

Article

Improving Emotion Perception in Children with Autism Spectrum Disorder with Computer-Based Training and Hearing Amplification

Joan H. Leung *, Suzanne C. Purdy  and Paul M. Corballis

School of Psychology, The University of Auckland, Auckland 1023, New Zealand;
sc.purdy@auckland.ac.nz (S.C.P.); p.corballis@auckland.ac.nz (P.M.C.)

* Correspondence: joan.leung@auckland.ac.nz

Abstract: Individuals with Autism Spectrum Disorder (ASD) experience challenges with social communication, often involving emotional elements of language. This may stem from underlying auditory processing difficulties, especially when incoming speech is nuanced or complex. This study explored the effects of auditory training on social perception abilities of children with ASD. The training combined use of a remote-microphone hearing system and computerized emotion perception training. At baseline, children with ASD had poorer social communication scores and delayed mismatch negativity (MMN) compared to typically developing children. Behavioral results, measured pre- and post-intervention, revealed increased social perception scores in children with ASD to the extent that they outperformed their typically developing peers post-intervention. Electrophysiology results revealed changes in neural responses to emotional speech stimuli. Post-intervention, mismatch responses of children with ASD more closely resembled their neurotypical peers, with shorter MMN latencies, a significantly heightened P2 wave, and greater differentiation of emotional stimuli, consistent with their improved behavioral results. This study sets the foundation for further investigation into connections between auditory processing difficulties and social perception and communication for individuals with ASD, and provides a promising indication that combining amplified hearing and computer-based targeted social perception training using emotional speech stimuli may have neuro-rehabilitative benefits.

Keywords: Autism Spectrum Disorder; speech prosody; auditory processing; hearing amplification; training intervention; cortical auditory evoked potentials



Citation: Leung, J.H.; Purdy, S.C.; Corballis, P.M. Improving Emotion Perception in Children with Autism Spectrum Disorder with Computer-Based Training and Hearing Amplification. *Brain Sci.* **2021**, *11*, 469. <https://doi.org/10.3390/brainsci11040469>

Received: 20 February 2021

Accepted: 6 April 2021

Published: 8 April 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Autism Spectrum Disorder (ASD) is a lifelong, pervasive, neurodevelopmental condition characterized by restrictive, repetitive patterns of behavior and deficits in social language and communication [1]. People with ASD often experience difficulty understanding irony, conflicting emotions, jealousy, social blunders, and others' intentions [2,3]. Accordingly, a number of researchers have suggested that the core language impairments in ASD reflect problems with language pragmatics [4–8]. Pragmatics, here, refers to the social, emotional, and communicative elements of language, which include nonverbal communicative cues such as prosody—the minimal distinctions in spoken language that convey the speaker's emotion and intent [9,10].

The perception of prosody is thought to depend on a number of acoustic parameters such as variations in fundamental frequency (pitch), intensity, timbre, and timing [11–14]. Electrophysiological studies indicate that the human auditory system engages immediately with prosodic cues while processing spoken language [15], and that specific cortical regions respond to variation in affective prosody (i.e., emotion) [16].

Because perception of prosody relies on the accurate processing of subtle variations in acoustic parameters, the perception of the subtle nuances in social or emotional speech

can be impaired by difficulties with auditory pattern recognition, auditory discrimination, sound localization, and temporal processing [17]. It is likely that at least a portion of the difficulties in prosodic processing experienced by people with ASD can be attributed to impaired central auditory processing of spoken language. Over the past two decades, a number of researchers have documented atypical auditory processing in individuals with ASD [18–21]. While some studies show that children with ASD have superior pitch discrimination and categorization in music processing [22,23], they do not show the same advantage in responses to speech [24].

Auditory processing challenges become more salient in individuals with ASD when the input is more complex, such as with tasks involving sentences and intonation [25], speech recognition in the presence of background noise [26,27], auditory filtering [28], competing speech streams [29,30], and when prosody-based cues are processed simultaneously with semantic information [31,32]. These challenges are more disruptive in children with ASD than they are in adults [21].

Electrophysiological studies of individuals with ASD have largely measured cortical auditory evoked potentials (CAEPs), and have mainly focused on ascertaining whether individuals with ASD possess the auditory processing abilities to detect prosodic changes, for example in phonemic intensity, pitch, and duration [33–35], and emotional undertones in the speaking voice [36–38]. Studies have demonstrated that delayed responses (longer latencies) across various CAEP components are correlated with impaired vocal affect recognition [39,40]. Studies investigating pre-attentive processing of vocal emotion [41–43] have typically calculated a mismatch response (MMR) from CAEPs recorded in an oddball experimental paradigm. The MMR is derived by subtracting evoked potentials elicited in response to a numerous repeated “standard” stimulus, from evoked potentials elicited in response to a less frequently and randomly presented “deviant” stimulus [44].

The foregoing suggests that training that targets central auditory processing—particularly for the acoustic discrimination abilities involved in prosodic processing—could ameliorate problems in social communication in people with ASD, perhaps addressing secondary problems such as academic underachievement, inattention, hypersensitivity, hyperactivity, and oppositional behavior [45], as well as noise sensitivity and anxiety [46]. Training studies with children with ASD [47–49] have demonstrated measurable post-training improvements, both in terms of improved behavioral performance [47,48] and in faster neural responses [49].

Several research groups have pursued interventions using sound amplification technology—remote-microphone hearing systems (RMHSs)—with the aim of managing the auditory functioning of children with ASD and other language and learning disorders [50–56]. These studies have reported benefits to academic performance in noisy classroom environments, reduced listening difficulties, improved speech recognition, and improved spatial and temporal processing of auditory stimuli [50–54,56], as well as decreased cortisol levels, indicative of reduced stress [55].

In this study, we trial an intervention in a group of children with high-functioning ASD in which we combine and RMHSs with a computerized training program in emotion perception. There is growing interest in this approach of enhancing traditional auditory training tasks with hearing amplification devices, with innovative studies from the University of North Texas reporting behavioral [53] and neuro-electrophysiological improvements [56].

We seek to expand on the emerging view that the difficulties people with ASD experience in perceiving and understanding subtle nuances in speech may, in part, stem from underlying auditory processing difficulties. We present behavioral and electrophysiological results, measured pre- and post-intervention, from the ASD group and a control group of typically developing (TD) peers who were only assessed at one time point. We aimed to address the following questions:

1. Do social perception abilities differ between the TD and ASD groups?
2. Do these differences change after the training intervention?
3. Do neural responses to changes in emotional undertones in speech differ between the TD and ASD groups?

4. Do neural responses change after the training intervention?

2. Materials and Methods

2.1. Ethical Approval

All participants gave their written informed consent for inclusion before they participated in this study. This study was approved by the University of Auckland Human Participants Ethics Committee (reference number 9657).

2.2. Participants

We recruited two groups of children—an ASD group and a TD control group. The ASD group consisted of 12 children with existing diagnoses of ASD previously made by an individual or multidisciplinary team of health care practitioners. As part of the eligibility criteria for this study, the children with ASD were also evaluated with a rating scale of autism symptoms to re-confirm their diagnosis. There were 9 males and 3 females ($M = 9.91$ years, $SD = 2.23$ years, $Range = 7–13$ years). Two children in this group had comorbid diagnoses of dyslexia (they were twins), and one child had ADHD. Of the 12 children with ASD, two did not assent to the evoked potential recording sessions, therefore CAEP results are only available for 10 of the 12 children with ASD.

The TD group consisted of 14 children with no current or previous diagnosis of any clinical disorders, as confirmed by their parents/caregivers. There were 4 males and 10 females ($M = 9.43$ years, $SD = 1.87$ years, $Range = 7–12$ years).

Participants in this study were all born in New Zealand, and would have been exposed to verbal cues, and prosodic intonation and expression specific to New Zealand English. All participants were enrolled in mainstream school. Two children with 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 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, since the public schools (where students enroll by residential zones) had the resources to cater for students with learning difficulties.

2.3. Study Design

We used a prospective longitudinal study design to evaluate intervention outcomes for children with ASD. We assessed participants with ASD four times, twice before and twice after an intervention period of three weeks. Session 1 served as a baseline assessment. Evaluations of autism symptoms and communication behaviors were completed. The children completed a hearing screen, a behavioral social perception assessment, and we recorded their baseline CAEPs.

Two weeks later, Session 2 involved the re-assessment of the behavioral social perception task only and served as a second baseline reference point. During this session, we fitted and verified RMHSs for each child with ASD in preparation for the 3 week intervention period.

Session 2 was followed by a week-long familiarization period, during which the ASD group were introduced to using RMHS and briefed about the upcoming intervention. They also received various aids to support this familiarization, including social stories, step-by-step instructions, and a demonstration video (contact corresponding author to request material).

For the 3 week intervention period, all children with ASD completed nine computerized training sessions. These were 20–30 min in duration and occurred three times a week. The activities were all completed on a Dell Latitude laptop in the presence of the researcher. The children wore their RMHS while they were engaged in these training sessions. With agreement from each child with ASD, and the support of parents and teachers, the ASD group also wore their RMHS in school for the duration of the intervention period to maximize exposure to amplified hearing. Full details on the development of intervention and CAEP materials and processes are outlined in the supplementary material.

Session 3 took place the week immediately after the intervention period to evaluate intervention-related effects. The behavioral social perception assessment was re-administered, and CAEPs re-recorded. Finally, Session 4 occurred after a further two weeks. During this session, the behavioral social perception assessment was administered for the final time as an evaluation of the retention of any intervention-related effects.

The children from the TD control group only participated in a one-off session that included the hearing screen, parental report of communication behaviors, the behavioral social perception assessment, and recordings of their CAEPs. It is noteworthy that we were only able to test the TD children once each, rather than the four times for the ASD group, so the TD group cannot be considered a true control group. Rather, the TD group provided data that we used for cross-sectional comparison of behavioral and CAEP results.

