidate the complex MMR patterns recorded in the TD adults and children, which would facilitate understanding of the atypical MMRs of the children with HF-ASD.
Chapter 8. General Discussion and Conclusion
Overview of research

This research was motivated by evidence that individuals with ASD experience significant auditory processing difficulties (Alcántara et al., 2004; Haesen et al., 2011; O’Connor, 2012), and it was hypothesised that they would also have behavioural and neurophysiological evidence of impaired affective prosody. Published evidence shows that individuals with ASD demonstrate considerable difficulties with semantic, interpretive, and pragmatic aspects of language (Frith, 1989; Frith & Happé, 1994; Minshew et al., 1995; Tager-Flusberg, 1996). They display impairments in discriminating prosodic cues such as stress, pitch, and affect, which hinders their understanding of what others are trying to convey (Peppé et al., 2006). It is thought that deficits in prosody perception also influence prosody production, resulting in atypical use of pitch (Green & Tobin, 2009; Sharda et al., 2010), stress patterns (McAlpine et al., 2014; Paul et al., 2005; Shriberg et al., 2011), and articulation (Shriberg et al., 2011) in the speech of individuals with ASD.

A series of studies were undertaken to examine the assessment and training of affective prosodic perception and production in children with HF-ASD using a range of methodologies. The first study (Chapter 4) examined the effects of implementing a computer-based emotion training program, in conjunction with a trial of a RMHA system. Several published studies report positive auditory processing and academic outcomes for RMHAs in children with ASD (Rance et al., 2014; Schafer et al., 2013), but none has examined affective prosody in children with ASD using RMHAs. Participants were assessed on prosody perception and production, twice before and twice after a 3-week intervention period. Facial expression and affective prosody perception were measured using a standardised set of social perception subtests. Prosody production was evaluated by acoustically analysing, and having raters subjectively judge, speech samples recorded from the participants for angry, happy,
and sad emotional tones of voice. One of the groups of children with HF-ASD were provided with RMHAs to use at school during classroom teaching time for the duration of the intervention period. Students’ and teachers’ reports of listening experiences in class were obtained before and after the three weeks, in addition to the behavioural and neurophysiological measurements.

The auditory detection of changing affective prosodic cues was measured objectively via the neurophysiological recording of CAEPs. Natural speech stimuli spoken with angry, happy, sad, and neutral emotional expressions (based on the perceptions of a group of normal hearing adult listeners) were sampled from existing recordings of a male speaker (Ross et al., 1991). These stimuli have not been utilised in any previous research. One exemplar for each emotion was used to elicit pre-attentive change detection responses using a passive listening mismatch paradigm. CAEP data were collected from TD adult to establish that the stimuli evoked robust mismatch responses that differed across emotion, and was reported in a feasibility study detailed in Chapter 6. In the third study (Chapter 7), which examined the effects of prosody training and RMHA use on prosody perception and production, CAEP responses and MMRs were recorded from TD control children and children with HF-ASD, and compared. Children in the TD group were tested CAEPs and MMRs were recorded from the children with HF-ASD children before and after their participation in the intervention. The effect of the intervention on neurophysiological responses was examined for the children with HF-ASD.

Summary of key findings

Children with HF-ASD were able to engage in computer training activities using images and speech samples depicting angry, happy, and sad emotions. The training approach was based on the concepts underlying video modelling training (Cardon, 2013; Charlop &
Milstein, 1989; Smith et al., 2014), and auditory discrimination training (Kraus et al., 1995a; Tong et al., 2009; Tremblay et al., 2001), with repeated exposure to the task. It is likely that the children’s engagement with the task was facilitated by the characteristic ASD symptom of preference towards repetitive and restricted behaviours (APA, 2013). This approach has been shown to be effective in teaching both children and adults with ASD to rapidly learn new target skills such as appropriate conversational speech (Charlop & Milstein, 1989) and communicative gestures (Cardon, 2013).

Does the implementation of a computer-based emotion training program have an effect on affective prosody perception in children with HF-ASD?
Implementing computer-based emotion training activities in a structured 3-week intervention period was associated with significant improvements in social perception. Pre-intervention, children with HF-ASD scored significantly more poorly than TD peers on both ACS Social Perception subtests – facial expression identification, and the matching of facial expressions to affective speech. Post-intervention, the children with ASD demonstrated a significant increase in performance scores on the subtests, to the point where they superseded those of the TD control who did not receive any training. This improvement was unlikely due to simple task practice effects, because repeating the task within the 2-week baseline period during the pre-intervention assessments did not result in significant change in scores. Performance on the social perception subtests was maintained at a high level when the children were measured again after a 2-week follow up period.

Does the implementation of a computer-based emotion training program with RMHAs have an effect on affective prosody perception in children with HF-ASD? Do the effects differ from those seen as a result of computer training only?
A second group of children with HF-ASD received the computer-based emotion training, and in addition they wore RMHAs during the training and at school. There were no significant differences in social perception between the computer training only and the computer training with RMHA groups. However, results for prosody perception suggest that the simultaneous use of RMHAs with computer training was associated with better retention of newly acquired emotion perception skills, compared to computer training alone. This was demonstrated via the maintenance of stable performance during the follow up period post-intervention for the RMHA group but not for the computer training only group.

Comparisons between acoustic measures and raters’ judgements revealed a number of common parameters that were sensitive to emotion differences in the children’s speech, which warrant further investigation as they could potentially facilitate objective identification of angry, happy, and sad speech. These acoustic factors were mean fundamental frequency (pitch), various measures of pitch range (measured in Hertz and semitones), mean intensity (loudness measured in decibels), and various measures of intensity range (measured as standard deviations from the mean, or as absolute range between minimum and maximum intensity). Other acoustic measures that were investigated were not correlated with emotion and hence do not appear to be critical for the perception of emotion in speech.

*Does wearing a RMHA system at school have an effect on students’ reports of listening experiences, and teachers’ reports of listening behaviours in the classroom?*

Children and teachers responded well to the training in the use of the RMHAs and all 11 participants who agreed to the school trial (out of 12 in the RMHA group) completed the trial. After the 3-week experience with a personal amplification device, children with HF-ASD in the RMHA group reported improving listening experiences in the classroom (especially listening in noise). The teachers reported improved listening behaviours, and gave
positive ratings for the effectiveness of the intervention. Because the use of this technology has a clear impact on the speech signal, it is not possible to conduct blinded trials with RMHAs, and hence an approach that has been used in the literature is to compare RMHAs to other treatment options (e.g. M. Sharma et al., 2012). RMHAs are currently the preferred evidence-based amplification treatment option for children with APD (Esplin & Wright, 2014; Reynolds et al., 2016), with positive effects on language and learning, objective and subjective measures of auditory processing, academic performance and psychosocial measures (Keith & Purdy, 2014; Schafer et al., 2014). Other studies have reported similar benefits for school children with hearing loss (Socklingam et al., 2007), and reading delay (Hornickel et al., 2012; Purdy et al., 2009), as well as improved speech-in-noise performance for adults recovering from acquired auditory processing deficits after a stroke (Koohi et al., 2016). The encouraging results from the current study support existing studies of RMHA use in children with ASD (Rance et al., 2014; Schafer et al., 2013), and support the view that RMHAs are a viable and effective treatment option for children with HF-ASD.

What are the characteristics of CAEPs recorded in response to emotional speech stimuli? In what ways are these similar or different to what is reported in the literature?

In the study of TD adults, CAEP components P1, N1, P2 and N2 evoked as obligatory responses to the emotional speech monosyllables were consistent in morphology and latency with CAEPs evoked by other non-speech auditory stimuli (Näätänen, 1992). The CAEP components recorded in the children differed, as expected, from the adult waveforms (Ponton et al., 2002; Wunderlich et al., 2006). The children’s responses consisted predominantly of a large P1-N250 complex. All participant groups –TD adults, children, and those with HF-ASD – exhibited a late N400 component, which has been associated with the cognitive appraisal of semantic and emotional information in stimuli (Holt et al., 2009; Kanske et al., 2011; Kotz & Paulmann, 2007; Kutas & Hillyard, 1984; Wambacq & Jerger, 2004). This suggests that the
auditory stimuli were successful in engaging neural processes associated with emotion processing.

**Did MMR morphology differ between emotions in adult participants?**

MMRs to the natural emotional monosyllables were very different to the commonly reported MMN observed in studies involving non-speech sounds such as tones and clicks, or simpler manipulations of speech stimuli, such as loudness, duration, and overall pitch (Alho et al., 1994; Kujala et al., 2001; Näätänen et al., 1989; Näätänen et al., 1992; Näätänen & Alho, 1995; Novak et al., 1990; Paavilainen et al., 1991; Tiitinen et al., 1994; Winkler et al., 1995).

The adult data showed significantly different MMRs between emotions that included both positive and negative mismatch peak regions. This was evident MMR waveforms were derived using the standard method of subtracting emotions-as-deviant from neutral-as-standard, as well as using the flipped method of subtracting neutral-as-deviant from emotions-as-standard.

**Were there differences in CAEP and MMR morphology between TD children and children with HF-ASD (measured before intervention)?**

MMR patterns from both TD and HF-ASD groups of children showed more than one latency region of MMN. MMR was present in several post-stimulus latency regions and differed in magnitude across emotions. The natural speech stimuli used to evoke MMR recordings were selected because they were consistently judged by a group of 14 adult listeners as sounding angry, happy, sad and neutral, respectively. They differed acoustically and in their emotional tone, but had matched overall duration and root mean square intensity. These stimuli had not been used in previous research. The morphology of the observed MMRs in this study differs from phonetic change detection results for non-emotional speech sounds in children (e.g. Kraus et al., 1993), suggesting that the results obtained do not simply reflect acoustic
differences between stimuli and instead reflect, at least in part, differences in emotional content.

*Was the combined intervention of computer-based emotion training and RMHA use in children with HF-ASD associated with a change in CAEPs and MMRs?*

MMR patterns elicited from children with HF-ASD changed significantly after they participated in the 3-week intervention of computer-based emotion training combined with RMHA use during training and at school. Pre-intervention, children with HF-ASD displayed two large MMR negativities that occurred significantly later than those recorded for the TD group. Post-intervention, the children with HF-ASD showed early and late negativities at latencies that were similar to TD controls. In addition, they showed an enhanced positivity at around 300 ms, which is consistent with increases in deviant response amplitudes that have been previously associated with post auditory discrimination effects (Kraus et al., 1995a; Tremblay et al., 2001; Tremblay & Kraus, 2002).

Although the participants were instructed not to pay attention to the auditory stimuli during the experiment, it is not possible to gauge attention to the stimuli in a passive listening paradigm during CAEP recording. TD adults and children, and HF-ASD children pre-intervention, did not show an increased positivity in the standard-deviant difference waveform latency region around 300 ms. Thus, there was no evidence for elicitation of a P300 component that is commonly associated with active attention towards a target or novel stimulus (Picton, 1992). It is possible that after receiving the intervention the children with HF-ASD were more aware of the deviations in the sequence of sounds, and the changes in MMR, including the enhanced positivity at about 300 ms, may have reflected greater attentional focus on the deviant sounds.
Post-intervention, children with HF-ASD also displayed significantly greater divergence in the responses to the three emotional deviants in their mismatch (standard-deviant) difference waveforms. This separation of MMRs across emotions was not present in the MMRs in the HF-ASD group recorded before the intervention. These results suggest that computer-based emotion training with the use of RMHAs contributed to the increased perception of differences between angry, happy, and sad emotions, consistent with the behavioural data, which was manifested as distinctly different waveforms that varied in response magnitude and timing across emotions.

TD adults and children only showed differences in MMR between emotions in their early negative MMRs, whereas HF-ASD children (post-intervention) showed a difference in MMR for both their early and late latency regions. One possible interpretation of this is, although the intervention resulted in an increase in the perception of emotion differences in the children with ASD, they were still using atypical processes for auditory change detection. For TD adults and children, emotion discrimination may be more automatic and pre-attentive, as MMR differences between emotions were evident early in in the process based on the latency of the MMR. Post-intervention, it is possible that the children with HF-ASD were engaging in a more conscious and effortful process to differentiate between emotions, hence mismatch effects were also present later in the evoked potential waveform in latency regions that are more synonymous with semantic and emotional cognitive processing.

**Novel aspects of this research**

Published studies have documented the positive effects of training individuals with ASD on cross-modal (visual-auditory) emotion perception (Baron-Cohen et al., 2002; Matsuda & Yamamoto, 2013), prosocial behaviour (Butler, 2011; Tse et al., 2007), and attention to semantic and prosodic content in speech (Ploog et al., 2009), all with some form
of computer- or technology-based methods. There is also evidence of the benefits using RMHAs with children with ASD for their auditory processing and academic performance (Rance et al., 2014; Schafer et al., 2013). However, no study to date has combined both aspects to investigate the joint effects of computer-training and auditory amplification using RMHAs on prosody perception and production. A novel computer-training intervention using audio-visual stimuli was developed by the researcher. A 3-week intervention period using this computer training, trialled with and without RMHAs, was associated with significant improvements in prosody perception but not prosody production, although some positive trends were noted.

Many studies have used either behavioural or neurophysiological measures to investigate prosody processing in people with ASD but few studies have examined both in the same participants. This research used monosyllables with emotional tones sampled from natural speech, as opposed to synthetic sources, as auditory stimuli for the evoked potential recordings. These have never been implemented in a neurophysiological study of affective prosody discrimination. The use of these speech stimuli may have contributed to the finding of group differences in CAEPs and MMR between controls and children with HF-ASD. Results in the literature are mixed regarding CAEP and MMR patterns obtained using a passive listening mismatch paradigm in children with ASD. Differences in stimuli are likely to contribute to this variation in results across studies. The speech stimuli used for ERP recordings had good ecological validity as they were natural speech tokens that were consistently judged by a large group of listeners to reflect angry, happy, sad, or neutral emotion.
Limitations of the research

Although the computer-based emotion training activities were effective in improving social perception, it is unclear whether the children with HF-ASD had increased their facial and prosodic perception skills, or whether they have mastered the matching-to-sample task. Similar training conducted by Matsuda and Yamamoto (2013) reported that repeated exposure to this task structure resulted in generalised improvements in face-voice pairings. Although performance was stable for the two baseline assessments the subsequent improvement could reflect mastery of the task with further practice. However, the finding of significant change in neurophysiological responses to affective speech stimuli after training suggests that there were changes in prosody perception in the children.

A limitation of this study was the absence of a HF-ASD group that only trialled the RMHAs without receiving computer-based emotion training. Future research in this area would benefit from including such a group, so that observed effects could be more clearly attributed to either the computer training, or the RMHAs, or indeed a combination of both interventions.

Another limitation of the RMHA trial was that the researcher did not monitor the exact number of hours the children wore the systems at school. General guidelines pertaining to the study were given to both the participants and their teachers stating that the children were to try and wear the RMHAs during classrooms hours throughout the day, and to assist them in removing the receivers during snack/lunch break and physical education class to ensure that the equipment was not damaged. However, anecdotal reports from individuals revealed that some teachers took the liberty in removing the receivers during other classes like drama, as well as school play rehearsals, and some reported that they removed the
receivers if the student was feeling uncomfortable wearing them and if it was interfering with their school work.

Neither approach to analysing prosody production – acoustic measures and raters’ judgements - revealed a significant effect of intervention on prosody production for the HF-ASD groups. One improvement in the acoustic analysis of the affective speech samples could include tracking changes in pitch and intensity contours, instead of taking a cross sectional average of the measures across the speech sample. This more fine-grained acoustic analysis approach has been found to provide more indicative information about fluctuations and atypical use of prosodic cues in speech (Falk et al., 2012; Murray & Arnott, 1993; Waaramaa & Kankare, 2013; Yildirim et al., 2004).

Improvements in the subjective approach to evaluating prosody production could be to engage naïve listeners to rate the samples, as their ratings might be more sensitive to inappropriate use of prosodic cues to convey emotion. There has been relatively little research in the area of subjective judgements of the accuracy or naturalness of affective speech for different conditions that affect prosody production (Chin, Bergeson & Phan, 2012; Peng, Tomblin & Turner, 2008; Samuelsson, 2011), including hearing loss and ASD. The focus of the training in the current study was on perception rather than production – future research could explore training paradigms that targeted both areas.

Both the behavioural data and CAEP data collected from children with HF-ASD were limited to those who have relatively less severe ASD symptoms and who were able to cooperate for testing. Therefore, the results cannot be presumed to be representative of the whole spectrum of children with ASD. This is a general problem for research in the area of ASD when the experimental paradigm requires active engagement of the child in a task, or co-operation for electrode application and ERP recordings.
In addition, although the researcher determined (via an introductory questionnaire) that none of the HF-ASD participants were currently engaged in any professional therapy services, families were not asked about past history of therapeutic experience (e.g. speech-language, music, play, or behavioural therapies), which would have had a significant influence on their abilities during baseline assessment, and their personal approach towards the intervention. Anecdotally, the researcher observed varying individual personalities of each child with HF-ASD. Some were more socially distant, while others were more anxious and emotional, and sought parental comfort. The children also varied in their sensitivities to noise, and differences in coping strategies may reflect in part history of therapy, and what strategies the child’s family uses to moderate behaviour.

Future revisions or extensions of this study would benefit from additional information about individual personality traits and history of treatment (age of onset, as well as years of engagement), to determine the impact of this on therapy outcomes. If sample size was sufficient this could allow for the separation of HF-ASD participants into different groups or the inclusion of various covariates into the analysis to better establish the effectiveness of the intervention on behavioural and ERP measures of prosody processing in children with different characteristics. Outcomes could also be compared to the parent ratings of autism features (CARS-2) and communication (CCC-2) questionnaires.

**Future directions**

A future direction for this research would be to further explore visual and auditory emotion perception training via an interactive computer game. The delivery of learning material for children with ASD and other special needs via a computer-based platform has received widespread attention (Ploog et al., 2013), and particular success has been documented with video games (Ploog et al., 2009). It would be useful to know if children
with ASD would be able to utilise their social perception knowledge acquired solely from playing the computer game in everyday life scenarios with real people. It is possible that they require task management, structure, and target goals in order for their newly acquired abilities to be generalised to real-life situations outside of their home or therapy session. In this case, computer-based interventions may be more suited as an accompaniment, instead of a replacement, to the one-on-one real-person interactions that individuals with HF-ASD receive in traditional therapy.

Future studies involving RMHA trials at school would benefit from asking teachers to keep track of the number of hours the child wore the system during the school day. In addition, the trial could include the wearing of RMHAs at home or during routine activities, like going to the supermarket, with monitoring by the parents of the number of hours the child wore the system. The amount of time spent using the RMHAs and being exposed to an enhanced auditory environment may be predictive of improvements in prosodic perception and social communication.

The acoustics-subjective correlation results from the analyses of prosody production may be useful in providing a foundation for future developments examining the acoustic features of typical and atypical affective speech. After pinpointing the key acoustic features that influence subjective emotion identification, a potential next step would be to manipulate these features systematically to vary the intensities of each emotion, for example having stimuli that ranges from ‘hot’ to ‘cold’ anger (Banse & Scherer, 1996) to determine the effect on listeners and social interactions. It would be interesting to conduct follow up evoked potential experiments using the mismatch paradigm to investigate whether TD individuals differ from those with ASD or other neurological conditions in the processing of emotional intensity, rather than category of emotion as was investigated here. It would also be
interesting to see if social perception and communication training would affect the neural representations of subtler emotional contrasts.

There is substantial overlap between children with APD and ASD in terms of auditory processing difficulties in noise (Alcántara et al., 2004, Foxe et al., 2013), difficulties with competing sensory input (Brooks & Ploog, 2013; Carpenter et al., 2014), and atypical sensory processing in general (Tomcheck & Dunn, 2007). Future studies could explore the ‘fine-tuning’ of auditory speech stimuli so that they reflect emotion perception versus simple acoustic contrasts in order to better differentiate auditory and affective prosody perception – this difference could potentially differentiate ASD from APD. Such research might show that deficits in emotion processing are dominant in the ASD population compared to children with APD who may show more generalised difficulties with auditory stimuli. In the future, we are likely to know much more about different genetic bases for ASD, which may allow the possibility of identifying differences in auditory and affective prosody processing across groups of children with ASD that show varying aetiologies and types and degrees of symptoms. It is the hope that this series of studies will lead to further research investigating atypical auditory processing via behavioural and neurophysiological methods in children with ASD, as well as the development of interventions targeting social perception and auditory processing for ASD.

The non-invasiveness of CAEP recording techniques, as well as the early manifestation of change-detection mismatch effects, makes this an ideal approach to identifying early neural markers of atypical auditory processing in children with ASD. A neurophysiological approach, involving CAEP recordings, strengthens the current predominant neurodevelopmental framework under which ASD is defined. This approach also has the potential to be considered alongside traditional clinical observations and
behavioural assessment to be part of the diagnostic and early screening procedures for the identification of ASD in young children.
Appendix 1: Ethics approval

UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS ETHICS COMMITTEE

05-Jun-2013

MEMORANDUM TO:

Prof Suzanne Purdy
Psychology

Re: Application for Ethics Approval (Our Ref. 9657)

The Committee considered your application for ethics approval for your project entitled Training affective prosody in children with autism spectrum disorder.

Ethics approval was given for a period of three years with the following comment(s):

1. The CF states that all results will be confidential, but the Committee considers confidential to mean that the results remain between researcher and participant, and therefore cannot be reported. Instead, participants can be assured that the questionnaires will only be accessed by the researcher, and in any publication, the results will be reported in aggregate and participants will be unable to be identified.

The expiry date for this approval is 05-Jun-2016.

If the project changes significantly you are required to resubmit a new application to UAHPEC for further consideration.

In order that an up-to-date record can be maintained, you are requested to notify UAHPEC once your project is completed.

The Chair and the members of UAHPEC would be happy to discuss general matters relating to ethics approvals if you wish to do so. Contact should be made through the UAHPEC Ethics Administrator at humanethics@auckland.ac.nz in the first instance.

All communication with the UAHPEC regarding this application should include this reference number: 9657.

(This is a computer generated letter. No signature required.)

Secretary
University of Auckland Human Participants Ethics Committee
c.c. Head of Department / School, Psychology

Additional information:

1. Should you need to make any changes to the project, write to the Committee giving full details including revised documentation.

2. Should you require an extension, write to the Committee before the expiry date giving full details along with revised documentation. An extension can be granted for up to three years, after which time you must make a new application.

3. At the end of three years, or if the project is completed before the expiry, you are requested to advise the Committee of its completion.

4. Do not forget to fill in the 'approval wording' on the Participant Information Sheets and Consent Forms, giving the dates of approval and the reference number, before you send them out to your participants.

5. Send a copy of this approval letter to the Manager - Funding Processes, Research Office if you have obtained funding other than from UniServices. For UniServices contract, send a copy of the approval letter to: Contract Manager, UniServices.

6. Please note that the Committee may from time to time conduct audits of approved projects to ensure that the research has been carried out according to the approval that was given.
MEMORANDUM TO:
Prof Suzanne Purdy
Psychology

Re: Request for change of Ethics Approval Ethics Approval (Our Ref. 9657): Amendments Approved

The Committee considered your request for change for your project entitled Training affective prosody in children with autism spectrum disorder and approval was granted for the following amendments on 02-Oct-2014.

The Committee approved the following amendments:
1) Inclusion of additional EEG session where the participants have a choice in participation.
2) Additional implementation of remote microphone hearing aids during the 3 week training period accompanied with computer tasks.
3) Provision of option for participants to use the remote microphone hearing aids at school.
4) To add Joan Leung, PhD student to the research study.

The expiry date for this approval is 05-Jun-2016.

If the project changes significantly you are required to resubmit a new application to the Committee for further consideration.

In order that an up-to-date record can be maintained, it would be appreciated if you could notify the Committee once your project is completed.

The Chair and the members of the Committee would be happy to discuss general matters relating to ethics approvals. If you wish to do so, please contact the UAHPEC Ethics Administrators at ethics@auckland.ac.nz in the first instance.

Please quote reference number: 9657 on all communication with the UAHPEC regarding this application.

(This is a computer generated letter. No signature required.)

UAHPEC Administrators
University of Auckland Human Participants Ethics Committee

C.c. Head of Department / School, Psychology
Prof Suzanne Purdy
Miss Joan Leung
Appendix 3: Recruitment advertisement (Study 1)

TRAINING AFFECTIVE PROSODY in CHILDREN with AUTISM SPECTRUM DISORDER (ASD) using a COMPUTER-BASED INTERVENTION

For our project, we are seeking children who have ASD, with no hearing difficulties, aged 7-13 years old. Studies have shown that processing emotional cues in facial expressions and in speech is an area of difficulty for people with ASD. We want to trial an iPad-based training application (app) to see if it will improve the perception and production of emotional cues in faces and speech.

The study requires a time commitment that spans 9 weeks. The researcher will be more than happy to conduct the sessions at a location that is most convenient for you.

There will be four assessment sessions, each taking about 1 hour. These will include:
- A standardised screening test, the “Childhood Autism Rating Scale” (CARS)
- A “Childhood Communication Checklist” questionnaire
- An assessment that requires listening to some segments of speech, and then matching them to an emotional facial expression
- Saying some sentences in different emotions, that will be recorded with a microphone
- And a story-telling activity, which will also be recorded

There will be two assessment sessions before, and two after, a 3-week training period. During this time, the child will get to participate in some training activities on an iPad. Three 30-minute sessions per week of the participants’ time will be required. The researcher brings the iPad to the session, and will go through specific activities with the child.

A $20 voucher will be gifted to you upon completion of your participation in the project.

All responses will remain completely confidential at all times and all data will be destroyed after 6 years by deleting the electronic files and shredding the paper files

By doing this research we hope to explore the feasibility of using an interactive and technological approach to better understand, and improve, communication difficulties in children with ASD.

If you are interested, or would like more information, please contact:
Joan Leung – jleu021@aucklanduni.ac.nz, or 021 842 349
Or Professor Suzanne Purdy – sc.purdy@auckland.ac.nz, ph: 3737599 ext. 82073

APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS ETHICS COMMITTEE ON June 5 2013 for 3 years from June 5 2013 to June 5 2016. Reference Number 9657
Appendix 4: Recruitment advertisement (Study 2)

DOES THE USE OF REMOTE-MICROPHONE HEARING AID TECHNOLOGY ENHANCE THE EFFECTIVENESS OF PROSODY PERCEPTION TRAINING IN CHILDREN WITH AUTISM SPECTRUM DISORDER (ASD)?

We are seeking children who have ASD, with no hearing difficulties, aged 7-13 years old.

Studies have shown that processing emotional cues in facial expressions and in speech is an area of difficulty for people with ASD. We want to pair the use of remote-microphone hearing aids (RM systems) with computer-based training tasks to see if it will improve the perception and production of emotional cues in faces and speech.

The study requires a time commitment that spans 10 weeks. The researcher will be more than happy to conduct the sessions at a location that is most convenient for you.

There will be four assessment sessions, each taking about 1 hour. These will include:
- A standardised screening test, the “Childhood Autism Rating Scale” (CARS)
- A “Childhood Communication Checklist” questionnaire
- An assessment that requires listening to some segments of speech, and then matching them to an emotional facial expression
- Saying some sentences in different emotions, that will be recorded with a microphone
- And listening to speech samples and identifying the underlying emotions

There will be two assessment sessions before, and two after, a 3-week training period. During this time, the child will first be given the opportunity to get used to wearing an RM system. They will then get to participate in some training activities on a laptop computer. Three 30-minute sessions per week of the participants’ time will be required. The child will also be given the option to try wearing the RM system at school.

In addition, there is an optional part of the study that involves recording the brain activity of the child in response to different emotional tones in speech. This part of the study is not mandatory.

A $20 voucher will be gifted to you upon completion of your participation in the project.

By doing this research we hope to gain a better understanding of communication difficulties in children with ASD; to assess the feasibility of using an interactive and technological approach to emotion training; and to explore whether social communication difficulties are based on auditory processing problems.

If you are interested, or would like more information, please contact:
Joan Leung – joan.leung@auckland.ac.nz, or 021 842 349
Or Professor Suzanne Purdy – sc.purdy@auckland.ac.nz, ph: (09) 373 7599 ext. 82073
Appendix 5: Recruitment advertisement (Controls)

TRAINING AFFECTIVE PROSODY in CHILDREN with AUTISM SPECTRUM DISORDER (ASD)

For our project, we are seeking children with no hearing difficulties, aged 7-13 years old.

Studies have shown that processing emotional cues in facial expressions and in speech is an area of difficulty for people with ASD. We want to compare the perception and production of emotional cues between children with ASD and typically developing children. By participating, you will be part of the comparison group.

The study will require you to travel to the Tamaki Campus of the University of Auckland for a 2 – 2.5 hours session. It will involve:

- A “Childhood Communication Checklist” questionnaire – to be filled in by parents/guardians
- An assessment that requires listening to some segments of speech, and then matching them to an emotional facial expression
- Saying some sentences in different emotions, that will be recorded with a microphone
- And a recording of brain activity in response to different emotional tones in speech

A $20 voucher will be given to you upon completion of your participation in the session.

All responses will remain completely confidential at all times and all data will be destroyed after 6 years by deleting the electronic files and shredding the paper files

By doing this research we hope to gain a better understanding of the extent of communication difficulties in children with ASD, and explore whether social communication difficulties are linked to auditory processing patterns at a neural level.

If you are interested, or would like more information, please contact:
Joan Leung – joan.leung@auckland.ac.nz, or 021 842 349
Or Professor Suzanne Purdy – sc.purdy@auckland.ac.nz, ph: 3737599 ext. 82073

APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS ETHICS COMMITTEE ON June 5 2013 for 3 years from June 5 2013 to June 5 2016. Reference Number 9657
PARTICIPANT INFORMATION FOR CHILDREN

Title: TRAINING AFFECTIVE PROSODY in CHILDREN with AUTISM SPECTRUM DISORDER (ASD)

Researchers: Joan Leung, and Professor Suzanne Purdy

Read out the following:

Suzanne, from the University of Auckland has been working with children for many years, especially those who have difficulties communicating. I (Joan), along with Suzanne, want to invite you to take part in our project! I’m going to explain a little about what is going to happen if you do decide to help us out.

I will be coming to visit you and your (parent/caregiver) over a period of 9 weeks. Starting with our first visit today, I am going to go through some picture matching, speaking and story-telling activities with you. For the speaking and story-telling activities, I will need to record your voice, and so I will put a microphone about 1-2 meters away from where you’re sitting. Today’s visit will take about 2 hours.

I will see you two weeks later to go through the same set of activities. By that time, you’ll be so good, it will only take about 1 hour. Your (parent/caregiver) will also be part of our team, and will be helping us out by answering some questions, while we do our activities. We can stop and take as many breaks as you like, if you get tired.

Then, I will come to see you more often – 3 times a week, each time staying for about 30 minutes. During these 3 weeks, you will get to play on my iPad. I have 3 different tasks on the iPad that we can go through together. I will explain how to do these tasks when we start, and you can ask me any questions to make sure you are fully ok with the activities before you do them.

After those 3 weeks, I will visit you two more times to run through the activities from the first two sessions again. We need to repeat these activities to get enough information for the project. But we’ll make sure we have lots of fun during each visit!

APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS ETHICS COMMITTEE ON June 5 2013 for 3 years from June 5 2013 to June 5 2016. Reference Number 9657
ASSENT FORM FOR CHILDREN

• I understand that Joan and Suzanne would like to go through some activities to do with matching facial expressions to speech.

• I understand that Joan and Suzanne would also like to record my speech with a microphone, when I am saying some sentences with different emotions, and when I am telling a story.

• I understand that I will need to repeat these activities 4 times in total.

• I understand that Joan and Suzanne would also like to go through some activities on an iPad during some short sessions, 3 times a week, for 3 weeks.

• I understand that Joan will be visiting, and completing activities with me, at different times during a period of 9 weeks.

• I know that I can tell Joan and Suzanne if I do not want to do the activities anymore, and that this will be fine.

• I understand that Joan will come to my house, or anywhere that is easy for me, to do these activities.

• I understand that Joan and Suzanne will not use my name when they talk about the results or when they write about the results.

• I agree to take part in this research project.

If you would like to do this, write or sign your name on the line below.
If you decide later that you don’t want to do this with us anymore, we can stop at any time.

My name: ______________________          My signature: ______________________

APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS ETHICS COMMITTEE ON June 5 2013 for 3 years from June 5 2013 to June 5 2016. Reference Number 9657
PARTICIPANT INFORMATION SHEET FOR PARENTS/CAREGIVERS

Title of Research:
TRAINING AFFECTIVE PROSODY in CHILDREN with AUTISM SPECTRUM DISORDER (ASD)

Researchers: Joan Leung, and Professor Suzanne Purdy

Thank you for taking an interest in this project!

Your child is invited to take part in this research being conducted by staff and students within the Department of Psychology (Speech Science). Before you decide, it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully and discuss it with others if you wish. Ask us if there is anything that is not clear or if you would like more information.

What is the purpose of the study?

Prosody refers to the variation in timing, pitch or stress pattern in speech that helps to convey meaning and emotion. Affective prosody, in particular, refers to both verbal (speech) and non-verbal (facial) emotional expression. Research has shown that individuals with ASD are impaired in the perception, as well as the production, of affective prosody. This project aims to trial an iPad-based intervention that trains children with ASD on emotional facial recognition, and perception of emotion in speech.

There is growing evidence that computerised tools, such as the iPad, offer an interactive platform for children with ASD to learn and express themselves. Many applications (apps) have already been developed to aid children with communicative difficulties to ‘speak’ through another medium. Apps also exist to help individuals with ASD, for example, organise a daily routine, and learn socially appropriate behaviour. So far, apps related to the recognition of social cues have mostly focused on facial expressions. The app from this project aims to introduce the additional aspect of having emotional cues present in speech. Our overall aim is to improve communication outcomes for children with ASD.
What is expected from you, and from your child?

1. Time commitment

Participation in this study would require an extended commitment that spans approximately 9 weeks. Before you continue reading the details below, we would like to inform you that the researcher would be more than happy to cater to your convenience, with regards to both your time and the location of where the sessions would take place.

First, there will be an introductory session that will last around 2 hours. During this time, you will receive all the information that you will need about the project. You and your child will also be asked to complete the first round of assessments, please see below for details. The 2\textsuperscript{nd} session would take place two weeks later. You will only be required to complete a round of assessments, exactly the same as the last ones, and so this would only take about 1 hour of your time.

Following this, there will be a 3-week period where your child gets to participate in some training activities on an iPad. We will require three 30-minute sessions per week of you and your child’s time. The researcher brings the iPad to the session, and will go through three activities with your child.

After this training period, you and your child will need to be assessed again for the 3\textsuperscript{rd} time. The same procedure (as the sessions before training) will once again require 1 hour of your time. And finally, the last session will take place two weeks after your 3\textsuperscript{rd} one. You and your child will be asked to complete a final round of the assessments. After this, you will be more than welcome to ask any further questions with regards to the project, and you will also be gifted with a $20 shopping voucher.

2. Assessments

\textbf{Questionnaires}

The study requires you, as a parent/caregiver, to complete two questionnaires (10-15 minutes per questionnaire). A routine screen for Autism symptoms will only need to be done once, during the first session. An assessment of your child’s communication abilities, as observed during the past week, will need to be completed four times, once for each assessment session. You can complete this while your child is being tested. Please be aware that the questionnaires are screening tools only. If the results are suggestive of previously undiagnosed communication problems or ASD, any necessary referrals will be made, and you can further discuss the matter with the research team. Please note that it is University of Auckland policy that we notify you regarding any significant findings from these questionnaires. You may choose not to participate if you do not wish to be informed of these results.