We offered all participants the choice of having their behavioral assessments and training sessions (for the ASD group) at the University of Auckland's clinic space, or in a quiet space in their own homes. It was important to ensure that the children with ASD had the opportunity to be in a space where they were comfortable and least anxious. We recorded all CAEPs at the University of Auckland in a sound-attenuating chamber.

2.4. Assessment Materials and Apparatus

2.4.1. Hearing Screen

All participants received an evoked otoacoustic emissions screen measured via distortion product (DPOAE) using a Grason-Stadler GSI Audioscreener (version 3.21). OAE testing is a recommended hearing screening approach for young children [57,58]. The pass criterion was a signal-to-noise ratio (SNR) above 6 dB across five frequency levels (2–6 kHz) [59,60].

2.4.2. Autism Characteristics and Communication Behaviors

The primary researcher, who met the requirement of qualification level C, observed the ASD participants during Session 1 and completed the high-functioning version of the Childhood Autism Rating Scale (CARS-HF) [61]. Results from the CARS-2 questionnaire for parents/caregivers (CARS-2-QPC) were used to complement and validate the autism symptom severity ratings given by the researcher. Standardized T-scores were derived from the raw total and corresponded 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–moderate symptoms, and 51 and higher indicates severe symptoms of ASD.

Parents/caregivers of all children (TD and ASD) completed the second edition of the Childhood Communication Checklist (CCC-2) [62]. 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 behaviors commonly displayed by children with ASD. A General Communication Composite (GCC), scaled to individual age groups, indicated whether children may have clinically significant communication issues. Children with ASD (and specific language impairment) score below 55 according to validation data [62]. A Social Interaction Deviance Composite (SIDC), calculated separately, indicated 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 [62].

2.4.3. Behavioral Social Perception

We assessed social perception using the Wechsler Advanced Clinical Solutions (ACS) Social Perception Subtests [63]. 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.

We derived two scores from these subtests—an ‘Affect Naming’ score, and a ‘Social Perception Prosody’ score (face- and pair-matching tasks combined), as per the ACS guidelines. Raw scores were used because published normative data are not available for individuals younger than age 16 years [63].

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% [63]. The Social Perception Subtests have previously been successfully administered to adults with high-functioning ASD, Asperger’s syndrome, and a typically developing control group [64], and revealed significantly worse performance from those with ASD compared to controls, only for social perception face- and prosody-matching tasks, but not Affect Naming.

2.4.4. Cortical Auditory Evoked Potentials

Stimuli and Sequences

Speech stimuli were sampled from existing recordings of monosyllables (/ba/) produced by a male speaker in Angry, Happy, Sad, and Neutral emotional tones of voice [65]. See supplementary material for information on the selection of the four speech stimuli. All the speech stimuli were 200 milliseconds (ms) in length. The time waveforms in Figure 1 show that the speech stimuli were, on average, matched for root mean square intensity and duration but differed in their temporal characteristics.

We used a modified “oddball” paradigm to investigate auditory discrimination of Angry/Happy/Sad against Neutral. The stimuli were programmed into sequence blocks using the NeuroScan STIM2 Gentask software. Each block consisted of a ratio of 70 standard stimuli to 10:10:10 deviant stimuli. 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 randomized but was adjusted so that at least 2 or 3 standards occurred between each deviant. There was a 640 ms inter-stimulus interval.

Experimental Setup and Data Acquisition

CAEP recording sessions took place in a sound-treated two-room setup, with a leather reclining chair for the participants to sit in. Stimuli sequences were presented (via Gentask software on NeuroScan STIM2) at 70 dB SPL via an Australian Monitor Synergy SY400 power amplifier and Sabine Graphi-Q GRQ-3102 equalizer, connected to a Turbosound IMPACT 50 loudspeaker. A half-inch polarized 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.

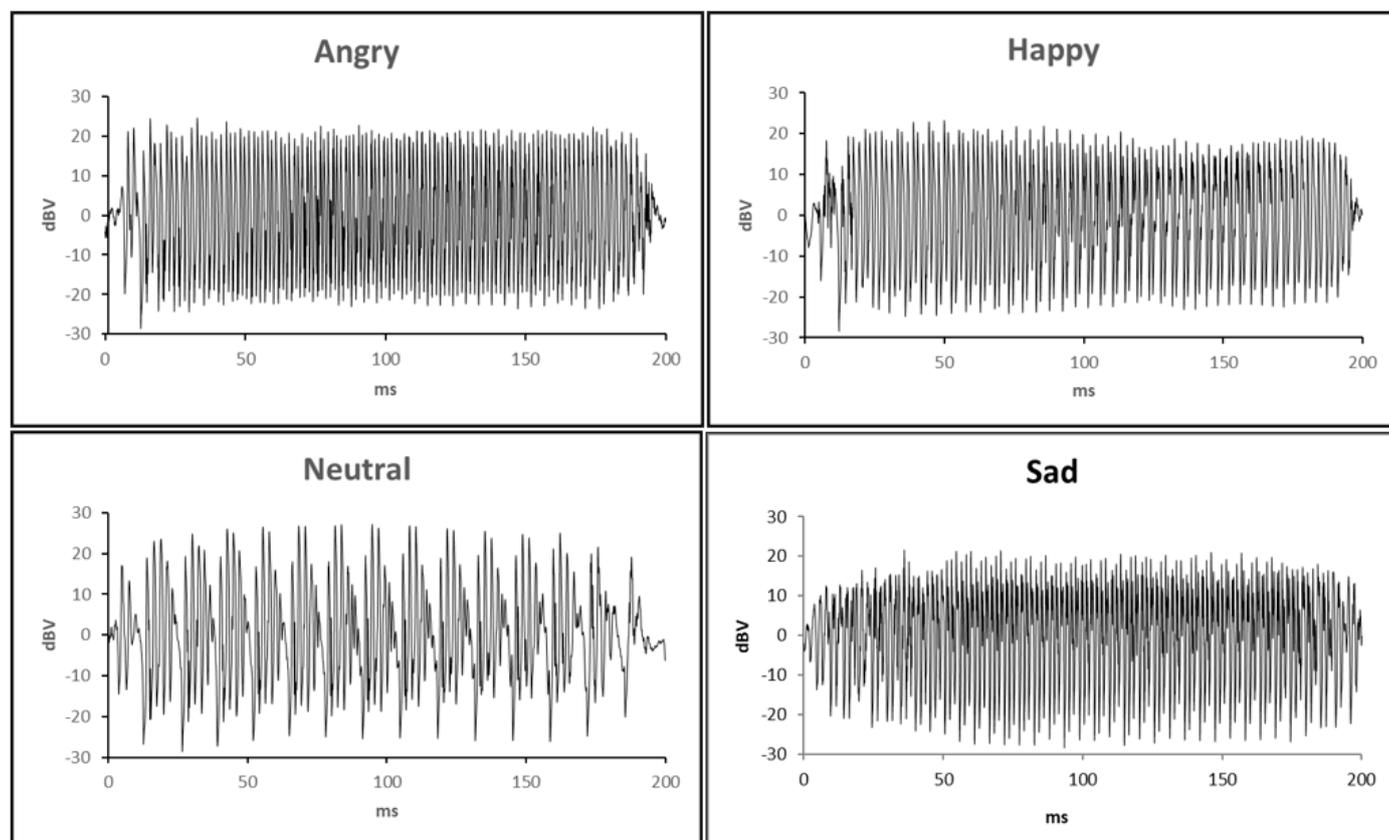


Figure 1. Electrical time waveforms (Adobe Audition CS6) depicting the monosyllabic stimulus /ba/ stimulus presented in four emotions. dBV represents relative sound intensity in decibel-voltages, plotted against time in milliseconds (ms).

We positioned the loudspeaker at a 150 cm distance at zero degrees azimuth in front of the participant seated on the recliner. Behind the loudspeaker was a television on a stand. We instructed participants to watch a movie of their choice with the audio turned off and the subtitles on, and to minimize their blinking and body movements during the recordings.

We recorded CAEPs using the NeuroScan Inc. Evoked Potential System (version 4.5) with a SynAmps 2 amplifier. 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, we linked the left and right mastoid electrodes, and re-referenced Cz, Fz, F3, and F4. We kept electrode impedance at or below 5 k Ω and used a sampling rate of 500 Hz and a bandpass filter setting of 0.1–100 Hz.

Data Processing

Post-acquisition, we performed further offline processing 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 μ V were rejected as artifacts. A minimum of 20 blinks were required to estimate an average blink. We filtered the data using a low-pass filter at 30 Hz (12 dB/octave slope, zero phase shift). We generated separate average files for each participant for each of the four emotions.

On average, approximately 20% of responses were rejected from the TD participants, as a result of ocular and other noise artefacts. For the children with ASD, an average of approximately 40% of responses were rejected.

All grand average waveforms are 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 complexities of using emotions tones as auditory stimuli contribute to the current dearth of evidence in the mismatch response literature, and there is little consensus on what a “typical mismatch waveform” looks like and what components are expected to be observed. The heterogeneous nature of CAEPs recorded from clinical populations like children with ASD adds to this, thus we took some additional steps to maximize consistency in the analysis of individual waveforms and quantify latencies and magnitudes of specific waveform components in a consistent manner.

Firstly, we averaged responses across individuals within each group of waveforms (TD, Pre-intervention ASD, and Post-intervention ASD) to create grand averaged waveforms for the standard (Neutral) and deviant (collapsed across Angry, Happy, and Sad) stimuli. We then subtracted the standard waveform from the deviant waveform to create a difference waveform for each participant group, which was then collapsed across electrode sites.

We conducted single-sample t -tests on these difference waveforms at each millisecond. We highlighted contiguous time periods of more than 30 ms where the waveform deviated at a significance level of $p \leq 0.001$ from 0 μ V. These contiguous time periods formed “mismatch windows”.

We superimposed the mismatch windows back onto original CAEP recordings from each electrode site for each individual participant. They were used as references for peak latency and mismatch response (MMR) magnitude quantification. MMR magnitudes were calculated by taking an absolute average of the amplitudes included within ± 20 ms either side of the peak latency for shorter windows (<100 ms), or by ± 50 ms either side of an approximate midpoint for the longer windows (>100 ms).

Only the selected windows were used to provide comparable data between groups. Each individual participant had the following data for each relevant mismatch window, for each electrode site:

- Peak latency for Combined-emotions-minus-Neutral difference waveform,
- Peak latency for Angry-minus-Neutral difference waveform,
- Peak latency for Happy-minus-Neutral difference waveform,
- Peak latency for Sad-minus-Neutral difference waveform,
- MMR magnitude for Combined-emotions-minus-Neutral difference waveform,
- MMR magnitude for Angry-minus-Neutral difference waveform,
- MMR magnitude for Happy-minus-Neutral difference waveform, and
- MMR magnitude for Sad-minus-Neutral difference waveform.

2.4.5. Statistical Analyses

The following statistical analyses (using IBM SPSS Statistics, v20.0) were conducted:

1. Related-samples analyses to determine whether behavioral results from the social perception assessment differed from each other at the two baseline time points, and at the two post-intervention time points.
2. Independent-samples analyses to compare behavioral social perception results between the TD and ASD groups (pre- and post-intervention vs. TD).
3. Related-samples analyses to explore the effects of the intervention on behavioral social perception within the ASD group.
4. Repeated-measures analyses of variance (ANOVAs) were conducted on peak latencies and MMR magnitudes within each mismatch window to explore differences between each electrode site (Cz, Fz, F4, and F3).
5. Independent-samples analyses to compare peak latencies and MMR magnitudes between the TD and ASD groups (pre- and post-intervention vs. TD), for Combined-emotions difference waveforms.

6. Related-samples analyses to explore the effects of the intervention on peak latencies and MMR magnitudes between the ASD pre- and post-intervention waveforms, for Combined-emotions difference waveforms.
7. Related-samples analyses to explore peak latencies and MMR magnitudes differences between emotions within each group of waveforms (TD, ASD pre-, and ASD post-intervention).

3. Results

3.1. Participant Characteristics

All participants passed the DPOAE screen, 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. TD and ASD groups did not differ significantly from each other, and both groups were within normal range [66].

Autism characteristics from the CARS-2 ratings showed that nine of the 12 children in the ASD group obtained standardized T-scores in line with “minimal severity” ($M = 33.44$, $SD = 3.21$), and three of the children scored in the mild–moderate ASD severity category ($M = 47.67$, $SD = 1.53$). Results from the CARS-2 re-confirmed the diagnoses of all the children in the ASD group.

Table 1 details mean DPOAE signal-to-noise ratios (SNR), from left and right ears separately, measured in decibels (dB) across frequencies 2–6 kHz; group mean, minimum, and maximum T-scores from the CARS-2 (from ASD participants only); and group mean and standard deviations of General Communication Composite (GCC) and Social Interaction Deviance Composite (SIDC) scores from the CCC-2 (communication behaviors). Parental reports from the ASD group highlighted significantly more communication difficulties compared to the TD group, for both the GCC ($t_{(24)} = -4.24$, $p < 0.001$), and the SIDC ($t_{(24)} = -2.55$, $p = 0.017$).

Table 1. DPOAE (dB), CARS-2, and CCC-2 results from TD and ASD participants.

Measure		TD Group ($n = 14$)	ASD Group ($n = 12$)
DPOAE SNRs (right ear)	2 kHz	12.66	12.89
	3 kHz	14.62	14.63
	4 kHz	13.11	14.82
	5 kHz	13.69	14.37
	6 kHz	14.97	15.49
DPOAE SNRs (left ear)	2 kHz	12.97	14.64
	3 kHz	14.73	15.17
	4 kHz	13.35	15.77
	5 kHz	13.64	14.97
	6 kHz	15.75	15.40
CARS-2	Mean	-	37.00
	Minimum	-	29.00
	Maximum	-	49.00
CCC-2 (GCC)	Mean	66.36	38.42
	SD	18.92	13.75
CCC-2 (SIDC)	Mean	-1.86	-8.50
	SD	5.61	7.62

3.2. Behavioral Social Perception

Behavioral results from the social perception assessment were all normally distributed, according to the Shapiro–Wilk test, with the exception of the Affect Naming scores at Session 3 and 4. Parametric paired-samples t -tests (and Wilcoxon signed ranks test for not normally distributed variables) were used to compare Affect Naming and Social Perception Prosody scores between Session 1 vs. 2, and Session 3 vs. 4. These scores (at baseline, and post-intervention) were not significantly different. Hence, results from Sessions 1 and

2 were averaged together to form a “pre-intervention” score; and Sessions 3 and 4 were averaged together to form a “post-intervention” score.

For the Affect Naming score, independent-samples t -test results showed that the TD group ($M = 19.29$, $SD = 2.23$) performed significantly better ($t_{(24)} = -4.33$, $p < 0.001$) than the ASD group pre-intervention ($M = 15.37$, $SD = 2.37$). Paired-samples t -test results showed that the ASD group improved their scores significantly after participating in the 3 week intervention period ($M = 21.33$, $SD = 0.94$) ($t_{(11)} = -9.71$, $p < 0.001$). Cohen’s d for repeated measures yielded an effect size of 2.96 for the pre- vs. post-intervention comparison. Independent-samples t -test results showed that the ASD group surpassed the TD group on Affect Naming performance post-intervention ($t_{(17.99)} = -3.12$, $p = 0.006$).

For the Social Perception Prosody score, independent-samples t -test results showed that the TD group ($M = 18.79$, $SD = 2.29$) performed significantly better ($t_{(24)} = -4.75$, $p < 0.001$) than the ASD group pre-intervention ($M = 13.75$, $SD = 3.10$). Paired-samples t -test results showed that the ASD group improved their scores significantly after participating in the 3 week intervention period ($M = 19.42$, $SD = 1.16$) ($t_{(11)} = -7.36$, $p < 0.001$). Cohen’s d for repeated measures yielded an effect size of 2.05 for the pre- vs. post-intervention comparison. Independent-samples t -test results showed that the ASD group did not significantly differ from the TD group on Social Perception Prosody performance post-intervention ($t_{(24)} = -0.86$, $p = 0.398$). Figure 2 illustrates these results.

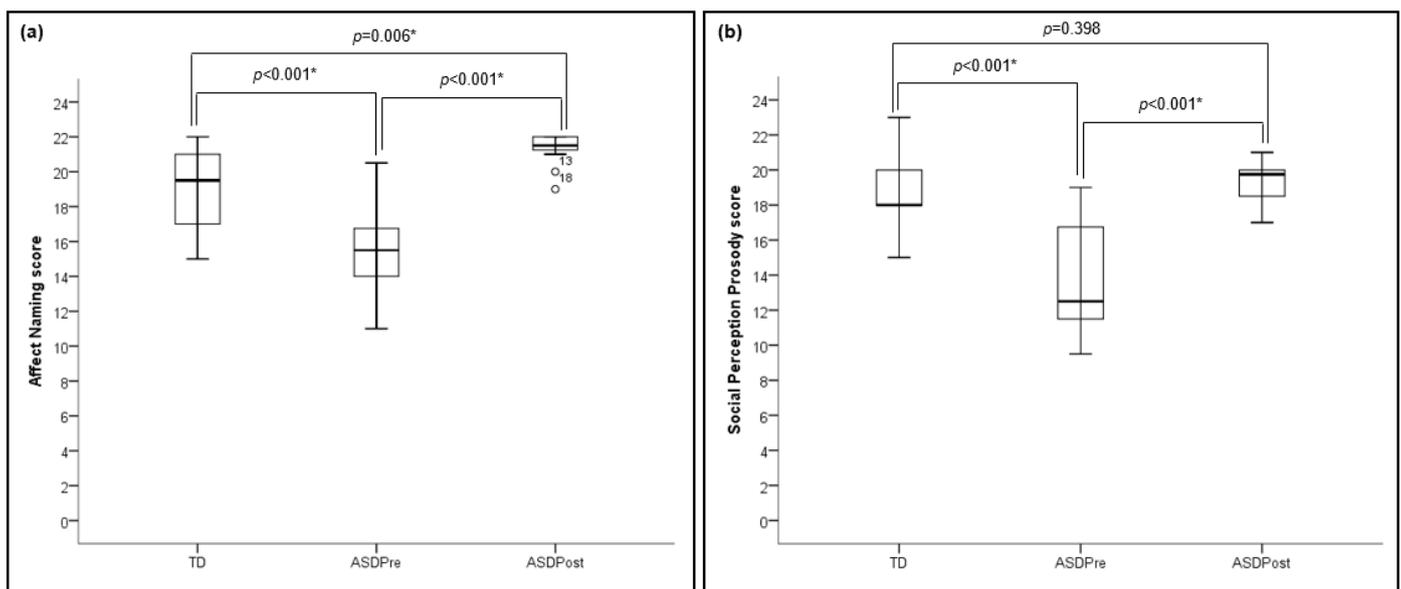


Figure 2. Behavioral social perception results from the Wechsler ACS Social Perception Subtests highlighting differences in performance between TD and ASD groups; and differences in performance pre- and post-intervention for the ASD group. (a) results for Affect Naming scores; and (b) results for Social Perception Prosody scores. * represent statistically significant differences.

3.3. Cortical Auditory Evoked Potentials

Figure 3 shows the Combined-emotions difference waveforms derived for each group of waveforms (TD, ASD pre-intervention, and ASD post-intervention). As detailed in the Methods, single-sample t -tests were conducted on these difference waveforms at each millisecond. Contiguous time periods where the waveform deviated from 0 μV at a significance level of $p \leq 0.001$ formed the “mismatch windows” illustrated in Figure 4.