\textbf{Assessment of recognition and matching of facial expressions to speech}

Your child will be asked to participate in an assessment that requires them to listen to some segments of speech, and then match them to one facial expression out of a selection of six photos. We assure you that the photos are just portraits of people, of different genders, ages and ethnicities, making different facial expressions. There are no offensive or
frightening photos. The speech segments also contain no offensive language or statements, and just consist of people saying various generic sentences with different emotions.

**Assessment of producing emotion in speech**

Your child will also be asked to participate in saying some sentences of their own, in different emotions. They will be asked to say “it is eleven o’clock” in either an angry, happy or sad tone. They will be asked to repeat these a few times, so that we get enough data. To record your child’s speech, a microphone will be propped up on a stand and positioned 1-2 meters away from where your child is sitting. No other devices will be attached to your child body or clothes.

**Assessment of continuous speech and production of emotion**

Finally, your child will be asked to participate in a story-telling activity. Once again, their speech will be recorded for analysis later on. Your child will be presented with a picture book – “The Red Balloon” – and just be asked to tell a story through the pictures, with as much detail and reference to the character’s emotions, as possible.

3. **Training**

There will be three different tasks to complete during each 30-minute training session. Your child will be asked to complete two matching tasks, similar to the first assessment task. They will need to identify specific emotions out of a selection of different expressions, and match them either to the researcher’s prompts, or to various speech samples.

There will also be a facial expression producing task, where your child will be encouraged to touch their finger on the surface of the iPad and morph a neutral-looking face into different emotional expressions, depending on what the researcher asks for.

**Participation**

Participation in the study is voluntary so it is up to you and your child to decide whether or not to take part. If you do decide to take part you will be given this information sheet to keep and you will be asked to sign a consent form for your child. If you decide to take part, you and your child are still free to withdraw at any time and without giving a reason. You can also request for your data to be destroyed until 31st October 2013.

**Will our taking part in this study be kept confidential?**

All participants will be given a unique number when they join the study. This will be used by us, instead of their name, on all of their measurements. If the information you provide is reported or published, this will be done in a way that does not identify you as its source.

You can be assured that the questionnaires will only be accessed by the researcher, and in any publication, the results will be reported in aggregate and participants will be unable to be identified.

**What will happen to the results of the research study?**
The results of the study will be used to help us design future studies. It is also normal practice for the results of these studies to be published in the literature. If you are interested then we will send you a summary of the results (please tick the box to request a summary on your consent form)

Who should you contact if you wish to obtain further information?

You can contact Professor Suzanne Purdy or Joan Leung (details are below)

Prof. Suzanne Purdy
Head of Speech Science
The University of Auckland
Private Bag 92019
Tel: 373 7599 ext. 82073
E-mail: sc.purdy@auckland.ac.nz

Joan Leung
Masters Student (main researcher)
Mob: 021 842 349
E-mail: jleu021@aucklanduni.ac.nz

The Head of the Department of Psychology is:
Associate Professor Doug Elliffe
Department of Psychology
University of Auckland
Tel. (09) 3737599 ext. 85262
Email: d.elliffe@auckland.ac.nz

For any queries regarding ethical concerns you may contact the Chair, The University of Auckland Human Participants Ethics Committee, The University of Auckland, Office of the Vice Chancellor, Private Bag 92019, Auckland 1142. Telephone 09 373-7599 extn. 83711

Thank you for reading this information sheet.

 APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS ETHICS COMMITTEE ON June 5 2013 for 3 years from June 5 2013 to June 5 2016. Reference Number 9657
CONSENT FORM FOR PARENTS/CAREGIVERS

Title of Research: TRAINING AFFECTIVE PROSODY in CHILDREN with AUTISM SPECTRUM DISORDER (ASD)

Researchers: Joan Leung, and Professor Suzanne Purdy

This consent form will be retained for six years in secure storage by the researchers before being destroyed

Please initial each box

1. I confirm that I have read and understood the information sheet for the above study and have had the opportunity to ask questions.

2. I understand that the experiment will take 9.5-10 hours over a period of 9 weeks.

3. I understand that I will be asked to participate by completing two questionnaires.

4. I understand that my participation and my child's participation are voluntary and that we are free to withdraw at any time, without giving any reason.
5. I understand that the study will include a screening questionnaire that looks for features of ASD. I understand that if any difficulties are identified in my child I will have the opportunity to discuss this with the research team.

6. I understand that I have the right to withdraw the data until 31/10/2013.

7. I agree for my child to take part in this study.

8. I understand that the questionnaires will only be accessed by the researcher, and in any publication, the results will be reported in aggregate and participants will be unable to be identified.

9. I would like a summary of the results

Please circle: YES / NO

PRINT NAME (Parent)

PRINT NAME (Child)

PRINT ADDRESS

TELEPHONE: EMAIL:

We need this information so we can arrange a convenient appointment time, as well as send you a summary of the results if you want.

________________________________________  __________________________
SIGNATURE DATE

APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS ETHICS COMMITTEE ON June 5 2013 for 3 years from June 5 2013 to June 5 2016. Reference Number 9657
PARTICIPANT INFORMATION FOR CHILDREN

Title: DOES THE USE OF REMOTE-MICROPHONE HEARING AID TECHNOLOGY ENHANCE THE EFFECTIVENESS OF PROSODY PERCEPTION TRAINING IN CHILDREN WITH AUTISM SPECTRUM DISORDER (ASD)?

Researchers: Joan Leung, and Professor Suzanne Purdy

Read out the following:

Suzanne, from the University of Auckland has been working with children for many years, especially those who have difficulties communicating. I (Joan), along with Suzanne, want to invite you to take part in our project! I’m going to explain a little about what is going to happen if you do decide to help us out.

I will be coming to visit you and your (parent/caregiver) over a period of 10 weeks. Starting with our first visit today, I am going to go through some picture matching, speaking, and listening activities with you. For the speaking activity, I will need to record your voice, and so I will put a microphone on your head and place it 5cm from your mouth. Today’s visit will take about 1 hour.

I will see you two weeks later to go through the same set of activities. By that time, you’ll be so good at them! Your (parent/caregiver) will also be part of our team, and will be helping us out by answering some questions, while we do our activities. We can stop and take as many breaks as you like, if you get tired.

Then, I will come to see you more often. I will bring along a little hearing device that I will teach you how to use. We will have one whole week to get used to this, and you can tell me anytime you feel uncomfortable while wearing it. After that, I will come to see you 3 times a week, each time staying for about 30 minutes. During these 3 weeks, you will get to play on my computer. I have 4 different tasks on the computer that we can go through together while you are wearing the hearing device. I will explain how to do these tasks when we start, and you can ask me any questions to make sure you are fully ok with them. You can also keep the hearing device and wear it at school if you want.

After those 3 weeks, I will visit you two more times to run through the activities from the first two sessions again. We need to repeat these activities to get enough information for the project. But we’ll make sure we have lots of fun during each visit!
Aside from that, I would also like to ask you if you would like to get your brain activity recorded. How we do this is: I will stick 8 small discs onto your head with special paste. It does not hurt at all. Then I will connect the discs to the computer so I can see what your brain is doing. I will then play some sounds from a speaker, but you do not have to do anything about them. All you need to do is sit on a comfy couch and watch your favourite movie (without the sound on) for about 30 minutes. It is ok if you do not want to do this.

APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS ETHICS COMMITTEE ON June 5 2013 for 3 years from June 5 2013 to June 5 2016. Reference Number 9657
ASSENT FORM FOR CHILDREN

- I understand that Joan and Suzanne would like to go through some activities to do with matching facial expressions to speech.

- I understand that Joan and Suzanne would also like to record my speech with a microphone, when I am saying some sentences with different emotions, and when I am telling a story.

- I understand that I will need to repeat these activities 4 times in total.

- I understand that Joan and Suzanne would like me to try wearing a small device in my ear to see if I will be able to hear better.

- I understand that I will get one week to get used to the device, and that Joan will help guide me through using it.

- I understand that Joan and Suzanne would like to go through some activities on a computer during some short sessions, 3 times a week, for 3 weeks, while I am wearing the hearing device.

- I understand that Joan will be visiting, and completing activities with me, at different times during a period of 9 weeks.

- I know that I can tell Joan and Suzanne if I do not want to do the activities anymore, and that this will be fine.

- I understand that Joan will come to my house, or anywhere that is easy for me, to do these activities.

- I understand that Joan and Suzanne will not use my name when they talk about the results or when they write about the results.

- I agree to take part in this research project.

- I understand what Joan has told me about recording my brain activity, and I would like to do that also:

  Please circle   YES / NO

If you would like to do this, write or sign your name on the line below.

If you decide later that you don’t want to do this with us anymore, we can stop at any time.

My name: ______________________   My signature: ______________________
Title of Research:
DOES THE USE OF REMOTE-MICROPHONE HEARING AID TECHNOLOGY ENHANCE THE EFFECTIVENESS OF PROSODY PERCEPTION TRAINING IN CHILDREN WITH AUTISM SPECTRUM DISORDER (ASD)?

Researchers: Joan Leung, and Professor Suzanne Purdy

Thank you for taking an interest in this project!

Your child is invited to take part in this research being conducted by staff and students within the Department of Psychology (Speech Science). Before you decide, it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully and discuss it with others if you wish. Ask us if there is anything that is not clear or if you would like more information.

What is the purpose of the study?

Prosody refers to the variation in timing, pitch or stress pattern in speech that helps to convey meaning and emotion. Affective prosody, in particular, refers to both verbal (speech) and non-verbal (facial) emotional expression. Research has shown that individuals with ASD are impaired in the perception, as well as the production, of affective prosody.

There is growing evidence that computerised tools offer an interactive platform for children with ASD to learn and express themselves. Many applications (apps) have already been developed to aid children with communicative difficulties to ‘speak’ through another medium. Apps and games also exist to help individuals with ASD, for example, organise a daily routine, and learn socially appropriate behaviour. So far, those related to the recognition of social cues have mostly focused on facial expressions, instead of prosodic cues in speech.

In recent years, there has also been a progressive increase in research studies involving remote microphone hearing aids (RM systems). The use of a personal RM system has been linked to 1) significant improvements in speech recognition in noise; 2) a significant increase in on-task behaviours in the classroom; and 3) an improvement in teacher-rated listening behaviours, for
Children with ASD/AHDH/or both. Improvements in spatial listening and temporal processing for school-aged children with ASD, after they underwent a 6-week trial of an RM system, has also been found.

This project aims to pair the use of remote-microphone hearing aids with a computer-based training intervention, to see if abilities in affective prosodic perception and production will improve in children with Autism Spectrum Disorder (ASD), and to see if there is an auditory processing issue at the basis of observed social communication difficulties in this population.

What is expected from you, and from your child?

1. Time commitment

Participation in this study would require an extended commitment that spans approximately 10 weeks. Before you continue reading the details below, we would like to inform you that the researcher would be more than happy to cater to your convenience, with regards to both your time and the location of where the sessions would take place.

First, there will be an introductory session that will last around 1 hour. During this time, you will receive all the information that you will need about the project. You and your child will also be asked to complete the first round of assessments, please see below for details. The 2nd session would take place two weeks later. You will again be required to complete a round of assessments, exactly the same as the last ones.

Following this, there will be a 1-week period where your child will be introduced to, and taught how to use an RM system. The researcher will come by every day, for 5 days, to go through 15-minute sessions aimed at familiarising your child with the hearing device. After that, there will be a 3-week period where your child gets to participate in some training activities on a laptop. We will require three 30-minute sessions per week of you and your child’s time. The researcher brings the laptop to the session, and will go through a series of activities with your child.

During this training period, you and your child will be given the opportunity to keep the RM system and trial it in school. Your support will be required to help your child wear the device, and to log the number of hours your child wears it.

After this training period, you and your child will need to be assessed again for the 3rd time. The same procedure (as the sessions before training) will once again require 1 hour of your time. And finally, the last session will take place two weeks later. After this, you will be more than welcome to ask any further questions with regards to the project, and you will also be gifted with a $20 shopping voucher.

You and your child will also be given the opportunity to participate in a part of the study that involves recording the brain activity in response to different emotional tones in speech. This will be a 1-1.5 hour session, once before assessment session 1, and once after assessment session 4. It will involve the researcher sticking 8 electrodes to the scalp of your child’s head, and playing a series of speech samples through a loud speaker, while your child sits on a couch and watches their movie of choice (with the sound turned off). This part of the study is entirely optional.
2. Assessments

*Questionnaires*

The study requires you, as a parent/caregiver, to complete two questionnaires (10-15 minutes per questionnaire). A routine screen for Autism symptoms, and an assessment of your child’s communication abilities, will be done once during the first session. You can complete this while your child is being tested. Please be aware that the questionnaires are screening tools only. If the results are suggestive of previously undiagnosed communication problems or ASD, any necessary referrals will be made, and you can further discuss the matter with the research team. Please note that it is University of Auckland policy that we notify you regarding any significant findings from these questionnaires. You may choose not to participate if you do not wish to be informed of these results.

*Assessment of recognition and matching of facial expressions to speech*

Your child will be asked to participate in an assessment that requires them to listen to some segments of speech, and then match them to one facial expression out of a selection of six photos. We assure you that the photos are just portraits of people, of different genders, ages and ethnicities, making different facial expressions. There are no offensive or frightening photos. The speech segments also contain no offensive language or statements, and just consist of people saying various generic sentences with different emotions.

*Assessment of producing emotion in speech*

Your child will also be asked to participate in saying some sentences of their own, in different emotions. They will be asked to say “it is eleven o’clock” in either an angry, happy or sad tone. They will be asked to repeat these a few times, so that we get enough data. To record your child’s speech, a microphone headset will be worn, with the receiver placed 5cm from the mouth. No other devices will be attached to your child body or clothes.

*Assessment of affective prosodic comprehension*

Finally, your child will be asked to listen to various emotional speech samples taken from audio children’s books. They will be asked to identify the underlying emotion of those samples.

3. Remote-microphone hearing aid (RM system) familiarisation and use in school:

Familiarisation period of 5 days (15 minutes session per day) – introduction to researcher and RM system; explaining, modelling, and using the RM system.

Include the option for participants to wear the RM system during their school day, during the 3-week training period. Support of parents and teachers will be obtained in order for them to support the child to wear the device during school. The number of hours worn per day will be logged by either the teacher (if the child wanted to take the device off while at
school), or the parent (when the child arrives home after wearing the device for a full day at school).

4. Training

There will be four different tasks to complete during each 30-minute training session. Your child will be asked to complete two matching tasks, similar to the first assessment task. They will need to identify specific emotions out of a selection of different expressions, and match them either to the researcher’s prompts, or to various speech samples.

They will also be required to listen to two speech samples at a time, and identify whether the underlying emotions match or do not match.

Finally, there will be a facial expression building task, where your child will be encouraged to put together pieces of a puzzle in order to come up with different portraits of emotional expressions.

Participation

Participation in the study is voluntary so it is up to you and your child to decide whether or not to take part. You are more than welcome to take part in the RM system and computer-training part only, if your child does not feel comfortable with the brain activity recordings. If you do decide to take part you will be given this information sheet to keep and you will be asked to sign a consent form for your child. If you decide to take part, you and your child are still free to withdraw at any time and without giving a reason. You can also request for your data to be destroyed until 29th February 2016.

Will our taking part in this study be kept confidential?

All participants will be given a unique number when they join the study. This will be used by us, instead of their name, on all of their measurements. If the information you provide is reported or published, this will be done in a way that does not identify you as its source.

You can be assured that the questionnaires will only be accessed by the researcher, and in any publication, the results will be reported in aggregate and participants will be unable to be identified.

What will happen to the results of the research study?

The results of the study will be used to help us design future studies. It is also normal practice for the results of these studies to be published in the literature. If you are interested then we will send you a summary of the results (please tick the box to request a summary on your consent form).

Who should you contact if you wish to obtain further information?

You can contact Professor Suzanne Purdy or Joan Leung (details are below)

Prof. Suzanne Purdy
Head of Speech Science
School of Psychology
The University of Auckland
Private Bag 92019, Auckland 1142
Tel: (09) 373 7599 ext. 82073
E-mail: sc.purdy@auckland.ac.nz

Joan Leung
PhD student (main researcher)
Mob: 021 842 349
E-mail: joan.leung@auckland.ac.nz

The Head of the Department of Psychology is:
Professor William Hayward, PhD
School of Psychology
The University of Auckland
Private Bag 92019, Auckland 1142
Tel: (09) 373 7599 ext. 88516
Email: w.hayward@auckland.ac.nz

For any queries regarding ethical concerns you may contact the Chair, The University of Auckland Human
Participants Ethics Committee, The University of Auckland, Office of the Vice Chancellor, Private Bag 92019,
Auckland 1142. Telephone 09 373-7599 extn. 83711

Thank you for reading this information sheet.

APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS ETHICS COMMITTEE ON June 5 2013 for
3 years from June 5 2013 to June 5 2016. Reference Number 9657
CONSENT FORM FOR PARENTS/CAREGIVERS

Title of Research: DOES THE USE OF REMOTE-MICROPHONE HEARING AID TECHNOLOGY ENHANCE THE EFFECTIVENESS OF PROSODY PERCEPTION TRAINING IN CHILDREN WITH AUTISM SPECTRUM DISORDER (ASD)?

Researchers: Joan Leung, and Professor Suzanne Purdy

This consent form will be retained for six years in secure storage by the researchers before being destroyed

Please initial each box

1. I confirm that I have read and understood the information sheet for the above study and have had the opportunity to ask questions.

2. I understand that the experiment will take 12 hours over a period of 10 weeks.

3. I understand that I will be asked to participate by completing two questionnaires on behalf of my child.

4. I understand that my child and I will be given the opportunity to:
   a) Participate in the brain activity recording part of the study (YES / NO)
   b) In addition to computer-based training at home, also trial the RM system at school (YES / NO)

5. I understand that my participation and my child's participation are voluntary and that we are free to withdraw at any time, without giving any reason.
6. I understand that the study will include a screening questionnaire that looks for features of ASD. I understand that if any difficulties are identified in my child I will have the opportunity to discuss this with the research team.