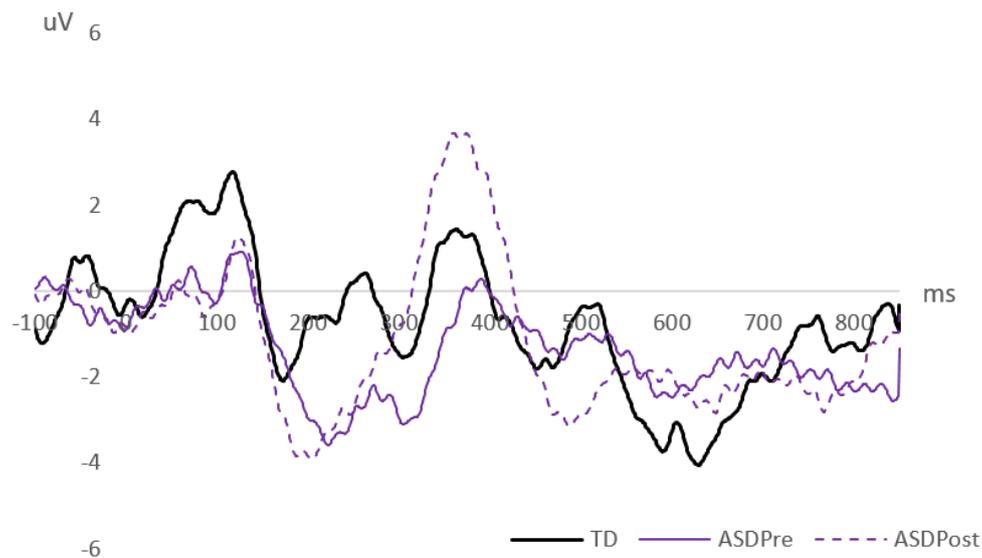


Figure 3. Difference waveforms (deviants-minus-standard), combined across emotions and electrode sites, resulting in a single grand average difference waveform for each group—TD children (thick solid [black] line), and ASD children pre-intervention (thin solid [purple] line) and post-intervention (thin dotted [purple] line). Response amplitudes are measured in microvolts (μV) and time in milliseconds (ms).

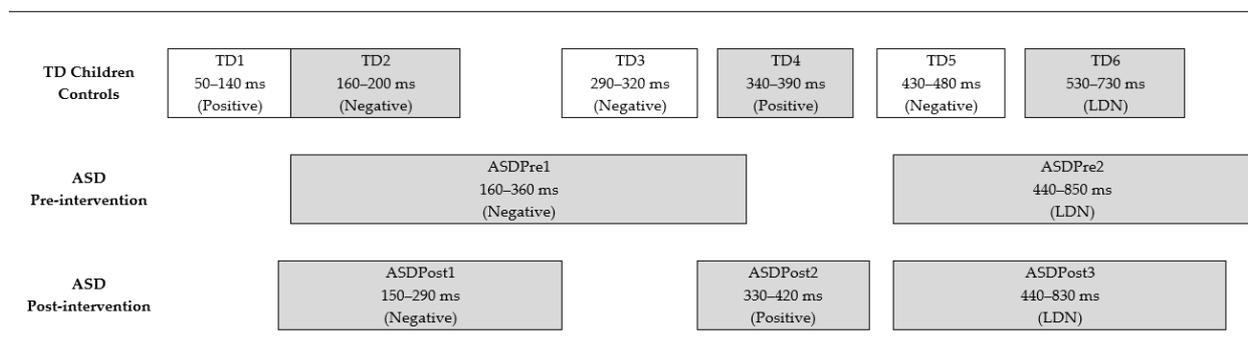


Figure 4. Timeline presentation of mismatch windows for each participant group, derived by conducting single *t*-tests at each time point of group average emotion-combined difference waveforms, and identifying contiguous periods of significant deviance from 0 μV .

Mismatch windows consisted of both positive and negative deviances, and ranged from 30 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 regard to language and speech processing [67–69]. Long deviance periods were clearly identified in both groups of children in this study.

The shaded mismatch windows in Figure 4 represent the selected ones where peak latency and MMR magnitude data were computed. The first negative window (“MMN”) has data for all three groups (TD, ASD pre-intervention, and ASD post-intervention). The following positive window (“MMP”) has data for TD and ASD post-intervention comparisons. Data for the last negative window “LDN” can be compared across all three groups of waveforms.

Repeated-measures ANOVAs were conducted within each mismatch window to explore differences between electrode sites (Cz, Fz, F4, and F3). For the ASD group (pre- and post-intervention), there were no significant differences between electrode sites. For the TD group, there was a significant main effect of electrode for MMR magnitudes at two mismatch windows and peak latencies at one mismatch window. However, Bonferroni

corrected post hoc comparisons revealed no significant differences between electrodes. Thus, peak latency and MMR magnitude data were averaged across electrode sites.

Table 2 details the results from the between-group (independent-samples *t*-tests) analyses comparing peak latencies and MMR magnitudes for Combined-emotions difference waveforms. Table 2 also details pre vs. post analyses (paired-samples *t*-tests) to explore the effect of the intervention on CAEPs for the children with ASD.

Table 2. Comparative analyses of peak latency and MMR magnitude data between three participant groups (TD, ASD pre-intervention, and ASD post-intervention) for combined emotions difference waveforms. * denotes group differences at a significance level of $p \leq 0.05$.

Difference Waveform	Measure	Group 1 (M, SD)	Group 2 (M, SD)	Statistical Results
Combined Emotions—Neutral	Peak latency at MMN	TD (179.32, 10.11)	ASD Pre (259.10, 18.59)	$t_{(12-81)} = 12.33, p < 0.001 *$
		TD (179.32, 10.11)	ASD Post (196.03, 9.01)	$t_{(22)} = -4.17, p < 0.001 *$
		ASD Pre (259.10, 18.59)	ASD Post (196.03, 9.01)	$t_{(9)} = 8.42, p < 0.001 *$
	Magnitude at MMN	TD (3.06, 1.60)	ASD Pre (5.05, 3.19)	$t_{(12-27)} = 1.81, p = 0.094$
		TD (3.06, 1.60)	ASD Post (4.71, 2.96)	$t_{(22)} = -1.77, p = 0.091$
		ASD Pre (5.05, 3.19)	ASD Post (4.71, 2.96)	$t_{(9)} = 0.34, p = 0.746$
	Peak latency at MMP	TD (356.27, 45.85)	ASD Post (361.85, 11.43)	$t_{(15-20)} = -0.44, p = 0.668$
		Magnitude at MMP	TD (4.46, 2.88)	ASD Post (5.27, 3.31)
	Peak latency at LDN	TD (618.20, 16.00)	ASD Pre (722.00, 38.76)	$t_{(11-21)} = 8.00, p < 0.001 *$
		TD (618.20, 16.00)	ASD Post (633.70, 24.10)	$t_{(22)} = -1.90, p = 0.071$
		ASD Pre (722.00, 38.76)	ASD Post (633.70, 24.10)	$t_{(9)} = 9.09, p < 0.001 *$
	Magnitude at LDN	TD (5.32, 2.03)	ASD Pre (3.82, 2.89)	$t_{(22)} = -1.50, p = 0.149$
TD (5.32, 2.03)		ASD Post (4.28, 3.64)	$t_{(22)} = 0.90, p = 0.378$	
ASD Pre (3.82, 2.89)		ASD Post (4.28, 3.64)	$t_{(9)} = -0.52, p = 0.615$	

Figures 3 and 4 show that at the first negative window—the MMN—the TD group demonstrated faster responses compared to the ASD group, when comparisons are conducted for both pre- and post-intervention waveforms. There were no significant differences in MMN amplitudes between groups. However, MMN latencies were significantly faster post-intervention for the children with ASD. At the positive window—the MMP—no statistically significant differences were found between responses from the TD group compared to the positive spike observed in the ASD waveform post-intervention. Finally, at the last negative window—the LDN—there were also no significant differences in the magnitude of responses.

Thus, the main group effect was for CAEP latencies; the ASD group pre-intervention showed significantly slower responses compared to the TD group. The significant increase in response speed post-intervention narrowed the gap, resulting in no significant latency differences between the TD and ASD groups after the ASD group received combined RMHS and computerized social perception training.

When the difference waveforms are examined for separate emotions, data were inconsistently normally distributed for both peak latency and MMR magnitude variables and hence non-parametric tests were used.

Figure 5 illustrates separate emotion difference waveforms for (a) the TD group, (b) the ASD group pre-intervention, and (c) the ASD group post-intervention. The shaded regions depict the MMN, MMP, and LDN windows where peak latency and MMR magnitude data were extracted and included in the related-samples analyses. Friedman tests revealed no significant latency differences between emotions for any group. Table 3 details Friedman analysis results for MMR magnitudes. For the TD group, there was a significant magnitude difference between emotions at the MMN window ($X^2_{(2)} = 7.02, p = 0.030$). For the ASD group pre-intervention, there were no magnitude differences between emotions for any mismatch window. For the ASD group post-intervention, there were significant magnitude differences between emotions at the MMP ($X^2_{(2)} = 16.20, p < 0.001$) and LDN ($X^2_{(2)} = 17.59, p < 0.001$) windows.