7. I understand that I have the right to withdraw the data until 29/02/2016.

8. I agree for my child to take part in this study.

9. I understand that the questionnaires will only be accessed by the researcher, and in any publication, the results will be reported in aggregate and participants will be unable to be identified.

10. I would like a summary of the results

   Please circle: YES / NO

   _______________________________________________________
   PRINT NAME (Parent)

   _______________________________________________________
   PRINT NAME (Child)

   _______________________________________________________
   PRINT ADDRESS

   TELEPHONE: EMAIL:

   We need this information so we can arrange a convenient appointment time, as well as send you a summary of the results if you want.

   ____________________________ __________________________
   SIGNATURE DATE

   APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS ETHICS COMMITTEE ON June 5 2013 for 3 years from June 5 2013 to June 5 2016. Reference Number 9657
PARTICIPANT INFORMATION FOR CHILDREN

Title: TRAINING AFFECTIVE PROSODY in CHILDREN with AUTISM SPECTRUM DISORDER (ASD)

Researchers: Joan Leung, and Professor Suzanne Purdy

Read out the following:

Suzanne, from the University of Auckland has been working with children for many years, especially those who have difficulties communicating. I (Joan), along with Suzanne, want to invite you to take part in our project! I’m going to explain a little about what is going to happen if you do decide to help us out.

You will come to visit me at the university on a day that you are free. This visit would take about 2.5 hours.

We will go through 2 activities. For the first one, I will need you to listen to some people speaking, and then match it to a photo of that person’s facial expression. For the second task, I will need to put a microphone around your head, and record you saying “it is eleven o’clock” in your best angry, happy and sad voices. We’ll do this a few times so that we get the best recording of your voice.

Lastly, if it is alright with you, we will have a go at recording your brain waves. I will stick 8 wires on to the skin of your head. This will not hurt at all. Then you can choose your favourite movie, and you can watch it with the sound turned off. While you are watching the movie, I will play some speech sounds through a speaker in the room. You do not have to do anything – you just sit back and enjoy your movie.

If you are ok with all of this, I would love it if you could help me out with this project!

APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS ETHICS COMMITTEE ON June 5 2013 for 3 years from June 5 2013 to June 5 2016. Reference Number 9657
ASSENT FORM FOR CHILDREN

- I understand that Joan and Suzanne would like to go through some activities to do with matching facial expressions to speech.

- I understand that Joan and Suzanne would also like to record my speech with a microphone when I am saying some sentences with different emotions.

- I understand that Joan will record my brain waves, and that it will require 8 wires to be stuck on to my head, which will not hurt.

- I understand that I will need to visit Joan at the university, and that this visit will last 2.5 hours.

- I know that I can tell Joan and Suzanne if I do not want to do the activities anymore, and that this will be fine.

- I understand that Joan and Suzanne will not use my name when they talk about the results or when they write about the results.

- I agree to take part in this research project.

If you would like to do this, write or sign your name on the line below.
If you decide later that you don’t want to do this with us anymore, we can stop at any time.

My name: ______________________    My signature: ______________________
Appendix 11: Participant information sheet (Study 2) – for parents/caregivers of Controls

PARTICIPANT INFORMATION SHEET FOR PARENTS/CAREGIVERS

Title of Research:
TRAINING AFFECTIVE PROSODY in CHILDREN with AUTISM SPECTRUM DISORDER (ASD)

Researchers: Joan Leung, and Professor Suzanne Purdy

Thank you for taking an interest in this project!

Your child is invited to take part in this research being conducted by staff and students within the School of Psychology (Speech Science). Before you decide, it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully and discuss it with others if you wish. Ask us if there is anything that is not clear or if you would like more information.

What is the purpose of the study?

Prosody refers to the variation in timing, pitch or stress pattern in speech that helps to convey meaning and emotion. Affective prosody, in particular, refers to both verbal (speech) and non-verbal (facial) emotional expression. Research has shown that individuals with ASD are impaired in the perception, as well as the production, of affective prosody. This project aims to better understand and improve communication outcomes for children with ASD.

What is expected from you, and from your child?

1. Your role

We want to compare the perception and production of emotional cues between children with ASD and typically developing children. By participating, you will be part of the comparison group, and will provide us with information such as the extent of communication difficulties in children with ASD compared to their peers.

2. Time commitment
Participation in this study would require a 2 – 2.5 hours visit to the Tamaki Campus of the University of Auckland (St Johns area). After this, you will be more than welcome to ask any further questions with regards to the project, and you will also be given a $20 shopping voucher for your time.

3. Assessments

Questionnaires

The study requires you, as a parent/caregiver, to complete a questionnaire about your child’s communication abilities, as observed during the past week. You can complete this while your child is being tested. Please be aware that the questionnaires are screening tools only. If the results are suggestive of previously undiagnosed communication problems, any necessary referrals will be made, and you can further discuss the matter with the research team. Please note that it is University of Auckland policy that we notify you regarding any significant findings from these questionnaires. You may choose not to participate if you do not wish to be informed of these results.

Assessment of recognition and matching of facial expressions to speech

Your child will be asked to participate in an assessment that requires them to listen to some segments of speech, and then match them to one facial expression out of a selection of six photos. We assure you that the photos are just portraits of people, of different genders, ages and ethnicities, making different facial expressions. There are no offensive or frightening photos. The speech segments also contain no offensive language or statements, and just consist of people saying various generic sentences with different emotions.

Assessment of producing emotion in speech

Your child will also be asked to participate in saying some sentences of their own, in different emotions. They will be asked to say “it is eleven o’clock” in either an angry, happy or sad tone. They will be asked to repeat these a few times, so that we get enough data. To record your child’s speech, a microphone will be worn, with the receiver placed 5cm away from the mouth. No other devices will be attached to your child body or clothes.

Recording of brain activity

Finally, your child will be asked to participate in a recording of their brain activity in response to different emotional tones in speech. This will involve the researcher sticking (non-permanent) 8 electrodes onto the scalp and forehead of your child. Your child will be invited to sit on a sofa and watch a movie of their choice, with the sound muted and the subtitles turned on. The researcher will play a series of speech sounds through a loud speaker, to which your child will be instructed to ignore and just enjoy the movie. This part of the session should take no longer than an hour.

Participation

Participation in the study is voluntary so it is up to you and your child to decide whether or not to take part. If you do decide to take part you will be given this information sheet to keep and you will be asked to sign a consent form for your child. If you decide to take part, you and your child are still
free to withdraw at any time and without giving a reason. You can also request for your data to be destroyed until 29th February 2016.

Will our taking part in this study be kept confidential?

All participants will be given a unique number when they join the study. This will be used by us, instead of their name, on all of their measurements. If the information you provide is reported or published, this will be done in a way that does not identify you as its source.

You can be assured that the questionnaires will only be accessed by the researcher, and in any publication, the results will be reported in aggregate and participants will be unable to be identified.

What will happen to the results of the research study?

The results of the study will be used to help us design future studies. It is also normal practice for the results of these studies to be published in the literature. If you are interested then we will send you a summary of the results (please tick the box to request a summary on your consent form)

Who should you contact if you wish to obtain further information?

You can contact Professor Suzanne Purdy or Joan Leung (details are below)

Professor Suzanne Purdy  
Head of Speech Science  
The University of Auckland  
Private Bag 92019  
Tel: 373 7599 ext. 82073  
E-mail: sc.purdy@auckland.ac.nz  
Joan Leung  
PhD student (main researcher)  
Mob: 021 842 349  
E-mail: joan.leung@auckland.ac.nz

The Head of the Department of Psychology is:  
Professor William Hayward, PhD  
School of Psychology  
The University of Auckland  
Private Bag 92019, Auckland 1142  
Tel: (09) 373 7599 ext. 88516  
Email: w.hayward@auckland.ac.nz

For any queries regarding ethical concerns you may contact the Chair, The University of Auckland Human Participants Ethics Committee, The University of Auckland, Office of the Vice Chancellor, Private Bag 92019, Auckland 1142. Telephone 09 373-7599 extn. 83711

Thank you for reading this information sheet.

APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS ETHICS COMMITTEE ON June 5 2013 for 3 years from June 5 2013 to June 5 2016. Reference Number 9657
CONSENT FORM FOR PARENTS/CAREGIVERS

Title of Research: TRAINING AFFECTIVE PROSODY in CHILDREN with AUTISM SPECTRUM DISORDER (ASD)

Researchers: Joan Leung, and Professor Suzanne Purdy

This consent form will be retained for six years in secure storage by the researchers before being destroyed

Please initial each box

1. I confirm that I have read and understood the information sheet for the above study and have had the opportunity to ask questions.

2. I understand that the experiment will take 2 – 2.5 hours.

3. I understand that I will be asked to participate by completing a questionnaire.

4. I understand that my participation and my child's participation are voluntary and that we are free to withdraw at any time, without giving any reason.

5. I understand that I have the right to withdraw the data until 29/02/2016.
6. I agree for my child to take part in this study.

7. I understand that the questionnaires will only be accessed by the researcher, and in any publication, the results will be reported in aggregate and participants will be unable to be identified.

8. I would like a summary of the results

   Please circle: YES / NO

__________________________________________________
PRINT NAME (Parent)
__________________________________________________
PRINT NAME (Child)

__________________________________________________
PRINT ADDRESS

TELEPHONE: EMAIL:

We need this information so we can send you a summary of the results if you want.

________________________________________   _______________________
SIGNATURE                      DATE

APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS ETHICS COMMITTEE ON June 5 2013 for 3 years from June 5 2013 to June 5 2016. Reference Number 9657
PARTICIPANT INFORMATION SHEET FOR ADULT CONTROL PARTICIPANTS

Title of Research:
TRAINING AFFECTIVE PROSODY in CHILDREN with AUTISM SPECTRUM DISORDER (ASD)

Researchers: Joan Leung, and Professor Suzanne Purdy

Thank you for taking an interest in this project!

You are invited to take part in this research being conducted by staff and students within the School of Psychology (Speech Science). Before you decide, it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully and discuss it with others if you wish. Ask us if there is anything that is not clear or if you would like more information.

What is the purpose of the study?

Prosody refers to the variation in timing, pitch or stress pattern in speech that helps to convey meaning and emotion. Affective prosody, in particular, refers to verbal (speech) emotional expression. Research has shown that individuals with ASD are impaired in the perception, as well as the production, of affective prosody. This project aims to better understand and improve communication outcomes for children with ASD.

What is expected from you, and from your child?

1. Your role

   We want to compare the perception of emotional cues between children with ASD, typically developing children, and typically developed adults. By participating, you will be part of the adult comparison group, and will provide us with information such as how there may be potential atypical neural activity in individuals with ASD that underlies these communication difficulties.

2. Time commitment
Participation in this study would require a 1.5 hour visit to the Tamaki Campus of the University of Auckland (St Johns area). After this, you will be more than welcome to ask any further questions with regards to the project, and you will also be given a $20 shopping voucher for your time.

3. Recording of brain activity

You will be asked to participate in a recording of your brain activity in response to different emotional tones in speech. This will involve the researcher sticking (non-permanent) 8 electrodes onto your scalp and forehead. You will be invited to sit on a sofa and watch a movie of your choice, with the sound muted and the subtitles turned on. The researcher will play a series of speech sounds through a loud speaker, to which you will be instructed to ignore and just enjoy the movie.

Participation

Participation in the study is voluntary so it is up to you to decide whether or not to take part. If you do decide to take part you will be given this information sheet to keep and you will be asked to sign a consent form. If you decide to take part, you are still free to withdraw at any time and without giving a reason. You can also request for your data to be destroyed until 29th February 2016.

Will our taking part in this study be kept confidential?

All participants will be given a unique number when they join the study. This will be used by us, instead of their name, on all measurements. If the information you provide is reported or published, this will be done in a way that does not identify you as its source.

You can be assured that the data will only be accessed by the researcher, and in any publication, the results will be reported in aggregate and participants will be unable to be identified.

What will happen to the results of the research study?

The results of the study will be used to help us design future studies. It is also normal practice for the results of these studies to be published in the literature. If you are interested then we will send you a summary of the results (please tick the box to request a summary on your consent form).

Who should you contact if you wish to obtain further information?

You can contact Professor Suzanne Purdy or Joan Leung (details are below)

Professor Suzanne Purdy
Head of Speech Science
The University of Auckland
Private Bag 92019
Tel: 373 7599 ext. 82073
E-mail: sc.purdy@auckland.ac.nz

Joan Leung
PhD Candidate (main researcher)
Mob: 021 842 349
E-mail: joan.leung@auckland.ac.nz

The Head of the Department of Psychology is:
Professor William Hayward, PhD
School of Psychology
The University of Auckland
Private Bag 92019, Auckland 1142
Tel: (09) 373 7599 ext. 88516
Email: w.hayward@auckland.ac.nz

For any queries regarding ethical concerns you may contact the Chair, The University of Auckland Human Participants Ethics Committee, The University of Auckland, Office of the Vice Chancellor, Private Bag 92019, Auckland 1142. Telephone 09 373-7599 extn. 83711

Thank you for reading this information sheet.

APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS ETHICS COMMITTEE ON June 5 2013 for 3 years from June 5 2013 to June 5 2016. Reference Number 9657
CONSENT FORM FOR ADULT CONTROL PARTICIPANTS

Title of Research: TRAINING AFFECTIVE PROSODY in CHILDREN with AUTISM SPECTRUM DISORDER (ASD)

Researchers: Joan Leung, and Professor Suzanne Purdy

This consent form will be retained for six years in secure storage by the researchers before being destroyed

Please initial each box

1. I confirm that I have read and understood the information sheet for the above study and have had the opportunity to ask questions.

2. I understand that the experiment will take 1.5 hours.

3. I understand that I will be asked to participate by having my brain activity recorded.

4. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason.

5. I understand that I have the right to withdraw the data until 29/02/2016.
6. I understand that the data will only be accessed by the researcher, and in any publication, the results will be reported in aggregate and participants will be unable to be identified.

7. I agree to take part in this study.

8. I would like a summary of the results

   Please circle: YES / NO

__________________________________________________
PRINT NAME

__________________________________________________
PRINT ADDRESS

TELEPHONE: EMAIL:

*We need this information so we can send you a summary of the results if you want.*

_________________________________          _________________________
SIGNATURE                          DATE

APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS ETHICS COMMITTEE ON June 5 2013 for 3 years from June 5 2013 to June 5 2016. Reference Number 9657
Appendix 13: Participant information sheet (Study 2) – for Teachers

PARTICIPANT INFORMATION SHEET FOR THE SCHOOL

Title of Research:
DOES THE USE OF REMOTE-MICROPHONE HEARING AID TECHNOLOGY ENHANCE THE EFFECTIVENESS OF PROSODY PERCEPTION TRAINING IN CHILDREN WITH AUTISM SPECTRUM DISORDER (ASD)?

Researchers: Joan Leung, and Professor Suzanne Purdy

Thank you for your willingness to support this project!

Your student is invited to take part in this research being conducted by staff and students within the Department of Psychology (Speech Science). Before you decide, it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully and discuss it with others if you wish. Ask us if there is anything that is not clear or if you would like more information.

What is the purpose of the study?

Prosody refers to the variation in timing, pitch or stress pattern in speech that helps to convey meaning and emotion. Affective prosody, in particular, refers to both verbal (speech) and non-verbal (facial) emotional expression. Research has shown that individuals with ASD are impaired in the perception, as well as the production, of affective prosody.

There is growing evidence that computerised tools offer an interactive platform for children with ASD to learn and express themselves. Many applications (apps) have already been developed to aid children with communicative difficulties to ‘speak’ through another medium. Apps and games also exist to help individuals with ASD, for example, organise a daily routine, and learn socially appropriate behaviour. So far, those related to the recognition of social cues have mostly focused on facial expressions, instead of prosodic cues in speech.

In recent years, there has also been a progressive increase in research studies involving remote microphone hearing aids (RM systems). The use of a personal RM system has been linked to 1) significant improvements in speech recognition in noise; 2) a significant increase in on-task behaviours in the classroom; and 3) an improvement in teacher-rated listening behaviours, for children with ASD/AHDH/or both. Improvements in spatial listening and temporal processing for
school-aged children with ASD, after they underwent a 6-week trial of an RM system, has also been found.

This project aims to pair the use of remote-microphone hearing aids with a computer-based training intervention, to see if abilities in affective prosodic perception and production will improve in children with Autism Spectrum Disorder (ASD), and to see if there is an auditory processing issue at the basis of observed social communication difficulties in this population.

**What is expected from the school?**

1. **RM system trial in school**

   As a part of the study, the child is encouraged to try wearing the RM system during their school day (except for lunch break and physical education), for a 3-week trial period.

   Support from the child’s teacher will be needed to ensure that the system stays on and is working in the classroom. The teacher will be required to wear a small transmitter around their neck, with a microphone clipped onto the collar of their clothes. This transmitter will send a clear signal of the teacher’s voice straight to 2 receivers, which will be worn by the child. If needed, the researcher is happy to visit the teacher at school to brief him/her about the system and the research.

   The teacher (and maybe accompanied by a teacher aide or the SENCO of the school) will also be invited to help by answering 2 rounds of questionnaires. Once before the RM trial, and once after. The questionnaires are split into an evaluation of listening behaviours in the classroom before- and after- an intervention is introduced.

2. **Other parts of the research study**

   Outside of school hours, the child will also be involved in:
   - Completing assessments about social communication
   - Completing 9 sessions of computer-based training activities designed to improve emotion perception in faces and speech.
   - Optional participation in the recording of brain activity in response to different emotional tones in speech.

**Participation**

Participation in the study is voluntary, for both the child, and the school. If you decide to take part, both the family and the school of the child are free to withdraw at any time and without giving a reason. You can also request for your data to be destroyed until 29th February 2016.