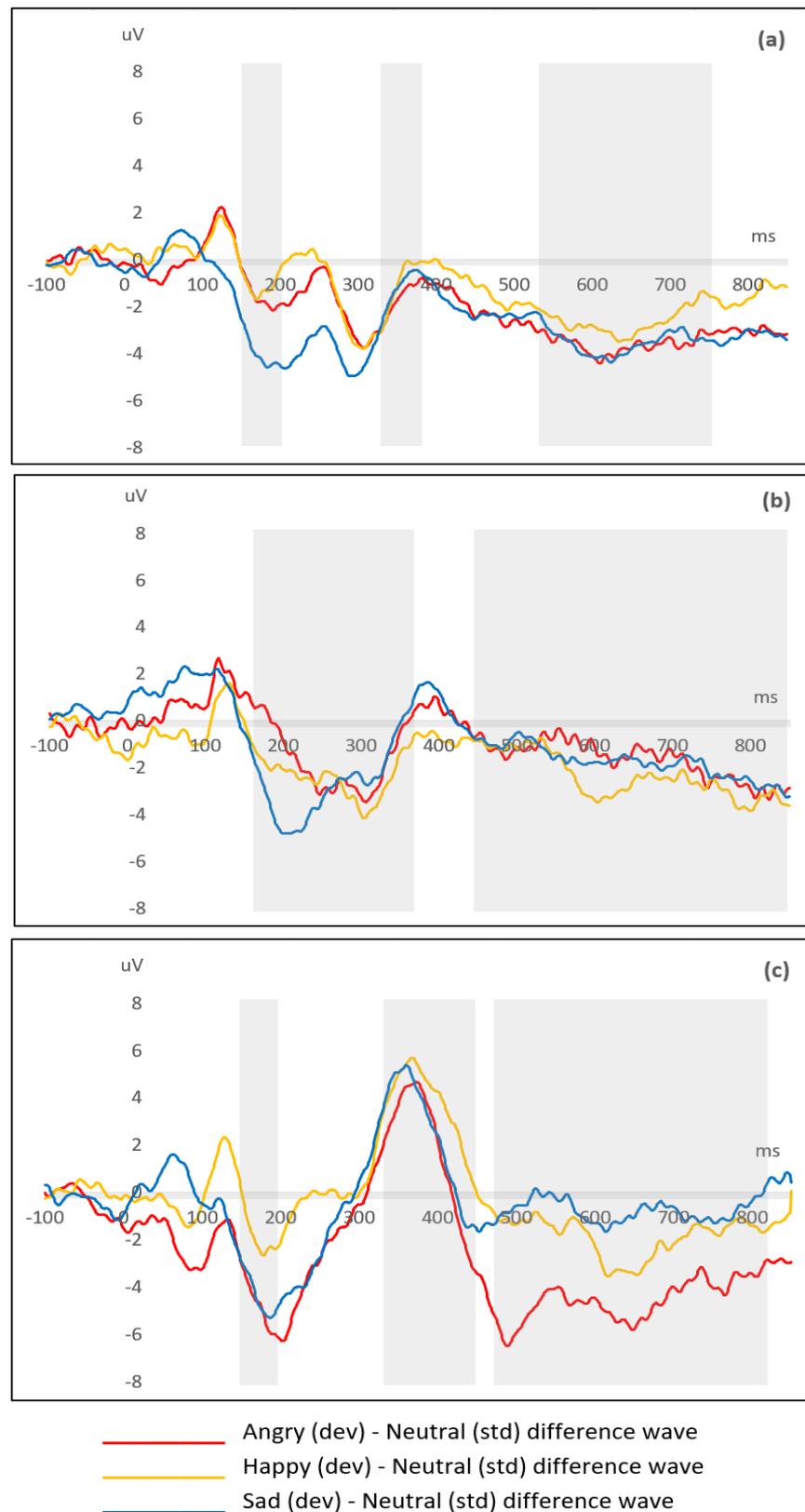


Figure 5. Separate emotion difference waveforms for (a) the TD group, (b) the ASD group pre-intervention, and (c) the ASD group post-intervention. Shaded regions depict the MMN, MMP, and LDN windows.

Table 3. Friedman analyses results (magnitude only) for all participant groups (TD, ASD pre-intervention, and ASD post-intervention) at each mismatch window. Additional information including median values, interquartile ranges, and individual Wilcoxon signed-ranks pair-wise comparison results are reported for significant main effects. * denotes emotion differences at a significance level of $p \leq 0.05$.

Participant Group	Region	Friedman Results	Emotion 1 (Mdn, IQR)	Emotion 2 (Mdn, IQR)	Wilcoxon Results
TD	MMN	$X^2_{(2)} = 7.02, p = 0.030 *$	Angry (2.50, 3.49) Angry (2.50, 3.49) Happy (2.89, 2.99)	Happy (2.89, 2.99) Sad (5.00, 3.19) Sad (5.00, 3.19)	$Z = -0.52, p = 0.600$ $Z = -2.67, p = 0.008 *$ $Z = -2.73, p = 0.006 *$
	MMP	$X^2_{(2)} = 1.71, p = 0.424$	-	-	-
	LDN	$X^2_{(2)} = 0.11, p = 0.947$	-	-	-
ASD Pre	MMN	$X^2_{(2)} = 1.40, p = 0.497$	-	-	-
	LDN	$X^2_{(2)} = 5.00, p = 0.082$	-	-	-
ASD Post	MMN	$X^2_{(2)} = 7.40, p = 0.025 *$	Angry (7.21, 7.06) Angry (7.21, 7.06) Happy (4.14, 3.22)	Happy (4.14, 3.22) Sad (8.06, 4.12) Sad (8.06, 4.12)	$Z = -2.50, p = 0.013 *$ $Z = -0.15, p = 0.878$ $Z = -2.09, p = 0.037 *$
	MMP	$X^2_{(2)} = 16.20, p < 0.001 *$	Angry (4.13, 2.32) Angry (4.13, 2.32) Happy (11.03, 4.31)	Happy (11.03, 4.31) Sad (7.73, 2.96) Sad (7.73, 2.96)	$Z = -2.80, p = 0.005 *$ $Z = -2.50, p = 0.013 *$ $Z = -2.40, p = 0.017 *$
	LDN	$X^2_{(2)} = 17.59, p < 0.001 *$	Angry (9.44, 6.39)	Happy (4.74, 3.41)	$Z = -2.67, p = 0.008 *$
			Angry (9.44, 6.39) Happy (4.74, 3.41)	Sad (3.96, 2.99) Sad (3.96, 2.99)	$Z = -2.80, p = 0.005 *$ $Z = -2.09, p = 0.037 *$

4. Discussion

This study investigated how the perception of affective prosody (i.e., emotion) differs between children with ASD and their TD peers. It was hypothesized that difficulties with perceiving and understanding subtle nuances in speech may stem from underlying auditory processing challenges, thus a second aim of this study was to evaluate the effects of an auditory-based intervention with children with ASD. Affective prosodic perception was investigated via a behavioral social perception task and cortical auditory evoked response recordings.

4.1. Behavioral Results and Implications

Overall behavioral results showed that TD and ASD children (pre-intervention) did exhibit differences in their abilities to identify facial expressions (Affect Naming score), as well as matching facial to vocal emotions (Social Perception Prosody score). Future work could expand the collection of normative data from TD children so that raw scores from the ACS Social Perception [62] test can be standardized for ages younger than 16 years, the current cut off point for available standardized data. All participants were able to complete the ACS tasks successfully, indicating the suitability of the tool for younger ages.

Significant improvement in social perception abilities was seen in the children with ASD after receiving computer-based training accompanied by amplified hearing via RMHS during the 3 week intervention period. Abilities improved to the extent that the children with ASD surpassed the TD group (who received no intervention) on Affect Naming and matched the performance of their TD peers on Social Perception Prosody scores. The TD group were not tested twice to determine whether their scores would be stable over time, however, the test-retest reliability of these measures has been previously reported [63].

A number of studies report similarly successful computer-based training for emotion recognition [70–73] and attention to prosodic cues [74,75] in individuals with ASD. A recent review [76] revealed that the largest training-related improvements for individuals with ASD result from specifically targeting, for example, speech prosody, and interventions using evidence-based practices and spanning across more than one treatment session. A clear future progression from many of the existing studies is the integration of auditory

with visual stimuli, as was done in this study, to improve social understanding and making the learnt skills more generalizable to the real world [45,77–79].

Studies that have investigated computer-based training alongside use of hearing amplification technology (RMHS) have largely focused on evaluating auditory processing specific outcomes measures, self-perceived hearing difficulties, and classroom listening behaviors [53,56]. The use of RMHSs in conjunction with prosodic-specific training has received little attention in the literature to date, and to our knowledge the effects on emotion recognition and social perception have not been reported previously.

4.2. Electrophysiological Results and Implications

This study demonstrated that the natural speech stimuli spoken with four different emotions, presented in an oddball paradigm, evoked measurable obligatory CAEP components and significant mismatch responses from children with and without ASD. Most previous auditory change-detection studies in the literature involve non-speech sounds and simple speech stimuli, which reliably elicit prominent mismatch negativities in children between 150 and 250 ms [80–82].

The presence of positive MMRs elicited by the TD children in this study is supported by other studies looking at the manipulation of speech-related factors [83,84]. Due to the complex nature of stimuli used in this study compared to earlier studies it was difficult to predict where mismatch would occur and hence the statistical identification of mismatch windows (Figures 3 and 4) was useful for documenting the responses in TD and ASD children.

TD children displayed a more complex mismatch response compared to the responses of ASD children pre-intervention who, in contrast, displayed a simpler pattern that consisted of two large mismatch negativities (Figures 3 and 4). The first negativity is consistent with the pattern reported in the literature for TD children for non-speech stimuli [80]. It is possible that the children with ASD processed the stimuli as simple sounds, as their mismatch response did not reflect the spectral and temporal complexity of the speech stimuli. Post intervention, the difference waveforms of the ASD children still showed mismatch negativity, but with significant reduction in latency, i.e., a more rapid neural response to affective prosodic changes (Table 2). This suggests that the combined intervention of computer-based training and a clearer speech signal received through the RMHSs was associated with improved auditory processing.

A large mismatch positivity, as observed in the TD group and the ASD group post-intervention, could be attributed to increased sensitivity towards changes in prosody. Consistent with this, other training studies involving TD children and different CAEP paradigms have also found evidence for improved post-training evoked responses in this latency region. For example, studies with normal hearing, neurotypical children have reported significantly enhanced P2 amplitudes after auditory discrimination training targeting different voice onset times [85,86], phoneme changes [87], and pitch [88]. A mismatch positivity was not present in the ASD group pre-intervention in the current study, but this emerged post-intervention, consistent with reports that the auditory evoked P2 is a potential biomarker of learning and plasticity [89].

In terms of individual emotion differentiation, TD children produced a significantly larger first mismatch negativity in response to the Sad emotion, whereas Angry and Happy response did not differ (Table 3). These findings suggest that emotion differentiation may primarily be pitch driven, as anger and happiness are characterized 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 [90,91]. Future work in this area would benefit from more sophisticated analyses of natural speech stimuli, which would better inform future studies regarding which parameters to manipulate. Delving deeper into whether emotion is primarily modulated by changes in pitch contours and temporal resolution, and its correlation with valence and

arousal strength [92], may deepen our understanding of the emotion perception difficulties of individuals with ASD and other neurological conditions.

Children with ASD did not show any differences between emotions in their MMRs pre-intervention (Table 3). The grand average waveforms suggest that Sad is separable from Angry and Happy waveforms, but there are large variations between individuals and the sample size is small, which may account for the lack of significant difference. These results are consistent, however, with earlier behavioral [93] and neurophysiological [36,38] studies that show impaired emotion discrimination in children with ASD.

Post-intervention, the MMRs of the ASD group looked substantially different. As highlighted in Table 3, the children now show responses that differ between the three different emotions. These differences could have been driven by perceptual differences in vocal pitch and intensity between the three emotional stimuli. Another possibility is that this change does not reflect altered auditory processing at all. For example, it could be the result of pitch- and intensity-related heightened anxiety in individuals with ASD resulting in an altered attentional effect on the mismatch response. Studies of visual attention and facial discrimination report a dominant reaction towards threatening environmental stimuli [94]. Although changes in anxiety or other factors could have contributed to MMR changes over time, behavioral scores were stable prior to and after training, suggesting that MMR differences were more likely to reflect training effects.

Emotion differentiation occurred in the earliest MMN for TD children, suggesting pre-attentive neural activity that does not engage later-occurring higher cognitive processes. This was not the case for children with ASD. Significant emotion differentiation was evident at the MMN, the MMP, and the LDN time windows (i.e., throughout the mismatch waveform) post-intervention, which suggests both pre-attentive and conscious appraisal of emotional differences. Thus, the intervention may have enabled the children with ASD to differentiate the emotions, but did not ultimately 'normalize' underlying auditory processing.

4.3. Limitations and Future Directions

The effects of the intervention on the behavioral and CAEP data should be interpreted with caution, as one of the major limitations of this study remains that there were no alternative ASD groups who received different versions of the intervention, nor was there a control group of children with ASD that received no intervention. The method of administering the behavioral test of social perception four times (twice pre- and twice post-intervention), and the resulting lack of statistical differences between Sessions 1 vs. 2, and Sessions 3 vs. 4, strongly suggests that the changes in behavioral performance are attributable to the intervention and not to test-retest effects. However, the same cannot be said for changes in the CAEPs and should be addressed in future extensions of this work.

Data from the TD group had a number of limitations, due to the need to a) administer a full range of assessment measures (both behavioral and electrophysiological) within a tight time frame, b) minimize participant fatigue, and c) minimize the imposition on TD volunteers. These limitations included not evaluating the TD children with the CARS-2 to rule out autism symptoms, only testing them at one time point, and not providing the TD children with an intervention. Future studies with sufficient time and funding could improve on this by conducting a randomized controlled trial of various interventions with ASD and TD groups.

Hyper or hyposensitivity to sound was not specifically measured in either ASD or TD groups. This is a potential confounding variable, as demonstrated in other studies investigating the neural processing of auditory information [95]. All ASD and TD children tolerated the hearing screening using DPOAEs well, which involved presentation of tones at 55–65 dB SPL across a range of frequencies from 2 to 6 kHz. We cannot, however, comment on whether the children had atypical sound sensitivity or loudness perception. Future studies in this area could benefit from including a subjective measure of sound sensitivity (e.g., the Sensory Experiences Questionnaire [96], or measurements of loudness perception

for the CAEP-evoking stimuli to determine whether these measures correlate with the mismatch response. There are also limitations to using DPOAEs as a measure of hearing, as it is most reliable for identifying the presence of moderate to severe hearing loss and does not offer a continuous measure of hearing thresholds [97]. Children with autism are at risk for peripheral hearing loss [98] as well as auditory processing difficulties [20], however with this clinical population, it is important to reduce the complexity of task demands to ensure that the quality of the data is upheld. Although desirable, use of the gold-standard method of pure tone audiometric testing may not be feasible for all children with autism and hence, in the current study, we relied on parent report and objective DPOAE measurement to screen for moderate or greater hearing loss. In future studies, it would be of interest to investigate the associations between sound sensitivity, hearing thresholds, and the neural processing of auditory information and vocal affect recognition [95,99]. This research could advance our understanding of electrophysiological biomarkers of autism, or quantifiable measurements of neural benefit from auditory-based interventions.

5. Conclusions

This study was motivated by evidence that individuals with ASD experience significant auditory processing difficulties [19,20,26]. It was hypothesized that this affects discrimination of prosodic cues such as stress, pitch, and emotion, which hinders understanding of affective speech [4,8,99]. Consistent with this, children with ASD had poorer social perception scores than TD children.

An intervention consisting of a combination of computer-based social perception training exercises, and a 3 week trial using RMHSs to provide an amplified hearing experience was administered to a group of children with ASD. Behavioral performance on measures of social perception significantly improved post-intervention. Electrophysiological results showed altered neural activity in response to changes in vocal emotion post-intervention in the ASD group. Improved performance and mismatch responses suggest that auditory training may improve the perception of affective cues in speech and that this may enhance social communication.

This study reports promising pilot data, but further work is required involving a larger sample size and a longitudinal study design. Control or sham conditions are also required to better understand the connection between underlying auditory processing difficulties, perceptual discrimination and neural processing of affective speech, and social perception and communication for individuals with ASD.

Supplementary Materials: A 3-part document is available online at <https://www.mdpi.com/article/10.3390/brainsci11040469/s1>. Part A includes details about remote-microphone hearing systems: product, verification, and familiarization period. Part B details the development of the computer-based social perception training. Part C describes the speech stimuli for cortical auditory evoked potential recordings.

Author Contributions: Conceptualization, J.H.L. and S.C.P.; methodology, P.M.C., J.H.L. and S.C.P.; formal analysis, J.H.L., P.M.C. and S.C.P.; investigation, J.H.L.; data curation, J.H.L.; writing—original draft preparation, J.H.L.; writing—review and editing, J.H.L., S.C.P. and P.M.C.; visualization, J.H.L.; supervision, S.C.P. and P.M.C.; project administration, J.H.L.; funding acquisition, J.H.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by SONOVA AG (PHONAK), grant number 3707537, in the form of a PhD scholarship.

Institutional Review Board Statement: This study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the UNIVERSITY OF AUCKLAND Human Participants Ethics Committee (reference number 9657, first submission on 5 June 2013, minor amendments on 2 October 2014).

Informed Consent Statement: Informed consent was obtained from the parents/caregivers of all children involved in this study. The children themselves gave their assent to participate in this study. Written informed consent has been obtained from all participants involved for the publication of aggregated data included in this paper.

Data Availability Statement: Data (participants de-identified) supporting reported results in this study can be supplied on request by emailing the corresponding author.

Acknowledgments: The authors would like to acknowledge the support of Stefan Launer (Sonova AG) for providing the remote-microphone hearing systems used in this research, and the support from David Crowhen (Phonak New Zealand). The authors would also like to acknowledge Erin Schafer (University of North Texas) for sharing her RMHS familiarization material as exemplars; and Autism New Zealand, the Wilson Home Trust, the New Zealand Herald, and the Minds 4 Minds Autism Network for their support with participant recruitment. Lastly, the authors would like to thank all the children with ASD and their families who have given time and energy towards supporting this research.

Conflicts of Interest: The authors declare no conflict of interest. The funders provided the hearing amplification equipment used in this study but had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*, 5th ed.; American Psychiatric Association: Arlington, VA, USA, 2013.
2. Kaland, N.; Mortensen, E.L.; Smith, L. Social communication impairments in children and adolescents with Asperger syndrome: Slow response time and the impact of prompting. *Res. Autism Spectr. Disord.* **2011**, *5*, 1129–1137. [[CrossRef](#)]
3. Wang, A.T.; Lee, S.S.; Sigman, M.; Dapretto, M. Neural basis of irony comprehension in children with Autism: The role of prosody and context. *Brain* **2006**, *129*, 932–943. [[CrossRef](#)]
4. Frith, U.; Happe, F. Language and communication in autistic disorders. *Philos. Trans. R. Soc. B* **1994**, *346*, 97–104. [[CrossRef](#)]
5. McCann, J.; Peppé, S.J.E. Prosody in Autism Spectrum Disorders: A critical review. *Int. J. Lang. Commun. Disord.* **2003**, *38*, 325–350. [[CrossRef](#)]
6. Paul, R.; Augustyn, A.; Klin, A.; Volkmar, F.R. Perception and production of prosody by speakers with Autism Spectrum Disorders. *J. Autism Dev. Disord.* **2005**, *35*, 205–220. [[CrossRef](#)] [[PubMed](#)]
7. Tager-Flusberg, H.; Caronna, E. Language disorders: Autism and other pervasive developmental disorders. *Pediatr. Clin. N. Am.* **2007**, *54*, 469–481. [[CrossRef](#)] [[PubMed](#)]
8. Crystal, D. *Prosodic Systems and Intonation in English*; University Press Archive: Cambridge, UK, 1976; Volume 1.
9. Kalathottukaren, R.T.; Purdy, S.C.; Ballard, E. Behavioral measures to evaluate prosodic skills: A review of assessment tools for children and adults. *Contemp. Issues Commun. Sci. Disord.* **2015**, *42*, 138–154. [[CrossRef](#)]
10. Johnstone, T.; Scherer, K.R. Chapter 14: Vocal communication of emotion. In *The Handbook of Emotion*; Lewis, M., Haviland, J., Eds.; Guilford: New York, NY, USA, 2000; pp. 220–235.
11. Juslin, P.N.; Laukka, P. Impact of intended emotion intensity on cue utilization and decoding accuracy in vocal expression of emotion. *Emotion* **2001**, *1*, 381–412. [[CrossRef](#)] [[PubMed](#)]
12. Murray, I.R.; Arnott, J.L. Toward the simulation of emotion in synthetic speech: A review of the literature on human vocal emotion. *J. Acoust. Soc. Am.* **1993**, *93*, 1097–1108. [[CrossRef](#)] [[PubMed](#)]
13. Peppé, S.J.E. Why is prosody in speech-language pathology so difficult? *Int. J. Speech Lang. Pathol.* **2009**, *11*, 258–271. [[CrossRef](#)]
14. Steinhauer, K.; Alter, K.; Friederici, A.D. Brain potentials indicate immediate use of prosodic cues in natural speech processing. *Nat. Neurosci.* **1999**, *2*, 191–196. [[CrossRef](#)] [[PubMed](#)]
15. Wittman, J.; Van Heuven, V.J.P.; Schiller, N.O. Hearing feelings: A quantitative meta-analysis on the neuroimaging literature of emotional prosody perception. *Neuropsychologia* **2012**, *50*, 2752–2763. [[CrossRef](#)] [[PubMed](#)]
16. Bamiou, D.; Musiek, F.E.; Luxon, L.M. Aetiology and clinical presentations of auditory processing disorders—A review. *Arch. Dis. Child.* **2001**, *85*, 361–365. [[CrossRef](#)]
17. Dawes, P.; Bishop, D. Auditory processing disorder in relation to developmental disorders of language, communication and attention: A review and critique. *Int. J. Lang. Commun. Disord.* **2009**, *44*, 440–465. [[CrossRef](#)] [[PubMed](#)]
18. Haesen, B.; Boets, B.; Wagemans, J. A review of behavioural and electrophysiological studies on auditory processing and speech perception in Autism Spectrum Disorders. *Res. Autism Spectr. Disord.* **2011**, *5*, 701–714. [[CrossRef](#)]
19. O'Connor, K. Auditory processing in Autism Spectrum Disorder: A review. *Neurosci. Biobehav. Rev.* **2012**, *36*, 836–854. [[CrossRef](#)] [[PubMed](#)]
20. Schafer, E.C.; Mathews, L.; Gopal, K.; Canale, E.; Creech, A.; Manning, J.; Kaiser, K. Behavioral auditory processing in children and young adults with Autism Spectrum Disorder. *J. Am. Acad. Audiol.* **2020**. [[CrossRef](#)]