**Will our taking part in this study be kept confidential?**

All participants will be given a unique number when they join the study. This will be used by us, instead of their name, on all of their measurements. If the information you provide is reported or published, this will be done in a way that does not identify you as its source.

You can be assured that the questionnaires will only be accessed by the researcher, and in any publication, the results will be reported in aggregate and participants will be unable to be identified.
What will happen to the results of the research study?

The results of the study will be used to help us design future studies. It is also normal practice for the results of these studies to be published in the literature. If you are interested then we will send you a summary of the results (please tick the box to request a summary on your consent form).

Who should you contact if you wish to obtain further information?

You can contact Professor Suzanne Purdy or Joan Leung (details are below):

Prof. Suzanne Purdy  
Head of Speech Science  
School of Psychology  
The University of Auckland  
Private Bag 92019, Auckland 1142  
Tel: (09) 373 7599 ext. 82073  
E-mail: sc.purdy@auckland.ac.nz

Joan Leung  
PhD student (main researcher)  
Mob: 021 842 349  
E-mail: joan.leung@auckland.ac.nz

The Head of the Department of Psychology is:  
Professor William Hayward, PhD  
School of Psychology  
The University of Auckland  
Private Bag 92019, Auckland 1142  
Tel: (09) 373 7599 ext. 88516  
Email: w.hayward@auckland.ac.nz

For any queries regarding ethical concerns you may contact the Chair, The University of Auckland Human Participants Ethics Committee, The University of Auckland, Office of the Vice Chancellor, Private Bag 92019, Auckland 1142. Telephone 09 373-7599 extn. 83711

Thank you for reading this information sheet.

APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS ETHICS COMMITTEE ON June 5 2013 for 3 years from June 5 2013 to June 5 2016. Reference Number 9657
CONSENT FORM FOR THE SCHOOL

Title of Research: DOES THE USE OF REMOTE-MICROPHONE HEARING AID TECHNOLOGY ENHANCE THE EFFECTIVENESS OF PROSODY PERCEPTION TRAINING IN CHILDREN WITH AUTISM SPECTRUM DISORDER (ASD)?

Researchers: Joan Leung, and Professor Suzanne Purdy

This consent form will be retained for six years in secure storage by the researchers before being destroyed

Please initial each box

1. I confirm that I have read and understood the information sheet for the above study and have had the opportunity to ask questions.

2. I understand that the child will be given the opportunity to trial the RM system at school.

3. I understand that the school’s participation is voluntary and that we are free to withdraw at any time, without giving any reason.

4. I understand that I have the right to withdraw the data until 29/02/2016.

5. I agree for the school to take part in this study.

6. I understand that the questionnaires will only be accessed by the researcher, and in any publication, the results will be reported in aggregate and participants will be unable to be identified.
7. I would like a summary of the results

Please circle: YES / NO

PRINT NAME (Principal)

PRINT NAME (Child)

PRINT ADDRESS

TELEPHONE: EMAIL:

We need this information so we can arrange meetings with teachers if necessary, as well as to send you a summary of the results if you want.

SIGNATURE DATE

APPROVED BY THE UNIVERSITY OF AUCKLAND HUMAN PARTICIPANTS ETHICS COMMITTEE ON June 5 2013 for 3 years from June 5 2013 to June 5 2016. Reference Number 9657
Appendix 14: Participant questionnaire

PARTICIPANT NUMBER: ____________
DATE: ____________

TRAINING AFFECTIVE PROSODY in CHILDREN with AUTISM SPECTRUM DISORDER (ASD)

PARTICIPANT QUESTIONNAIRE

About You

1. Please indicate if you are male or female
   M  F

2. Please tell us your age in years
   ____________

3. Please indicate your ethnic identity (tick all that apply)
   NZ European
   NZ Maori
   Pacific peoples
   Asian
   European
   Other
   Please indicate
   ____________

4. Is English your first language?
   Yes
   No
   If no, please indicate your first language:
   ____________

5. Other conditions
   Yes  No
   If yes, please indicate:
   E.g. ADHD, dyslexia
   ____________

6. Have you ever been tested for hearing loss?
   Yes  No
   If yes, please provide report
   ____________
   If no, DPOAE screen result:
   ____________

7. What activities do you do during the week?
   Type  Hours per week
   ____________
   ____________
   ____________
Appendix 15: Childhood Autism Rating Scale (CARS-2) – Parents/Caregivers questionnaire

INSTRUCTIONS

This form asks about behaviors in several areas where people may have difficulty. The person you are rating may or may not have ever shown these behaviors.

For each behavior listed, please make a check mark under the description that best describes the person you are rating. Check the box under Don’t Know if you do not have enough information about a behavior to give a rating. It is important to provide an answer for every behavior. After each section, there is space for you to give one or more brief, specific examples that relate to your ratings in that section. Use the blank page at the end of the form if you need extra space. The final section of this questionnaire provides spaces where you can describe any other behaviors that you would like us to know about.
### SECTION 1

**How does the person you are rating communicate?**

<table>
<thead>
<tr>
<th>1. Imitates sounds, words, and movements of others</th>
<th>□ □ □ □ □</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. Responds to facial expressions, gestures, and different tones of voice used by others</td>
<td>□ □ □ □ □</td>
</tr>
<tr>
<td>3. Responds to his or her name being called by turning and making eye contact with the person calling his or her name</td>
<td>□ □ □ □ □</td>
</tr>
<tr>
<td>4. Directs facial expressions to others to show the emotions he or she is feeling</td>
<td>□ □ □ □ □</td>
</tr>
<tr>
<td>5. Uses a variety of gestures (pointing, nodding the head, showing the size of something) that are coordinated with words or used to explain things when he or she doesn’t have the words to do so</td>
<td>□ □ □ □ □</td>
</tr>
</tbody>
</table>

*If the person you are rating is not using words, skip ahead to Section 2.*

<table>
<thead>
<tr>
<th>6. Uses made-up words or repeats specific words or phrases</th>
<th>□ □ □ □ □</th>
</tr>
</thead>
<tbody>
<tr>
<td>7. Has an unusual tone, rhythm, loudness, or rate of speech</td>
<td>□ □ □ □ □</td>
</tr>
<tr>
<td>8. Speech is overly formal; for example, uses vocabulary that seems more sophisticated than usual for a person of his or her age or for the situation</td>
<td>□ □ □ □ □</td>
</tr>
<tr>
<td>9. Carries on a conversation with another person that flows back and forth, at a level you would expect for someone of his or her age</td>
<td>□ □ □ □ □</td>
</tr>
<tr>
<td>10. Can talk with another person about that person’s interests</td>
<td>□ □ □ □ □</td>
</tr>
</tbody>
</table>

**Examples:** Give one or more brief but specific examples of the problem behaviors rated above. If you need more space to write, use the blank page at the end of this form.

[continue on next page...]

250
SECTION 2

How does the person you are rating relate to others and show emotion?

1. Makes eye contact when speaking with or listening to another person ..........................................

2. Points to and shares things of interest with others .................................................................

3. Follows another person’s gaze or points toward an object that is out of reach ..........................................

4. Is responsive to social initiations from others ..............................................................................

5. Initiates social interactions with adults and peers (not just to get a basic need met) ..................

6. Sustains an interaction with others in an easy, flowing, back-and-forth manner .........................

7. Makes and maintains friendships with peers of same developmental level .................................

8. Shows a range of emotional expressions that match the situation (for example, smiles, frowns, conveys different emotions through eyes and facial expressions, etc.) .................................................................

9. Understands and responds to how another person may be thinking or feeling (for example, tries to comfort someone in distress, does something because he or she thinks the other person will like it) .................................................................

Examples: Give one or more brief but specific examples of the problem behaviors rated above. If you need more space to write, use the blank page at the end of this form.
SECTION 3

How does the person you are rating move his or her body?

1. Has unusual ways of moving fingers, hands, arms, legs; or spins or rocks body ................................................................. □ □ □ □ □ □ □

2. Does things that might result in self-injury, like scratching, head banging, picking at his or her skin ................................................................. □ □ □ □ □ □ □

3. Is clumsy, stumbles, or has an awkward walk or run ................................................................. □ □ □ □ □ □ □

4. For school-aged children or adults: Has difficulty tying shoes or difficulty with handwriting or other tasks that require fine motor coordination ................................................................. □ □ □ □ □ □ □

SECTION 4

How does the person you are rating play?
(For an older individual, how did he or she play as a child?)

1. Uses only parts of toys instead of whole toys, or plays with objects (e.g., opens and closes toy barn doors, spins wheels on cars, wobbles or spins household objects) ................................................................. □ □ □ □ □ □ □

2. Plays with the same things in the same way over and over ................................................................. □ □ □ □ □ □ □

3. Uses toys or other materials to represent something they are not (e.g., uses a banana as a phone or a microphone) ................................................................. □ □ □ □ □ □ □

4. Engages in make-believe play, taking on a role (not based on scripts from movies or TV shows) ................................................................. □ □ □ □ □ □ □

Examples: Give one or more brief but specific examples of the problem behaviors rated above. If you need more space to write, use the blank page at the end of this form.
SECTION 5
How does the person you are rating react to new experiences and changes in routine?

1. May show anxiety or worry in facial expression or body movement, or by becoming overly impatient ................................................................. [ ] [ ] [ ] [ ] [ ]

2. May show worry about the same thing over and over again ................................................................. [ ] [ ] [ ] [ ] [ ]

3. Copes with changes in routine or the environment
   (for example, moving furniture) ................................................................. [ ] [ ] [ ] [ ] [ ]

4. Has specific routines or specific ways things must be done
   by self or others ................................................................................................................ [ ] [ ] [ ] [ ] [ ]

5. Has special interests or topics (for example, dinosaurs, trains, clocks, weather, license plates) ................................................................. [ ] [ ] [ ] [ ] [ ]

SECTION 6
How does the person you are rating use his or her senses of vision, hearing, touch, and smell?

1. Tends to look at objects from unusual angles or out of the corner of his or her eyes ................................................................. [ ] [ ] [ ] [ ] [ ]

2. Is overly interested in light from mirrors or light reflecting off objects ................................................................. [ ] [ ] [ ] [ ] [ ]

3. Is overly sensitive to some sounds, smells, or textures; seeks some out, actively avoids others ................................................................. [ ] [ ] [ ] [ ] [ ]

4. Has an unusual response to touch; may overreact to touch or pain or may not respond to things that others would find uncomfortable or painful ................................................................. [ ] [ ] [ ] [ ] [ ]

Examples: Give one or more brief but specific examples of the problem behaviors rated above. If you need more space to write, use the blank page at the end of this form.
SECTION 7
Other Behaviors

1. Does this individual have any extremely unusual mathematical, reading, or artistic abilities?  No  Yes (please explain)

2. Are there other unusual behaviors you have noticed that you would like to tell us about? Please list the specific behavior, and give an example or two.

Additional Behavior Examples or Comments:

Please specify the number of the question that is related to your example or comment: __________________________

Additional Behavior Examples or Comments:

Please specify the number of the question that is related to your example or comment: __________________________

Additional Behavior Examples or Comments:

Please specify the number of the question that is related to your example or comment: __________________________
Appendix 16: Childhood Autism Rating Scale (CARS-2) – High functioning version scoring guide

**Summary**

**Category Ratings**

1. Social-Emotional Understanding
   - Median: 2.5
2. Emotional Expression and Regulation of Emotions
   - Median: 2.5
3. Relating to People
   - Median: 2.5
4. Body Use
   - Median: 2.0
5. Object Use in Play
   - Median: 2.0
6. Adaptation to Change/Restricted Interests
   - Median: 2.5
7. Visual Response
   - Median: 2.0
8. Auditory Response
   - Median: 2.0
9. Taste, Smell, and Touch Response and Use
   - Median: 2.0
10. Fear or Anxiety
    - Median: 2.0
11. Verbal Communication
    - Median: 2.5
12. Nonverbal Communication
    - Median: 2.0
13. Thinking/Cognitive Integration Skills
    - Median: 2.0
14. Level and Consistency of Intellectual Response
    - Median: 2.0
15. General Impressions
    - Median: 2.5

**Total raw score =**

**Severity Group**

- Minimal-to-No Symptoms of Autism Spectrum Disorder (15–27.5)
- Mild-to-Moderate Symptoms of Autism Spectrum Disorder (28–33.5)
- Severe Symptoms of Autism Spectrum Disorder (34 and higher)

**Note:** SEM = 0.73
1. Social-Emotional Understanding

Social-emotional understanding addresses a person's cognitive understanding of others' communication, behaviors, and differing perspectives. The dimensions of social understanding that are included in this item are the ability to read the nonverbal cues of others and the ability to take another person's perspective. This item does not reflect whether someone has friends or is in a relationship. Rather, it deals with a person's ability to perceive and articulate how another person may feel or what his or her perspective may be on a given situation.

1. Age-appropriate social-emotional understanding: Clearly understands facial expressions, gestures, tone of voice, and body language of others. Able to understand that others may have a different perspective and what that perspective may be.

2. Mildly impaired social-emotional understanding: Responds to most facial expressions and expressions of emotion in others' gestures and body language, but these cues may need to be slightly exaggerated. Some subtle expressions such as relief, surprise, disgust, or embarrassment are sometimes not understood. The ability to take another's perspective is inconsistent.

3. Moderately impaired social-emotional understanding: Shows an understanding of facial expressions, tone of voice, and body language only when these cues are exaggerated. Is likely to interpret or misunderstand expressions or perspectives of others.

4. Severely impaired social-emotional understanding: Virtually no ability to understand appropriate facial expressions, gestures, tone of voice, or body language. Difficulty recognizing the perspective, understanding, or expression of others might differ.

2. Emotional Expression and Regulation of Emotions

This item refers to the capacity to express feelings and regulate one's emotions. This item is based on both direct observation and the reports of others who have witnessed this person's behavior in other settings.

1. Age-appropriate and situation-appropriate emotional response: Shows appropriate type and degree of emotional response, both verbal and nonverbal, including emotional variation such as happy, sad, angry, scared, and calm, and related internal states.

2. Mildly abnormal emotional response: Emotional expressions are relatively flat, distorted, or slightly exaggerated. Nonverbal expression of emotions does not always match verbal content. Able to describe general emotions in self but limited compared to developmental level. May have inconsistent emotional regulation problems.

3. Moderately abnormal emotional response: Expression of emotions is flat, excessive, or frequently inconsistent with situation or context of verbalized topic. May display greater emotion than expected about topics of personal or idiosyncratic concern. Difficulty to describe or understand emotional states is self-limited. Serious problems with emotional regulation that occur frequently is at least one setting.

4. Severely abnormal emotional response: Extreme problems with emotional regulation that occur in more than one setting. Responses are extreme or utilize inappropriate to situation or context of discussion. Shows extreme mood swings that are difficult to change. Expressions of emotions are exaggerated and may present as a particular emotion without understanding.
3. Relating to People

This item is related to the first two items, which also rate aspects of social relationships. This item differs in that it is confined to dimensions related to direct interpersonal interactions and the person's expression and reaction to another person. The two dimensions that are rated in this item are the person's initiation of interactions and the reciprocal nature of the interactions.

1. No evidence of difficulty or abnormality in relating to people. Age-appropriate initiation of interactions to get help, to have needs met, and for purely social purposes. Interactions with others are fixed and show a maternal, basic and forthright pattern.

2. Mildly abnormal relationships. Initiates interactions only to get obvious needs met or around special interests. Some give and take noted in interactions, but lacks continuity or fluidity or appropriateness. Aware of other people of same age and interested in interactions, but may have difficulty initiating or managing interactions. Minimal initiative for purely social purposes that does not involve special interests.

3. Moderately abnormal relationships. Initiates interactions almost totally around his or her special interests, with little attempt to engage others in these interests. Responds to overtures from others, but lacks social glue and color or responds in ways that are unusual and not always related to original overtures. Unable to maintain an interaction beyond initial overtures.

4. Severely abnormal relationships. Does not initiate any directed interactions and shows minimal response to overtures from others. Only the most persistent attempts to get the person to engage have any effect.

4. Body Use

This item represents grossly deviant body movements and also subtler forms of fine motor and coordination problems. Any obvious current deviant behaviors—including posturing, spinning, rocking, toe-walking, and self-directed aggression—automatically merit a rating of 3 or higher, depending on the persistence of the behavior. Difficulties with handwriting and tying shoes are rated on this item, with higher ratings given for problems that are so severe that the person actively resists these tasks. While this item can be scored using another's report, it is best to base your rating on current behavior. Directly observed behaviors should be given more weight than those from another's report.

1. Age-appropriate body use. Moves with the same ease, agility, and coordination as a typical person of the same age.

2. Mildly abnormal body use. Some minor peculiarities may be present, such as clumsiness, repetitive movements, poor coordination, or poor balance. May have fine motor difficulties, such as problems with handwriting or tying shoes, compared to others of the same developmental level.

3. Moderately abnormal body use. Currently displays any unusual body posture or stance, hand or finger movements, flapping, and directed aggression, picking at body, nodding, spinning, or toe-walking. Fine motor difficulties or obvious handwriting difficulties are present, which may result in persistence to written tasks.

4. Severely abnormal body use. Intense or frequent movements of the type listed above are signs of severe abnormal body use.
5. Object Use in Play

This rating includes the person's interest in and use of objects. In addition to the traditional issues related to repetitive use of parts of objects, the focus of this item also includes the degree to which the person engages in imaginative symbolic play and the degree to which toy figures are used as agents. For older persons, the rating may need to be based on the parent interview. Any obvious inappropriate or repetitive use of objects or obvious interest in parts of objects as opposed to the whole should be rated 3 or higher, depending on the persistence.

1. Appropriate interest in, and creative use of, toys and other objects. Able to spontaneously use toys in age-appropriate imaginative symbolic play and able to use objects to represent something else for whom shown interest in a variety of toys and objects.

2. Mildly inappropriate interest in, or use of, toys and other objects. Play or imaginative themes tend to be repetitive or appear to reflect things seen in media or on TV. Some use of toy people as agents, for example, has an action figure or fake puppet play and may use objects or use objects to represent something else. Response to others' attempts to engage him or her in pretend play, but limited spontaneous initiation of imaginative play. Interests may be in a specific activity or toy or object.

3. Moderately inappropriate interest in, or use of, toys and other objects. Limited imaginative creative play, either spontaneously or in response to others. People typically not used as agents, and limited use of objects to represent something else. No original themes in play. May show some repetitive, inappropriate use of objects or interest in parts of objects, interest in play or objects that are not unusual or unusual.

4. Severely inappropriate interest in, or use of, toys and other objects. No creative play. Toys or objects used to represent or inappropriate manner.

6. Adaptation to Change/Restricted Interests

This item includes difficulty coping with change, ritualistic behaviors, and restricted special interests. The rating is based on the most severe level of difficulty in any one specific area.

1. Age-appropriate response to changes in routine. May notice or comment on changes in routines, but accepts these changes without undue stress. Shows a wide variety of interests, with no single interest or theme predominating.

2. Mildly abnormal adaptation to change/variety of interests. Unusually quick to develop new routines or, when others try to change tasks, the person may continue the same activity or one for more or less interest and may be directed to change if needed. May show preference for specific activities or toys or topics of conversation, though can be directed to other topics.

3. Moderately abnormal adaptation to change/variety of interests. Has definite special interests or preferences for specific activities, toys, objects, or topics. Adult needs to actively want to engage him or her in other topics or activities. Shows displeasure and may resist change or refusal to engage in routine. May become distressed by attempts to interrupt or change topic or activity. May have rigid rules or routines that have to be done in a particular way. May report subjective feelings of distress about change or interruptions, or may become overly fixed on schedule, checklist, or linking of events.

4. Severely abnormal adaptation to change/variety of interests. Has definite special interests or preferences, or has severe reaction to change. Resists with extreme anxiety, anger, or resistance to attempts to change activity or topic or routine.
7. Visual Response

This item covers use of vision in three areas: visual fascination, the ease with which the person can shift visual attention, and the degree to which the person’s eye contact is integrated with actions and communication.

1. Age-appropriate visual response. Visual behavior is normal and appropriate for his or her age. Eye contact is good and integrated with verbal and nonverbal communication skills. Easily shifts visual attention.

2. Mildly abnormal visual response. May stare inappropriately at objects. Eye contact is not consistently integrated with verbalizations. Included at this level is any inconsistency in eye contact, regardless of the presence of eyes he or she makes eye contact. May show more interest than is typical in objects in room or in looking at specific objects, such as moving parts, lights, or mirrors.

3. Moderately abnormal visual response. Eye contact is not integrated with verbalizations. Obvious visual fascination with objects, lights, mirrors, spinning toys, and toys. May use peripheral vision to look at things. Obvious difficulty shifting visual attention from high-interest items.

4. Severely abnormal visual response. Persistent avoidance of eye contact. Excessive interest in looking at specific objects or in looking at objects in a passive way.

8. Listening Response

This rating is based on the person’s response to sounds and how the listening response is coordinated with the use of other senses. The person’s response to his or her own name is scored on this item. Emphasis is placed on unusual over- or under-interest in sounds, as opposed to distractibility. Older individuals should be asked directly about this item.

1. Age-appropriate listening response. Listening behavior is normal and appropriate for his or her age. Listening is used together with other senses; for example, child looks toward person who is speaking, person responds to name.

2. Mildly abnormal listening response. Some difficulty responding to verbal stimuli when background noise present. Responds in course after repeated attempts to get attention. There may be some lack of response or mild overreaction to certain sounds. Typical listening responses are apparent either in direct observation or by report from outside sources, but not both.

3. Moderately abnormal listening response. Responses to sounds or verbal stimuli are inconsistent. May show marked reaction to some sounds or complete disregard for others. Sudden reactions in some cases as means of getting attention. Unusual responses are obvious across settings, based on some combination at direct report of persons, witness report, and direct observation.

4. Severely abnormal listening response. Overreacts or underreacts to sounds to an extremely varied degree. Is noticeably less responsive to verbalization than to noises made by objects. Does not respond to repeated attempts to get his or her attention by calling his or her name. Unusual responses are evident across settings.
9. Taste, Smell, and Touch Response and Use

This item addresses the person’s response to stimulation of the near receptors of taste, smell, touch, and pain. Subtle aspects of the stimulation of these senses include responses to the texture of clothing or food such that the person wears a limited variety of clothes or eats a limited variety of foods. Self-report of issues in this area should be considered, especially for adults.

1. Normal use of, and response to, taste, smell, and touch. Explores new objects in age-appropriate way, generally by looking and feeling. Responds appropriately to pain or touch from others. Reacts to taste pokes or licks by showing appropriate discomfort, but does not overreact. Wears a variety of textures of clothing and eats a wide variety of foods.

2. Mildly abnormal use of, and response to, taste, smell, and touch. May occasionally explore things by using attempts to smell or taste the object, or rub the object against part of his or her face or body. May show mild over- or underreaction to touch or pain. May have obvious eating or food preferences, but is easily encouraged to try new things. Unusual sensory responses are apparent in direct observation or by report from adults who know the child, but not both.

3. Moderately abnormal use of, and response to, taste, smell, and touch. Obstructively explores objects by chewing or touching them, or rubbing the object against part of his or her face or body. Over- or underreacts to touch or pain to a moderate degree. The person has limited clothing he or she will wear or food he or she will eat. Unusual eating patterns such as chewing and/or food preferences are obvious across settings, and the person reports these difficulties. They are obvious on direct observation. Sensory responses are difficult to modify and create stress or require adaptation in the everyday environment.

4. Severely abnormal use of, and response to, taste, smell, and touch. Has extreme limits on the clothes he or she will wear or the foods he or she will eat. Unusual eating patterns such as chewing and/or food preferences are obvious across settings, and the person shows persistent preoccupation with chewing, touching, or tasting things. Unusual sensory responses are obvious across settings. Unusual responses are evident across settings.

10. Fear or Anxiety

This item focuses on the degree to which the person has unusual fear or anxiety compared to what is appropriate for the situation or context.

1. Normal fear or anxiety. Behavior is appropriate both to the situation and for his or her age.

2. Mildly abnormal fear or anxiety. Occasionally shows too much or too little fear or anxiety compared to the reaction of a typical person of the same age in a similar situation. The abnormal response is only evident in one setting—for example, either on direct observation or on report from a witness in another setting, but not both.

3. Moderately abnormal fear or anxiety. Shows either quite a bit more or quite a bit less fear or anxiety than is typical even for a younger person in a similar situation. The abnormal response is apparent across more than one setting, and the person either reports these difficulties or they are obvious on direct observation.

4. Severely abnormal fear or anxiety. Fear or anxiety is pervasive across all settings and persists even after repeated explanations or experiences with harmless events or objects. It is extremely difficult to calm or comfort the person. Conversely, may show persistent and pervasive disregard for hazards that others of same age avoid.
11. Verbal Communication

This is a rating of two facets of the person's speech and language skills, and is best evaluated by a direct interaction with the person. This item includes verbal oddities—such as formal language, unusual tone or inflection, and repetitive or made-up phrases—and the ability to carry on a reciprocal conversation.

| 1 | Normal verbal communication, age and situation appropriate. Able to carry on an age-appropriate conversation with another person, he or she is able to respond to others' utterances while also adding additional information in at least a four-element sentence. No evidence of unusual speech inflection, volume, or tone. No evidence of made-up words or repetitive or made-up phrases. |
| 1.5 | Mildly abnormal verbal communication. Conversation exchanges are more limited than expected for this age. Occasional use of made-up words or repetitive, stereotyped speech. At times may display unusual vocal intonation or rate of speech. Ratings at this level indicate that the person has problems with conversation or verbal ability, but not both. |
| 2 | Moderately abnormal verbal communication. Normal initiation of conversation during direct interaction. Verbalizations include overly formal language or repetitive phrases. Little reciprocal conversation; may talk on own topic but little sense of interaction. Vocal intonation or rate of speech when unusual. Some use of unusual words or narrative speech. Some apparent difficulties in carrying on a reciprocal conversation and displays some type of verbal oddity. |
| 3 | Severely abnormal verbal communication. Inability to have a conversation with another person. May respond to specific questions, but does not engage in a two-way and non-interactive conversation. May talk on own topic but little sense of interaction. Vocal intonation or rate of speech when unusual. Complete use of unusual words or narrative speech. Significant difficulties in both areas of expressive communication—reciprocal conversation and verbal oddity. |

12. Nonverbal Communication

This item rates all forms of nonverbal communication. While both use of and response to nonverbal cues are considered, greater emphasis is placed on their use. Attention is given to the use of gaze to regulate and understand interactions and the use of facial expressions and gestures in combination with verbalizations for a variety of communication functions—Instrumental, descriptive, and emphatic.

| 1 | Normal use of nonverbal communication, age and situation appropriate. Uses a variety of facial expressions and instrumental, descriptive, and emphatic gestures that are well integrated with verbalizations. Responses to facial expressions and gestures from others. Gaze is used to regulate interactions with others. |
| 1.5 | Mildly abnormal use of nonverbal communication. Uses instrumental gestures such as pointing or reaching to indicate needs. Descriptive gestures are used infrequently and are not well integrated with verbalizations. Responds to very obvious facial expressions or gestures from others. May show too little or exaggerated facial expressions at times, though generally shows appropriate expression. Inconsistent use of gaze to regulate interaction with others. |
| 2 | Moderately abnormal use of nonverbal communication. Facial expressions are often flat or exaggerated. Uses limited instrumental gestures, and these gestures are not well integrated with verbalizations. Rarely uses descriptive or emphatic gestures. Shows limited response to nonverbal communication from others. Joint attention is rare, the person seldom uses or responds to gaze or gestures as a means of sharing attention to an object or activity. |
| 3 | Severely abnormal use of nonverbal communication. Facial expressions are either flat or exaggerated. Does not use instrumental, descriptive, or emphatic gestures and shows no awareness of nonverbal communication from others. No evidence of using gaze to regulate activities with others. |
13. Thinking/Cognitive Integration Skills

This is a rating of the person's ability to understand the meaning of larger concepts and the ability of the person to integrate relevant details into a meaningful overview (central coherence). Part of this process involves the person's ability to discriminate between relevant and irrelevant details.

1. Age-appropriate thinking/cognitive integration skills. Able to understand meaning of information presented either pictorially, in writing, or verbally. He or she demonstrates central coherence, that is, the ability to attend to relevant versus irrelevant details and integrate this information into a meaningful overview.

2. Mildly impaired in specific thinking/cognitive integration skills. Delays in thinking compared to persons of same age. Difficulties may be seen in distinguishing relevant from irrelevant cues for conceptualizing or person can verbalize an overall understanding but is unable to articulate how meaning was derived. At times supportive presence of another person helps with comprehension.

3. Moderately impaired in specific thinking/cognitive integration skills. Notable difficulties comprehending meaning and integrating information into overall conceptualization, but shows great attention to specific things and concrete details. Frequently requires specific prompts from others to attend to relevant details or grasp the larger conceptual concepts.

4. Severe delay in specific thinking/cognitive integration skills. Shows repeated and consistent difficulty distinguishing relevant from irrelevant details. Even with persistent efforts from another, may not be able to conceptualize the overall meaning of information.

14. Level and Consistency of Intellectual Response

THE RATER MUST read the complete description in the Manual before rating this item. This rating is concerned with the discrepancies in and consistency of the individual's skills across different areas, as well as the person's general level of intellectual functioning. By definition, this instrument is appropriate only for individuals whose overall IQ score is above 80, so the descriptors make this assumption. Unless the individual has a savant skill, which always receives a rating of 4, all individuals whose adaptive skills are appropriate for their age and intellectual abilities should receive a rating of 1, regardless of intellectual level or variability in skills.

1. Intelligence is at least normal and reasonably consistent across various areas. Has at least average intellectual abilities and does not have any unusual intellectual skills or problems. IQ score is average or higher (90) with limited discrepancies. Adaptive skills are appropriate for age and intellectual abilities.

2. Mildly abnormal intellectual functioning. Not as smart as typical person of same age; skills appear evenly delayed across all areas. IQ score is in the low-average range (80 to 90) with limited discrepancies. Adaptive skills are less than expected for cognitive level.

3. Moderately abnormal intellectual functioning. Not quite as smart as typical person of same age; skills appear evenly delayed across all areas. IQ score is in the low-average range (80 to 90) with limited discrepancies. Adaptive skills are less than expected for level of intelligence.

4. Severely abnormal intellectual functioning. Not a skill that is significantly better than an expected for his or her level of intelligence and better than that exhibited by typical peers (savan skill). At least low-average intelligence (80). Adaptive skills are typically less than expected for level of intelligence, but in rare instances may be appropriate for cognitive level.
15. General Impressions

This is intended to be an overall rating of autism based on your subjective impression of the degree to which the person has autism as defined by the other 14 items. This rating should be made without recourse to averaging the other ratings. As with the other items, this rating should be made by taking into account all available data from such sources as the case history, test results, parent and other interviews, or past records.

1. No autism spectrum disorder. Shows none of the symptoms characteristic of an autism spectrum disorder.

2. Mild autism spectrum disorder. Shows only a few symptoms or only a mild degree of an autism spectrum disorder—mild interference with daily functioning.

3. Moderate autism spectrum disorder. Shows a number of symptoms or a moderate degree of an autism spectrum disorder—moderate interference with daily functioning.

4. Severe autism spectrum disorder. Shows many symptoms or an extreme degree of an autism spectrum disorder—extreme interference with daily functioning.
Appendix 17: Children's Communication Checklist (CCC-2)

The Children's Communication Checklist
Second Edition

By D.V.M. Bishop

INSTRUCTIONS

The CCC-2 was developed to help us understand more about communication strengths and difficulties in children. Although we can get an idea of how a child communicates by using language tests, it is helpful to also find out how the child behaves in an everyday setting. You can help us do this by completing the items on the next three pages.

This checklist contains a series of statements describing how children communicate. For each statement, you are asked to give information about the child whose name (or code number) appears below. You are asked to judge whether you have observed that behaviour:

0. less than once a week (or never)
1. at least once a week, but not every day
2. once or twice a day
3. several times (more than twice) a day (or always)

Please write the number in the box for each item, choosing the response that, in your judgement, best describes the child. If you find it hard to make up your mind, think over the last week, and try to remember how often you have observed the child behaving this way.

Please read each item carefully. Do not leave any items blank. If you are really unable to make a judgement, please put an X against that item, and add a comment if you wish.

Name or code number of child: ___________________________ Gender: ______

Date of birth: _______________ Today's date: _______________ Age: ______

Your name (person completing the checklist): ___________________________

Your relation to the child [i.e. parent, speech therapist, etc.]: _______________

(For respondents other than parents) How long have you known this child? _______________

Has s/he ever had a permanent hearing loss diagnosed? YES □ NO □

If YES, please give further details below.

Has s/he any permanent physical handicap or chronic illness? YES □ NO □

If YES, please give further details below.

Is English the main language spoken at home? YES □ NO □

If NO, please give further details below.

Is s/he able to string words together in sentences? YES □ NO □

The CCC-2 is intended to be used with children who can talk in simple sentences, so if you have ticked NO, please do not complete any further questions:

Additional details:

264
| 1 Gets mixed up between he and she so might say “he” when talking about a girl, or “she” when talking about a boy |
| 2 Simplifies words by leaving out some sounds, e.g. “crocodile” pronounced as “cockadile”, or “stranger” as “staynger” |
| 3 Appears anxious in the company of other children |
| 4 Makes false starts, and appears to grope for the right words; e.g., might say “can I – can I – can I have an – have an ice-cream.” |
| 5 Talks repetitively about things that no-one is interested in |
| 6 Forgets words s/he knows – e.g. instead of “rhinoceros” may say “you know, the animal with the horn on its nose…” |
| 7 With familiar adults, seems inattentive, distant or preoccupied |
| 8 Looks blank in a situation where most children would show a clear facial expression – e.g. when angry, fearful or happy |
| 9 When given the opportunity to do what s/he likes, chooses the same favourite activity (e.g. playing a specific computer game) |
| 10 Uses terms like “he” or “it” without making it clear what s/he is talking about. For instance, when talking about a film, might say “he was really great” without explaining who “he” is |
| 11 Says things that s/he does not seem to fully understand (may appear to be repeating something s/he’s heard an adult say). So, for instance, a 5-year-old may be heard to say of a teacher “s/he’s got a very good reputation” |
| 12 Mixes up words of similar meaning, e.g., might say “dog” for “fox”, or “screwdriver” for “hammer” |
| 13 Is babied, teased, or bullied by other children |
| 14 Does not look at the person s/he is talking to |
| 15 Mixes the point of jokes and puns (though may be amused by nonverbal humour such as slapstick) |
| 16 Is left out of joint activities by other children |
| 17 Gets mixed up between he/him or she/her, so might say “him is working” rather than “he is working”, or “her have a cake” rather than “she has a cake” |
| 18 Uses favourite phrases, sentences or longer sequences in rather inappropriate contexts. E.g., might say “all of a sudden” rather than “then”, as in “we went to the park and all of a sudden we had a picnic”. Or might habitually start utterances with “by the way” |
| 19 Gets confused when a word is used with a different meaning from usual; e.g. might fail to understand if an unfriendly person was described as “cold” (and would assume they were shivering!) |
| 20 Stands too close to other people when talking to them |
| 21 Talks to people too readily: e.g. without any encouragement, starts up a conversation with a stranger |
| 22 Talks about lists of things s/he has memorised e.g., the names of the capitals of the world, or the names of varieties of dinosaurs |
| 23 Pronounces words in an over-precise manner: accent may sound affected or “put-on”, as if child is mimicking a TV personality rather than talking like those around him/her |
| 24 Pronounces words in a babyish way, such as “chimbley” for “chimney” or “bokkle” for “bottle” |
| 25 Can be hard to tell if s/he is talking about something real or make-believe |
| 26 Moves the conversation to a favourite topic, even if others don’t seem interested in it |
### Children's Communication Checklist (CCC-2)

by D.V.M. Bishop

Please enter a number in the box in the right-hand column, as follows:
- 0 = less than once a week (or never)
- 1 = at least once a week, but not every day
- 2 = once or twice a day
- 3 = several times (more than twice) a day (or always)