21. Bonnel, A.; Mottron, L.; Peretz, I.; Trudel, M.; Gallun, E.; Bonnel, A. Enhanced pitch sensitivity in individuals with Autism: A signal detection analysis. *J. Cogn. Neurosci.* **2003**, *15*, 226–235. [[CrossRef](#)]
22. Ouimet, T.; Foster, N.E.V.; Tryfon, A.; Hyde, K.L. Auditory-musical processing in Autism Spectrum Disorders: A review of behavioral and brain imaging studies. *Ann. N. Y. Acad. Sci.* **2012**, *1252*, 325–331. [[CrossRef](#)] [[PubMed](#)]
23. Järvinen-Pasley, A.; Heaton, P. Evidence for reduced domain-specificity in auditory processing in Autism. *Dev. Sci.* **2007**, *10*, 786–793. [[CrossRef](#)]
24. Järvinen-Pasley, A.; Peppé, S.J.E.; King-Smith, G.; Heaton, P. The relationship between form and function level receptive prosodic abilities in Autism. *J. Autism Dev. Disord.* **2008**, *38*, 1328–1340. [[CrossRef](#)]
25. Alcántara, J.I.; Weisblatt, E.J.L.; Moore, B.C.J.; Bolton, P.F. Speech-in-noise perception in high-functioning individuals with Autism or Asperger’s syndrome. *J. Child Psychol. Psychiatry* **2004**, *45*, 1107–1114. [[CrossRef](#)]
26. Alcántara, J.I.; Cope, T.E.; Cope, W.; Weisblatt, E.J. Auditory temporal-envelope processing in high-functioning children with Autism Spectrum Disorder. *Neuropsychologia* **2012**, *50*, 1235–1251. [[CrossRef](#)]
27. Tomchek, S.D.; Dunn, W. Sensory processing in children with and without Autism: A comparative study using the short sensory profile. *Am. J. Occup. Ther.* **2007**, *61*, 190–200. [[CrossRef](#)]
28. Carpenter, M.L.; Estrem, T.L.; Crowell, R.L.; Edrisinha, C.D. (Central) auditory processing skills in young adults with Autism Spectrum Disorder. *Commun. Disord. Deaf Stud. Hear. Aids* **2014**, *2*, 1–8. [[CrossRef](#)]
29. DePape, A.R.; Hall, G.B.C.; Tillmann, B.; Trainor, L.J. Auditory processing in high-functioning adolescents with Autism Spectrum Disorder. *PLoS ONE* **2012**, *7*, e44084. [[CrossRef](#)] [[PubMed](#)]
30. Brooks, P.J.; Ploog, B.O. Attention to emotional tone of voice in speech perception in children with Autism. *Res. Autism Spectr. Disord.* **2013**, *7*, 845–857. [[CrossRef](#)]
31. Singh, L.; Harrow, M.S. Influences of semantic and prosodic cues on word repetition and categorization in Autism. *J. Speech Lang. Hear. Res.* **2014**, *57*, 1764–1778. [[CrossRef](#)]
32. Jansson-Verkasalo, E.; Ceponienè, R.; Kielinen, M.; Suominen, K.; Jäntti, V.; Linna, S.L.; Moilanen, I.; Näätänen, R. Deficient auditory processing in children with Asperger syndrome, as indexed by event-related potentials. *Neurosci. Lett.* **2003**, *338*, 197–200. [[CrossRef](#)]
33. Kujala, T.; Kuuluvainen, S.; Saalasti, S.; Jansson-Verkasalo, E.; von Wendt, L.; Lepistö, T. Speech-feature discrimination in children with Asperger syndrome as determined with the multi-feature mismatch negativity paradigm. *Clin. Neurophysiol.* **2010**, *121*, 1410–1419. [[CrossRef](#)] [[PubMed](#)]
34. Lepistö, T.; Nieminen-von Wendt, T.; von Wendt, L.; Näätänen, R.; Kujala, T. Auditory cortical change detection in adults with Asperger syndrome. *Neurosci. Lett.* **2007**, *414*, 136–140. [[CrossRef](#)] [[PubMed](#)]
35. Fan, Y.; Cheng, Y. Atypical mismatch negativity in response to emotional voices in people with Autism Spectrum Conditions. *PLoS ONE* **2014**, *9*, e102471. [[CrossRef](#)] [[PubMed](#)]
36. Korpilahti, P.; Jansson-Verkasalo, E.; Mattila, M.L.; Kuusikko, S.; Suominen, K.; Rytty, S.; Pauls, D.L.; Moilanen, I. Processing of affective speech prosody is impaired in Asperger syndrome. *J. Autism Dev. Disord.* **2006**, *37*, 1539–1549. [[CrossRef](#)] [[PubMed](#)]
37. Kujala, T.; Lepistö, T.; Nieminen-von Wendt, T.; Näätänen, P.; Näätänen, R. Neurophysiological evidence for cortical discrimination impairment of prosody in Asperger syndrome. *Neurosci. Lett.* **2005**, *383*, 260–265. [[CrossRef](#)] [[PubMed](#)]
38. Demopoulos, C.; Hopkins, J.; Kopald, B.E.; Paulson, K.; Doyle, L.; Andrews, W.E.; Lewine, J.D. Deficits in auditory processing contribute to impairments in vocal affect recognition in Autism Spectrum Disorders: A MEG study. *Neuropsychology* **2015**, *29*, 1–27. [[CrossRef](#)] [[PubMed](#)]
39. Lerner, M.; McPartland, J.C.; Morris, J.P. Multimodal emotion processing in Autism Spectrum Disorders: An event-related potential study. *Dev. Cogn. Neurosci.* **2013**, *3*, 11–21. [[CrossRef](#)] [[PubMed](#)]
40. Goydke, K.N.; Altenmüller, E.; Möller, J.; Münte, T.F. Changes in emotional tone and instrumental timbre are reflected by the mismatch negativity. *Cogn. Brain Res.* **2004**, *21*, 351–359. [[CrossRef](#)]
41. Schirmer, A.; Striano, T.; Friederici, A.D. Sex differences in the preattentive processing of vocal emotional expressions. *Cogn. Neurosci. Neuropsychol.* **2005**, *16*, 635–639. [[CrossRef](#)]
42. Schirmer, A.; Escoffier, N. Emotional MMN: Anxiety and heart rate correlate with the ERP signature for auditory change detection. *Clin. Neurophysiol.* **2010**, *121*, 53–59. [[CrossRef](#)] [[PubMed](#)]
43. Csépe, V. On the origin and development of the mismatch negativity. *Ear Hear.* **1995**, *16*, 91–104. [[CrossRef](#)]
44. Ashburner, J.; Ziviani, J.; Rodger, S. Sensory processing and classroom emotional, behavioral, and educational outcomes in children with Autism Spectrum Disorder. *Am. J. Occup. Ther.* **2008**, *62*, 564–573. [[CrossRef](#)] [[PubMed](#)]
45. Landon, J.; Shepherd, D.; Lodhia, V. A qualitative study of noise sensitivity in adults with Autism Spectrum Disorder. *Res. Autism Spectr. Disord.* **2016**, *32*, 43–52. [[CrossRef](#)]
46. Irwin, J.; Preston, J.; Brancazio, L.; D’angelo, M.; Turcios, J. Development of an audiovisual speech perception app for children with Autism Spectrum Disorders. *Clin. Linguist. Phon.* **2015**, *29*, 76–83. [[CrossRef](#)] [[PubMed](#)]
47. Kozou, H.; Azouz, H.G.; Abdou, R.M.; Shaltout, A. Evaluation and remediation of central auditory processing disorders in children with Autism Spectrum Disorders. *Int. J. Pediatr. Otorhinolaryngol.* **2018**, *104*, 36–42. [[CrossRef](#)] [[PubMed](#)]
48. Russo, N.M.; Hornickel, J.; Nicol, T.; Zecker, S.; Kraus, N. Biological changes in auditory function following training in children with Autism Spectrum Disorders. *Behav. Brain Funct.* **2010**, *6*, 60–67. [[CrossRef](#)]