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>27</td>
<td>Produces utterances that sound babyish because they are just 2 or 3 words long, such as &quot;me got ball&quot; instead of &quot;I've got a ball&quot; or &quot;give dolly&quot; instead of &quot;give me the dolly&quot;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>Ability to communicate varies from situation to situation — e.g., may cope well when talking one-to-one with a familiar adult, but have difficulty expressing himself/herself in a group of children</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>Leaves off beginnings or ends of words, e.g., says &quot;toe&quot; instead of &quot;road&quot; or &quot;Nama&quot; instead of &quot;banana&quot;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>Repeats back what others have just said. For instance, if you ask, &quot;what did you eat?&quot; might say, &quot;what did I eat?&quot;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>Ignores conversational overtures from others (e.g., if asked, &quot;What are you making?&quot; does not look up and just continues working)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>Mixes up words that sound similar. e.g., might say &quot;telephone&quot; for &quot;television&quot; or &quot;magician&quot; for &quot;musician&quot;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>Hurts or upsets other children without meaning to</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>34</td>
<td>Takes in just 1-2 words in a sentence, and so misinterprets what has been said. E.g., if someone says &quot;I want to go skating next week&quot;, s/he may think they've been skating, or want to go now</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>It's difficult to stop him/her from talking</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>36</td>
<td>Leaves off past tense — ed endings on words, so might say &quot;John kicked the ball&quot; instead of &quot;John kicked the ball&quot;, or &quot;Sally play over there&quot; instead of &quot;Sally played over there&quot;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>37</td>
<td>Tells people things they know already</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>38</td>
<td>Makes mistakes in pronouncing long words; e.g., says &quot;vegetable&quot; rather than &quot;vegetable&quot; or &quot;treadle&quot; rather than &quot;telescope&quot;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>39</td>
<td>Fails to recognise when other people are upset or angry</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>Gets the sequence of events muddled up when trying to tell a story or describe a recent event. E.g., if describing a film, might talk about the end before the beginning</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>41</td>
<td>Is over-literal, sometimes with (unintentionally) humorous results. E.g., a child who was asked &quot;Do you find it hard to get up in the morning?&quot; replied &quot;No. You just put one leg out of the bed and then the other and stand up.&quot; Another child who was told &quot;watch your hands&quot; when using scissors, proceeded to stare at his fingers.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>42</td>
<td>Includes over-precise information (e.g., exact date or time) in his/her talk, e.g., when asked &quot;when did you go on holiday?&quot; may say &quot;1st July 1995&quot; rather than &quot;in the summer&quot;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>43</td>
<td>Leaves out &quot;is&quot;, and so says &quot;Daddy going to work&quot; rather than &quot;Daddy's going to work&quot; or &quot;Daddy is going to work&quot;. Or might say &quot;The boy big&quot; rather than &quot;The boy is big&quot;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>44</td>
<td>Mispronounces &quot;th&quot; for &quot;s&quot; or &quot;v&quot; for &quot;r&quot;. E.g., says &quot;thwap&quot; instead of &quot;soap&quot; or &quot;rahbit&quot; instead of &quot;rabbit&quot;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>45</td>
<td>Asks a question, even though s/he has been given the answer</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>46</td>
<td>Is vague in choice of words, making it unclear what s/he is talking about. E.g., saying &quot;that thing&quot; rather than &quot;kettle&quot;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>47</td>
<td>Shows interest in things or activities that most people would find unusual, such as traffic lights, washing machines, lamp-posts</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>48</td>
<td>Doesn't explain what s/he is talking about to someone who doesn't share his/her experiences; for instance, might talk about &quot;Johnny&quot; without explaining who he is</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>49</td>
<td>Surprises people by his/her knowledge of unusual words — uses terms you'd expect to hear from an adult rather than child</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>It is hard to make sense of what s/he is saying (even though the words are clearly spoken)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Children’s Communication Checklist (CCC-2) by D.V.M. Bishop

**Please enter a number in the box in the right hand column, as follows:**
0 = less than once a week (or never); 1 = at least once a week, but not every day
2 = once or twice a day; 3 = several times (more than twice) a day (or always)

The questions so far have asked about difficulties children may have that affect communication. The remaining questions ask about communicative strengths.

Please respond 0 to 3, as before, but remember that now a 0 response would mean that a child lacks this strength, and a 3 would indicate good communicative skill.

<table>
<thead>
<tr>
<th>Question</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>51 Speaks clearly so that the words can easily be understood by someone who doesn't know him/her very well</td>
<td></td>
</tr>
<tr>
<td>52 Reacts positively when a new and unfamiliar activity is suggested</td>
<td></td>
</tr>
<tr>
<td>53 Talks clearly about what s/he plans to do in the future (e.g. what s/he will do tomorrow, or plans for going on holiday)</td>
<td></td>
</tr>
<tr>
<td>54 Appreciates the humour expressed by irony. Would be amused rather than confused if someone said “isn’t it a lovely day!” when it is pouring with rain</td>
<td></td>
</tr>
<tr>
<td>55 Produces long and complicated sentences such as: “When we went to the park I had a go on the swings”; “I saw this man standing on the corner”</td>
<td></td>
</tr>
<tr>
<td>56 Makes good use of gestures to get his/her meaning across</td>
<td></td>
</tr>
<tr>
<td>57 Shows concern when other people are upset</td>
<td></td>
</tr>
<tr>
<td>58 Speaks fluently and clearly, producing all speech sounds accurately and without any hesitation</td>
<td></td>
</tr>
<tr>
<td>59 Keeps quiet in situations where someone else is trying to talk or concentrate (e.g. when someone else is watching TV, or during formal occasions such as school assembly or a religious ceremony)</td>
<td></td>
</tr>
<tr>
<td>60 Realises the need to be polite – would pretend to be pleased if given a present if she did not really like, and would avoid making personal comments about strangers</td>
<td></td>
</tr>
<tr>
<td>61 When answering a question, provides enough information without being over-precise</td>
<td></td>
</tr>
<tr>
<td>62 You can have an enjoyable, interesting conversation with him/her</td>
<td></td>
</tr>
<tr>
<td>63 Shows flexibility in adapting to unexpected situations: e.g. does not get upset if s/he planned to play on the computer, but has to do something else because it isn’t working</td>
<td></td>
</tr>
<tr>
<td>64 Uses abstract words that refer to general concepts rather than something you can see – e.g. “knowledge”, “politics”, “courage”</td>
<td></td>
</tr>
<tr>
<td>65 Smiles appropriately when talking to people</td>
<td></td>
</tr>
<tr>
<td>66 Uses words that refer to whole classes of objects, rather than a specific item. E.g. refers to a table, chair and drawers as “furniture”, or to apples, bananas and peas as “fruit”</td>
<td></td>
</tr>
<tr>
<td>67 Talks about his/her friends; shows interest in what they do and say</td>
<td></td>
</tr>
<tr>
<td>68 Explains a past event (e.g. what s/he did at school, or what happened at a football game) clearly</td>
<td></td>
</tr>
<tr>
<td>69 Produces sentences containing “because” such as “John had a cake because it was his birthday”</td>
<td></td>
</tr>
<tr>
<td>70 Talks to others about their interests, rather than his/her own</td>
<td></td>
</tr>
</tbody>
</table>
# Appendix 18: Children's Communication Checklist (CCC-2) – Score sheet

## CCC-2 Summary Sheet

<table>
<thead>
<tr>
<th>Scale</th>
<th>W Items (51-70)</th>
<th>Pos: $=\frac{1}{6}\cdot W$</th>
<th>X Items (1-26)</th>
<th>Y Items (27-50)</th>
<th>Neg: $=X+Y$</th>
<th>M Items Missing</th>
<th>Sum: Neg + Pos</th>
<th>Scaled Score</th>
<th>Percentile</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Speech</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B. Syntax</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. Semantics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D. Coherence</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E. Inappropriate Initiation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F. Stereotyped Language</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G. Use of Context</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H. Nonverbal Communication</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I. Social Relations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>J. Interests</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Sum</strong></td>
<td><strong>Positive Sum</strong></td>
<td></td>
<td><strong>Negative Sum</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Validity check:** Are positive sum and negative sum consistent?  
(See Appendix 3B) □ YES □ NO □  
1. See manual for instructions if you have missing values on the 'positive' items for any scale.

---

**CCC-2**

The Psychological Corporation®  
Meeting Your Assessment Needs

Published by The Psychological Corporation Limited, 32 Jamestown Road, London NW1 7BY. Copyright © The Psychological Corporation 2003.  
All rights reserved. No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage or retrieval system, without permission in writing from the publisher.  
Printed in the United Kingdom ISBN 0 7494 2062 3
Appendix 19: Listening Inventory for Education – Student version (LIFE-7, adapted for NZ)

LIFE-7 (NZ) (Purdy et al., 2011)

Instructions: What do you think it is like to be in the picture below? You have to look at the picture carefully and decide how easy it is to hear the teacher. Draw a cross through the box to show your answer.

For example: If you think it is mostly easy to hear the words the teacher is saying mark the box like this:

<table>
<thead>
<tr>
<th>always easy</th>
<th>mostly easy</th>
<th>sometimes difficult</th>
<th>mostly difficult</th>
<th>always difficult</th>
</tr>
</thead>
<tbody>
<tr>
<td>😊</td>
<td>☹️</td>
<td>😞</td>
<td>😞</td>
<td>😞</td>
</tr>
</tbody>
</table>

LIFE-7 (NZ) Questionnaire

1. It is a quiet day and there is no noise from outside the classroom.

How well can you hear the teacher’s words?

<table>
<thead>
<tr>
<th>always easy</th>
<th>mostly easy</th>
<th>sometimes difficult</th>
<th>mostly difficult</th>
<th>always difficult</th>
</tr>
</thead>
<tbody>
<tr>
<td>😊</td>
<td>😊</td>
<td>😞</td>
<td>😞</td>
<td>😞</td>
</tr>
</tbody>
</table>
2. The teacher is talking but there are children making a noise outside your classroom.

How well can you hear the teacher's words?

<table>
<thead>
<tr>
<th>always easy</th>
<th>mostly easy</th>
<th>sometimes difficult</th>
<th>mostly difficult</th>
<th>always difficult</th>
</tr>
</thead>
<tbody>
<tr>
<td>☑️</td>
<td>☑️</td>
<td>☐</td>
<td>☒□</td>
<td>☒□</td>
</tr>
</tbody>
</table>

3. The teacher has asked a question to the whole class. Someone is giving an answer.

How well can you hear the answer?

<table>
<thead>
<tr>
<th>always easy</th>
<th>mostly easy</th>
<th>sometimes difficult</th>
<th>mostly difficult</th>
<th>always difficult</th>
</tr>
</thead>
<tbody>
<tr>
<td>☑️</td>
<td>☑️</td>
<td>☐</td>
<td>☒□</td>
<td>☒□</td>
</tr>
</tbody>
</table>

4. The teacher is talking and moving around the room.

How well can you hear the teacher's words?

<table>
<thead>
<tr>
<th>always easy</th>
<th>mostly easy</th>
<th>sometimes difficult</th>
<th>mostly difficult</th>
<th>always difficult</th>
</tr>
</thead>
<tbody>
<tr>
<td>☑️</td>
<td>☑️</td>
<td>☐</td>
<td>☒□</td>
<td>☒□</td>
</tr>
</tbody>
</table>
5  The teacher is giving a test to the class.

<table>
<thead>
<tr>
<th>always</th>
<th>mostly</th>
<th>sometimes</th>
<th>mostly</th>
<th>always</th>
</tr>
</thead>
<tbody>
<tr>
<td>easy</td>
<td>easy</td>
<td>difficult</td>
<td>difficult</td>
<td>difficult</td>
</tr>
<tr>
<td>😊</td>
<td>😊</td>
<td>😞</td>
<td>😞</td>
<td>😞</td>
</tr>
</tbody>
</table>

How well can you hear the teacher's words?

6  There are two teachers in the class. They are both talking. One of the teachers is talking to you from the front of the class. You need to listen to this teacher.

<table>
<thead>
<tr>
<th>always</th>
<th>mostly</th>
<th>sometimes</th>
<th>mostly</th>
<th>always</th>
</tr>
</thead>
<tbody>
<tr>
<td>easy</td>
<td>easy</td>
<td>difficult</td>
<td>difficult</td>
<td>difficult</td>
</tr>
<tr>
<td>😊</td>
<td>😊</td>
<td>😞</td>
<td>😞</td>
<td>😞</td>
</tr>
</tbody>
</table>

How well can you hear the teacher's words?

7  You are in assembly.

<table>
<thead>
<tr>
<th>always</th>
<th>mostly</th>
<th>sometimes</th>
<th>mostly</th>
<th>always</th>
</tr>
</thead>
<tbody>
<tr>
<td>easy</td>
<td>easy</td>
<td>difficult</td>
<td>difficult</td>
<td>difficult</td>
</tr>
<tr>
<td>😊</td>
<td>😊</td>
<td>😞</td>
<td>😞</td>
<td>😞</td>
</tr>
</tbody>
</table>

How well can you hear the teacher's words?
Background
The original LIFE (Anderson & Smaldino, 1999) was adapted and modified by researchers in the UK to reflect the educational context there (Canning, 1999). The questionnaire was adapted to focus on whole-class listening situations only. The LIFE-UK has been used in various research trials in the UK including classroom soundfield amplification, paediatric hearing aid use and children with hearing impairment. In 2007 MSLTPrac student Amanda Morgan collected normative data for Auckland children (n=83, 42% boys) aged 7.4-12.9 years (mean 9.91, SD 1.33 years) with no listening concerns (see Table 1). These children came from low-mid decile (72%) and high decile (28%) schools. Seven items with high factor loadings are included in the LIFE-NZ.

References
**Scoring of questionnaire**
The LIFE questionnaire has a 5-point Likert scale using smiling and frowning faces to rate ease of listening; ranging from Always Easy to Always Difficult. Numbers from 1 to 5 are assigned to responses, with higher ratings indicating more difficulty with listening [Always easy=1; Mostly easy=2; Sometimes difficult=3; Mostly difficult=4; Always difficult=5]. The table below can be used to calculate scores for the three factors and overall. Numbers exceeding the norm indicate greater self-reported listening difficulty.

<table>
<thead>
<tr>
<th>NAME:</th>
<th>DATE:</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Question</th>
<th>Content</th>
<th>Factor 1</th>
<th>Factor 2</th>
<th>Factor 3</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Listening in Noise</strong></td>
<td><strong>Listening in Quiet</strong></td>
<td><strong>Focussed Listening</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Teacher talking, noise outside the classroom</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Teacher is talking and moving around the room</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Two teachers talking, trying to listen to one of them</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Mean = (Sum ÷ 3)</strong></td>
<td><strong>Mean = (Sum ÷ 2)</strong></td>
<td></td>
<td><strong>Mean = (Sum ÷ 7)</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>NORM 2.14±.83</strong></td>
<td><strong>NORM 1.23±.46</strong></td>
<td></td>
<td><strong>NORM 1.60±.59</strong></td>
<td><strong>NORM 1.73±.51</strong></td>
</tr>
<tr>
<td>1</td>
<td>Quiet day, no noise from outside the classroom</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Teacher is giving a test to the whole class</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Teacher asks question to whole class. One person answering.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>In assembly. Hear teacher’s words?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix 20: Listening Inventory for Education – Teacher version (LIFE-UK)

**L.I.F.E. UK**
Listening Inventory for Education
Teacher Appraisal of Listening

<table>
<thead>
<tr>
<th>Name</th>
<th>Date</th>
<th>School</th>
<th>Teacher</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Classroom soundfield amplification trial: yes/no
Personal f.m. amplification trial: yes/no
Details of Amplification system used:

**Instructions:** Read the statement below and then circle the score that best describes the behaviour of the child (or children). You should make a judgement as to whether or not the following behaviours are a cause for concern or not.

<table>
<thead>
<tr>
<th>In your opinion</th>
<th>Very good</th>
<th>Satisfactory</th>
<th>Very Poor</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Following class directions</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Following individual directions</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Overall attention span</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. On task behaviour</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Rate of learning (speed of following instruction)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Involvement in class discussions (volunteers more, makes appropriate contributions)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Contributes when working in a group</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Paying attention to multimedia (e.g. video, OHP)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Willingness to answer questions</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Answering questions in an appropriate and relevant manner</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. Amount of repair behaviour (this refers to asking questions, to teacher or peer, in order to clarify what is required)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12* Overall noise levels in the class while working in groups</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13* Noise levels in the class during whole class teaching</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Additional Comments:

---

1 Based on LIFE by Kari Anderson and Joseph Smaldino 1997
* Only appropriate for classroom soundfield amplification systems

© d.canning@city.ac.uk 1999
## L.I.F.E. UK

### Listening Inventory for Education

**Teacher Appraisal of Impact on Listening**

**Name**

**School**

**Date**

**Teacher**

- Classroom soundfield amplification trial: yes/no
- Personal f.m. amplification trial: yes/no
- Length of trial: 10 Weeks
- Length of trial:
- Details of Amplification system used:

**Instructions:** Read the statement below and then circle the score that best describes the behaviour following the use of the amplification system. The scores are ranked from ‘improvement’ on the left to most deterioration on the right. When completed, add up your score and write it in the space at the bottom of the page. You should refer to your original ratings completed prior to the trial.

<table>
<thead>
<tr>
<th>Has there been a change in:</th>
<th>Improvement</th>
<th>No change</th>
<th>Deterioration</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Following class directions</td>
<td>(2)</td>
<td>(1)</td>
</tr>
<tr>
<td>2</td>
<td>Following individual directions</td>
<td>(2)</td>
<td>(1)</td>
</tr>
<tr>
<td>3</td>
<td>Overall attention span</td>
<td>(2)</td>
<td>(1)</td>
</tr>
<tr>
<td>4</td>
<td>On task behaviour</td>
<td>(2)</td>
<td>(1)</td>
</tr>
<tr>
<td>5</td>
<td>Rate of learning (quicker to comprehend instructions)</td>
<td>(2)</td>
<td>(1)</td>
</tr>
<tr>
<td>6</td>
<td>Involvement in class discussions (volunteers more, makes appropriate contributions)</td>
<td>(2)</td>
<td>(1)</td>
</tr>
<tr>
<td>7</td>
<td>Contributes when working in a group</td>
<td>(2)</td>
<td>(1)</td>
</tr>
<tr>
<td>8</td>
<td>Paying attention to multimedia (e.g. video, OHP)</td>
<td>(2)</td>
<td>(1)</td>
</tr>
<tr>
<td>9</td>
<td>Willingness to answer questions</td>
<td>(2)</td>
<td>(1)</td>
</tr>
<tr>
<td>10</td>
<td>Answering questions in an appropriate and relevant manner</td>
<td>(2)</td>
<td>(1)</td>
</tr>
<tr>
<td>11</td>
<td>Frequency of repair behaviour (this refers to the frequency of asking questions, to teacher or peer, in order to clarify what is required)</td>
<td>(2)</td>
<td>(1)</td>
</tr>
<tr>
<td>12*</td>
<td>Noise levels in the class during whole class teaching</td>
<td>(2)</td>
<td>(1)</td>
</tr>
<tr>
<td>13*</td>
<td>Overall noise levels in the class while working in groups</td>
<td>(2)</td>
<td>(1)</td>
</tr>
<tr>
<td></td>
<td>Based on my knowledge and observations I believe that the amplification system is beneficial to the student’s overall attention, listening and learning in the classroom</td>
<td>(5)</td>
<td>(2)</td>
</tr>
</tbody>
</table>

**Total Appraisal score**

- Highly successful: 20-27
- Successful: 10-19
- Minimally successful: 1-9
- Unsuccessful: -ve value

---

1 Based on LIFE by Karen Anderson and Joseph Smaildio 1997

* Only appropriate for classroom soundfield amplification systems

©David Canning 1999
Appendix 21: Listening Inventory for Education – Teacher version (LIFE-R)

### Listening Inventory For Education-Revised (L.I.F.E.-R.)

**Teacher Appraisal of Listening Difficulty**

By Karen L. Anderson, PhD, Joseph J. Smaldino, PhD, & Carrie Spangler, AuD

<table>
<thead>
<tr>
<th>Name</th>
<th>Grade</th>
<th>School</th>
<th>Teacher</th>
<th>Hearing Aid</th>
<th>CI User</th>
<th>Date LIFE Completed</th>
<th>Type of Classroom Hearing Technology</th>
</tr>
</thead>
</table>

#### L.I.F.E Classroom Listening Situations

**Instructions:** Based on your observations, please mark the response that best describes the student's level of challenge when listening and learning in each of the situations described below. If you have no idea how to answer an item, leave the item blank. Thank you for your assistance.

<table>
<thead>
<tr>
<th>No challenge or very rare</th>
<th>Occasionally challenged</th>
<th>Sometimes challenged</th>
<th>Often or regularly challenged</th>
<th>Almost always challenged</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

1. Student's ability to focus on follow large group verbal instruction (i.e., teacher front of room):

2. Student's ability to focus or follow verbal instruction when you are moving about the room:

3. Student's ability to focus or understand verbal responses by other students seated across the classroom from him/her:
   - Check one: ☐ With FM mic used by student ☐ Without FM mic

4. Ability to attend when listening to directions presented to the whole class (focus):

5. Ease of following directions provided to large group (hesitation before beginning work):

6. Ability to attend to class activities (distractibility, fidgety, typical level of attention):

7. Ability to stay on task (re: need for individual redirection):

8. Level of hesitation when volunteering to answer class questions in relation to peers:

9. Ability to answer questions appropriately (shows understanding of question and reasonable response):

10. Ability to understand information presented via instructional media (videos, computer, etc.):

11. Ability to focus on and understand morning announcements or large group assemblies:

12. Ability to attend to verbal instruction and understand when noise is present (i.e., transitions):

13. Ability to focus or understand peer comments during small group work:

14. Comfort during social involvement/informal peer conversations in comparison to peers:

15. Overall rate of listening/learning in comparison to class peers (rate of comprehension):

**Comments:** (absences, equipment use problems, etc.)

<table>
<thead>
<tr>
<th>CLASSROOM LISTENING SCORE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sum of items 1-15 (75 possible) Pretest _____ Post-test _____</td>
</tr>
</tbody>
</table>

No listening challenges or very rare | Occasional listening challenges | Sometimes experiences listening challenges | Often or regularly has listening challenges | Almost always has listening challenges |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>75</td>
<td>60</td>
<td>45</td>
<td>30</td>
<td>15</td>
</tr>
</tbody>
</table>

© 2011 by Karen L. Anderson, PhD, Joseph J. Smaldino, PhD, & Carrie Spangler, AuD
Refer to www.successforkidswithhearingloss.com for Instruction Manual
### Social Perception

**Affect Naming**

<table>
<thead>
<tr>
<th>Item</th>
<th>Affect</th>
<th>Emotion</th>
<th>Affect</th>
<th>Emotion</th>
<th>Affect</th>
<th>Emotion</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Happy</td>
<td>Sad</td>
<td>Angry</td>
<td>Afraid</td>
<td>Surprised</td>
<td>Disgusted</td>
<td>No Feeling/ Neutral</td>
</tr>
<tr>
<td>2.</td>
<td>Happy</td>
<td>Sad</td>
<td>Angry</td>
<td>Afraid</td>
<td>Surprised</td>
<td>Disgusted</td>
<td>No Feeling/ Neutral</td>
</tr>
<tr>
<td>3.</td>
<td>Happy</td>
<td>Sad</td>
<td>Angry</td>
<td>Afraid</td>
<td>Surprised</td>
<td>Disgusted</td>
<td>No Feeling/ Neutral</td>
</tr>
<tr>
<td>4.</td>
<td>Happy</td>
<td>Sad</td>
<td>Angry</td>
<td>Afraid</td>
<td>Surprised</td>
<td>Disgusted</td>
<td>No Feeling/ Neutral</td>
</tr>
<tr>
<td>5.</td>
<td>Happy</td>
<td>Sad</td>
<td>Angry</td>
<td>Afraid</td>
<td>Surprised</td>
<td>Disgusted</td>
<td>No Feeling/ Neutral</td>
</tr>
<tr>
<td>6.</td>
<td>Happy</td>
<td>Sad</td>
<td>Angry</td>
<td>Afraid</td>
<td>Surprised</td>
<td>Disgusted</td>
<td>No Feeling/ Neutral</td>
</tr>
<tr>
<td>7.</td>
<td>Happy</td>
<td>Sad</td>
<td>Angry</td>
<td>Afraid</td>
<td>Surprised</td>
<td>Disgusted</td>
<td>No Feeling/ Neutral</td>
</tr>
<tr>
<td>8.</td>
<td>Happy</td>
<td>Sad</td>
<td>Angry</td>
<td>Afraid</td>
<td>Surprised</td>
<td>Disgusted</td>
<td>No Feeling/ Neutral</td>
</tr>
<tr>
<td>9.</td>
<td>Happy</td>
<td>Sad</td>
<td>Angry</td>
<td>Afraid</td>
<td>Surprised</td>
<td>Disgusted</td>
<td>No Feeling/ Neutral</td>
</tr>
<tr>
<td>10.</td>
<td>Happy</td>
<td>Sad</td>
<td>Angry</td>
<td>Afraid</td>
<td>Surprised</td>
<td>Disgusted</td>
<td>No Feeling/ Neutral</td>
</tr>
<tr>
<td>11.</td>
<td>Happy</td>
<td>Sad</td>
<td>Angry</td>
<td>Afraid</td>
<td>Surprised</td>
<td>Disgusted</td>
<td>No Feeling/ Neutral</td>
</tr>
<tr>
<td>12.</td>
<td>Happy</td>
<td>Sad</td>
<td>Angry</td>
<td>Afraid</td>
<td>Surprised</td>
<td>Disgusted</td>
<td>No Feeling/ Neutral</td>
</tr>
<tr>
<td>13.</td>
<td>Happy</td>
<td>Sad</td>
<td>Angry</td>
<td>Afraid</td>
<td>Surprised</td>
<td>Disgusted</td>
<td>No Feeling/ Neutral</td>
</tr>
<tr>
<td>14.</td>
<td>Happy</td>
<td>Sad</td>
<td>Angry</td>
<td>Afraid</td>
<td>Surprised</td>
<td>Disgusted</td>
<td>No Feeling/ Neutral</td>
</tr>
<tr>
<td>15.</td>
<td>Happy</td>
<td>Sad</td>
<td>Angry</td>
<td>Afraid</td>
<td>Surprised</td>
<td>Disgusted</td>
<td>No Feeling/ Neutral</td>
</tr>
<tr>
<td>16.</td>
<td>Happy</td>
<td>Sad</td>
<td>Angry</td>
<td>Afraid</td>
<td>Surprised</td>
<td>Disgusted</td>
<td>No Feeling/ Neutral</td>
</tr>
<tr>
<td>17.</td>
<td>Happy</td>
<td>Sad</td>
<td>Angry</td>
<td>Afraid</td>
<td>Surprised</td>
<td>Disgusted</td>
<td>No Feeling/ Neutral</td>
</tr>
<tr>
<td>18.</td>
<td>Happy</td>
<td>Sad</td>
<td>Angry</td>
<td>Afraid</td>
<td>Surprised</td>
<td>Disgusted</td>
<td>No Feeling/ Neutral</td>
</tr>
<tr>
<td>19.</td>
<td>Happy</td>
<td>Sad</td>
<td>Angry</td>
<td>Afraid</td>
<td>Surprised</td>
<td>Disgusted</td>
<td>No Feeling/ Neutral</td>
</tr>
<tr>
<td>20.</td>
<td>Happy</td>
<td>Sad</td>
<td>Angry</td>
<td>Afraid</td>
<td>Surprised</td>
<td>Disgusted</td>
<td>No Feeling/ Neutral</td>
</tr>
<tr>
<td>21.</td>
<td>Happy</td>
<td>Sad</td>
<td>Angry</td>
<td>Afraid</td>
<td>Surprised</td>
<td>Disgusted</td>
<td>No Feeling/ Neutral</td>
</tr>
<tr>
<td>22.</td>
<td>Happy</td>
<td>Sad</td>
<td>Angry</td>
<td>Afraid</td>
<td>Surprised</td>
<td>Disgusted</td>
<td>No Feeling/ Neutral</td>
</tr>
<tr>
<td>23.</td>
<td>Happy</td>
<td>Sad</td>
<td>Angry</td>
<td>Afraid</td>
<td>Surprised</td>
<td>Disgusted</td>
<td>No Feeling/ Neutral</td>
</tr>
<tr>
<td>24.</td>
<td>Happy</td>
<td>Sad</td>
<td>Angry</td>
<td>Afraid</td>
<td>Surprised</td>
<td>Disgusted</td>
<td>No Feeling/ Neutral</td>
</tr>
</tbody>
</table>

**Prosody-Face Matching**

<table>
<thead>
<tr>
<th>Item</th>
<th>Task</th>
<th>Emotion</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>2.</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>3.</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>4.</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
</tbody>
</table>

**Score**

Affect Naming and Prosody-Face Matching: Score 1 point for each response. Prosody-Face Matching: See Administration and Scoring Manual.

2 ACS Social Cognition Record Form
<table>
<thead>
<tr>
<th>#</th>
<th>Speaker</th>
<th>Prosody-Pair Matching</th>
<th>Conduction</th>
<th>Meaning Change</th>
<th>Speaker Meaning Response</th>
<th>Score</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>14</td>
<td>Back off, don't you ever speak that way to me.</td>
<td>1 2 3 4</td>
<td></td>
<td></td>
<td></td>
<td>Y N</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Did you finish the report?</td>
<td>1 2 3 4 0 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Don't work too hard.</td>
<td>1 2 3 4 0 1</td>
<td></td>
<td></td>
<td></td>
<td>Y N</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>Don't hurt yourself; that's heavy.</td>
<td>1 2 3 4 0 1</td>
<td></td>
<td></td>
<td></td>
<td>Y N</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>Thanks for your help.</td>
<td>1 2 3 4 0 1</td>
<td></td>
<td></td>
<td></td>
<td>Y N</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>Your brother is a real gent.</td>
<td>1 2 3 4 0 1</td>
<td></td>
<td></td>
<td></td>
<td>Y N</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>You didn't get the tickets.</td>
<td>1 2 3 4 0 1</td>
<td></td>
<td></td>
<td></td>
<td>Y N</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>I think you should shoot him with that.</td>
<td>1 2 3 4 0 1</td>
<td></td>
<td></td>
<td></td>
<td>Y N</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>That's a big dog.</td>
<td>1 2 3 4 0 1</td>
<td></td>
<td></td>
<td></td>
<td>Y N</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>If you do that again, I'm going to punch you.</td>
<td>1 2 3 4 0 1</td>
<td></td>
<td></td>
<td></td>
<td>Y N</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>I'm sure that will make you happy.</td>
<td>1 2 3 4 0 1</td>
<td></td>
<td></td>
<td></td>
<td>Y N</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>He dropped the ball again.</td>
<td>1 2 3 4 0 1</td>
<td></td>
<td></td>
<td></td>
<td>Y N</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>I can really count on him.</td>
<td>1 2 3 4 0 1</td>
<td></td>
<td></td>
<td></td>
<td>Y N</td>
<td></td>
</tr>
</tbody>
</table>
Appendix 23: Social story 1 for RMHA familiarisation

Wearing My Roger System

Starting next week, I am going to wear a Roger system during class. One part of it goes over my ear and the other part goes in my ear.
I will ask my teacher for help to put it on, and I will tell them if it does not feel right.
I will put it on before class starts. I will try to wear it for the whole day, but I can take it off if it starts to feel uncomfortable at any time.
I will take it off before I go out to play during break time.
Appendix 24: Social story 2 for RMHA familiarisation

How My Roger System is Helping Me

When I am in class, I will wear my Roger system. Roger systems help students to hear the teacher more clearly. My teachers think the Roger system will help me to concentrate and understand. My teachers want to see how well I can listen and do my work when it is on. My Roger system is not too loud and it is comfortable to wear.

My Roger System
Appendix 25: Script for RMHA demonstration and introduction video

Section 1: Hello! Alex, Molly, and Alice are going to teach you how to use your new Roger system!

Researcher: Today in the classroom...

Alex: Hi, my name is Alex (waving).

Molly: Hi, my name is Molly (waving).

Molly: This is my Roger system (showing receiver worn in ear). Sometimes, I have trouble hearing, so I wear my Roger system to help me. It helps me hear by being not too loud and not too quiet, and being just right so I can hear my teacher.

Section 2: How to put on my Roger system.

Alex: Every morning before school, I wear my Roger system. First, I need to close the battery door. Secondly, I put the receiver behind my ear. Then, I put the little cone shape into my ear. Lastly, I curve the long thin tube into the inside of my ear. I wear both receiver throughout the day – red for my right ear, and blue for my left ear.

Section 3: Parts of my Roger system.

Molly: There are 3 parts to the Roger system. 2 receivers, which are in my ears (takes them out to show). And a transmitter, which my teacher wears (holds up to show).

Alex: The receivers in my ear make me feel like my teacher is right next to me talking to me. This helps me to hear instructions clearly and more directly. I can now learn new things with less trouble. My teacher’s voice goes right into my ear. My Roger system helps me to concentrate, and my parents are really proud of me now that I can pay attention and do better in school.

Molly: I will try to wear my Roger system at school for half the day until lunchtime. Before I go out to play during lunch break, I will ask my teacher to help me take off both earpieces. I will put them back on after lunch. I must remember to put my Roger system in my bag to take home at the end of the day.
Alex: I must never forget to check that my Roger system is in my bag to take home. I must have my ear pieces in the box, like this (demonstrate), and my transmitter must be safe in the bag, like this (demonstrate). Tomorrow, I will wear them in class again. It was really great telling you all about my Roger system!

Researcher: And in the home…

Alice: Hi, I am Alice (waving). I like my Roger system. It is a cool, new way to help me hear better.

Section 4: Parts of my Roger system.

Alice: There are 3 parts to the Roger system. 2 receivers (takes them out of ears to show). And a transmitter (holds up to show).

Section 5: How to put on my Roger system.

Alice: At home, I put on the Roger system like this. First, we close the battery door. Next, I put the receiver behind my ear. Then, I put the little cone shape into my ear. Lastly, I curve the long thin tube around the inside of my ear. I wear both receivers on my ears – I wear the red one on the right, and the blue one on the left.

Section 6: Using my Roger system at home.

Alice: At home, I like doing computer activities while wearing my Roger system. The transmitter is placed near the speakers (demonstrate clipping on of microphone). The activities are very easy when I’m wearing my Roger system! Sometimes, I have trouble hearing. My Roger system makes other people’s voices clearer and louder to help me focus. Thank you for letting me tell you all about my Roger system!

Section 7: Why I like my Roger system (not scripted, comments from children are compiled). Four questions that were asked by the researcher:

1) What was it like before you got your Roger system?

2) What has improved for you since you started wearing it?

3) Is it hard or uncomfortable to wear?

4) What do your friends think about your Roger system?


Näätänen, R. (2001). The perception of speech sounds by the human brain as reflected by the mismatch negativity (MMN) and its magnetic equivalent (MMNm). *Psychophysiology, 38*, 1-21.