49. Schafer, E.C.; Mathews, L.; Mehta, S.; Hill, M.; Munoz, A.; Bishop, R.; Moloney, M. Personal FM systems for children with Autism Spectrum Disorders: An initial investigation. *J. Commun. Disord.* **2013**, *46*, 30–52. [[CrossRef](#)]
50. Schafer, E.C.; Traber, J.; Layden, P.; Amin, A.; Sanders, K.; Bryant, D.; Baldus, N. Use of wireless technology for children with auditory processing disorders, attention-deficit hyperactivity disorder, and language disorders. *Semin. Hear.* **2014**, *35*, 193–205. [[CrossRef](#)]
51. Schafer, E.C.; Wright, S.; Anderson, C.; Jones, J.; Pitts, K.; Bryant, D.; Wastson, M.; Box, J.; Neve, M.; Mathews, L.; et al. Assistive technology evaluations: Remote-microphone technology for children with Autism Spectrum Disorder. *J. Commun. Disord.* **2016**, *64*, 1–17. [[CrossRef](#)] [[PubMed](#)]
52. Schafer, E.C.; Gopal, K.V.; Mathews, L.; Thompson, S.; Kaiser, K.; McCullough, S.; Jones, J.; Castillo, P.; Canale, E.; Hutcheson, A. Effects of auditory training and remote microphone technology on the behavioral performance of children and young adults who have Autism Spectrum Disorder. *J. Am. Acad. Audiol.* **2019**, *30*, 431–443. [[CrossRef](#)]
53. Rance, G.; Saunders, K.; Carew, P.; Johansson, M.; Tan, J. The use of listening devices to ameliorate auditory deficit in children with Autism. *J. Pediatr.* **2014**, *164*, 352–357. [[CrossRef](#)] [[PubMed](#)]
54. Rance, G.; Chisari, D.; Saunders, K.; Rault, J. Reducing listening-related stress in school-aged children with Autism Spectrum Disorder. *J. Autism Dev. Disord.* **2017**, *47*, 2010–2022. [[CrossRef](#)] [[PubMed](#)]
55. Gopal, K.V.; Schafer, E.C.; Mathews, L.; Nandy, R.; Beaudoin, D.; Schadt, L.; Brown, A.; Phillips, B.; Caldwell, J. Effects of auditory training on electrophysiological measures in individuals with Autism Spectrum Disorder. *J. Am. Acad. Audiol.* **2020**, *31*, 96–104. [[CrossRef](#)] [[PubMed](#)]
56. Eiserman, W.D.; Shisler, L.; Foust, T.; Buhrmann, J.; Winston, R.; White, K.R. Screening for hearing loss in early childhood programs. *Early Child. Res. Q.* **2007**, *22*, 105–117. [[CrossRef](#)]
57. Hall, J.W. Effective and efficient pre-school hearing screening: Essential for successful early hearing detection and intervention (EHDI). *J. Early Hear. Detect. Interv.* **2016**, *1*, 2–12.
58. Gorga, M.P.; Neely, S.T.; Dorn, P.A. Distortion product otoacoustic emission test performance for a priori criteria and for multifrequency audiometric standards. *Ear Hear.* **1999**, *20*, 345–362. [[CrossRef](#)]
59. Ramos, J.A.; Kristensen, S.G.B.; Beck, D.L. An overview of OAEs and normative data for DPOAEs. *Hear. Rev.* **2013**, *20*, 30–33.
60. Schopler, E.; Van Bourgondien, M.E.; Wellman, G.J.; Love, S.R. *Childhood Autism Rating Scale*, 2nd ed.; Western Psychological Services: Los Angeles, CA, USA, 2010.
61. Bishop, D.V.M. *The Children's Communication Checklist-2*; Psychological Corporation: London, UK, 2003.
62. Pearson. *Advanced Clinical Solutions for WAIS-IV and WMS-IV*; Pearson: San Antonio, TX, USA, 2009.
63. Holdnack, J.; Goldstein, G.; Drozdick, L. Social perception and WAIS-IV performance in adolescents and adults diagnosed with Asperger's syndrome and Autism. *Assessment* **2011**, *18*, 192–200. [[CrossRef](#)]
64. Ross, E.D.; Thompson, R.D.; Yenkosky, J. Lateralization of affective prosody in brain and the callosal integration of hemispheric language functions. *Brain Lang.* **1997**, *56*, 27–54. [[CrossRef](#)]
65. Keogh, T.; Kei, J.; Driscoll, C.; Smyth, V. Distortion-product otoacoustic emissions in schoolchildren: Effects of ear asymmetry, handedness, and gender. *J. Am. Acad. Audiol.* **2001**, *12*, 506–513. [[PubMed](#)]
66. Cheour, M.; Korpilahti, P.; Martynova, O.; Lang, A. Mismatch negativity and late discriminative negativity in investigating speech perception and learning in children and infants. *Audiol. Neurotol.* **2001**, *6*, 2–11. [[CrossRef](#)] [[PubMed](#)]
67. Halliday, L.F.; Barry, J.G.; Hardiman, M.J.; Bishop, D.V. Late, not early mismatch responses to changes in frequency are reduced or deviant in children with dyslexia: An event-related potential study. *J. Neurodev. Disord.* **2014**, *6*, 21–35. [[CrossRef](#)] [[PubMed](#)]
68. Korpilahti, P.; Krause, C.M.; Holopainen, I.; Lang, A.H. Early and late mismatch negativity elicited by words and speech-like stimuli in children. *Brain Lang.* **2001**, *76*, 332–339. [[CrossRef](#)] [[PubMed](#)]
69. Golan, O.; Ashwin, E.; Granader, Y.; McClintock, S.; Day, K.; Leggett, V.; Baron-Cohen, S. Enhancing emotion recognition in children with Autism Spectrum Conditions: An intervention using animated vehicles with real emotional faces. *J. Autism Dev. Disord.* **2010**, *40*, 269–279. [[CrossRef](#)]
70. Lacava, P.G.; Golan, O.; Baron-Cohen, S.; Smith-Myles, B. Using assistive technology to teach emotion recognition to students with Asperger syndrome. *Remedial Spec. Educ.* **2007**, *28*, 174–181. [[CrossRef](#)]
71. Lacava, P.G.; Rankin, A.; Mahlios, E.; Cook, K.; Simpson, R.L. A single case design evaluation of a software and tutor intervention addressing emotion recognition and social interaction in four boys with ASD. *Autism Int. J. Res. Pract.* **2010**, *14*, 161–178. [[CrossRef](#)] [[PubMed](#)]
72. Silver, M.; Oakes, P. Evaluation of a new computer intervention to teach people with Autism or Asperger syndrome to recognize and predict emotions in others. *Autism Int. J. Res. Pract.* **2001**, *5*, 299–316. [[CrossRef](#)]
73. Matsuda, S.; Yamamoto, J. Intervention for increasing the comprehension of affective prosody in children with Autism Spectrum Disorders. *Res. Autism Spectr. Disord.* **2013**, *7*, 938–946. [[CrossRef](#)]
74. Ploog, B.O.; Banerjee, S.; Brooks, P.J. Attention to prosody (intonation) and content in children with Autism and in typical children using spoken sentences in a computer game. *Res. Autism Spectr. Disord.* **2009**, *3*, 743–758. [[CrossRef](#)]
75. Holbrook, S.; Israelsen, M. Speech prosody interventions for persons with Autism Spectrum Disorders: A systematic review. *Am. J. Speech-Lang. Pathol.* **2020**, *29*, 2189–2205. [[CrossRef](#)] [[PubMed](#)]
76. Didehbani, N.; Allen, T.; Kandalaf, M.; Krawczyk, D.; Chapman, S. Virtual reality social cognition training for children with high functioning Autism. *Comput. Hum. Behav.* **2016**, *62*, 703–711. [[CrossRef](#)]

77. Golan, O.; Baron-Cohen, S.; Hill, J.J.; Golan, Y. The reading the mind in films task: Complex emotion recognition in adults with and without Autism Spectrum Conditions. *Soc. Neurosci.* **2006**, *1*, 111–123. [[CrossRef](#)] [[PubMed](#)]
78. Kandalafi, M.R.; Didehbani, N.; Krawczyk, D.C.; Allen, T.T.; Chapman, S.B. Virtual reality social cognition training for young adults with high-functioning Autism. *J. Autism Dev. Disord.* **2013**, *43*, 34–44. [[CrossRef](#)]
79. Ferreira, D.A.; Bueno, C.D.; De Costa, S.S.; Sleifer, P. Mismatch negativity in children: Reference values. *Int. Arch. Otorhinolaryngol.* **2019**, *23*, 142–146. [[CrossRef](#)]
80. Gomot, M.; Giard, M.; Roux, S.; Barthélemy, C.; Bruneau, N. Maturation of frontal and temporal components of mismatch negativity (MMN) in children. *Neuroreport* **2000**, *11*, 3109–3112. [[CrossRef](#)] [[PubMed](#)]
81. Shafer, V.L.; Morr, M.L.; Kreuzer, J.A.; Kurtzberg, D. Maturation of mismatch negativity in school-age children. *Ear Hear.* **2000**, *21*, 242–251. [[CrossRef](#)] [[PubMed](#)]
82. Cheng, Y.; Wu, H.; Tzeng, Y.; Yang, M.; Zhao, L.; Lee, C. The development of mismatch responses to mandarin lexical tones in early infancy. *Dev. Neuropsychol.* **2013**, *38*, 281–300. [[CrossRef](#)] [[PubMed](#)]
83. Shafer, V.L.; Yu, Y.H.; Datta, H. Maturation of speech discrimination in 4- to 7-yr-old children as indexed by event-related potential mismatch responses. *Ear Hear.* **2010**, *31*, 735–745. [[CrossRef](#)] [[PubMed](#)]
84. Tremblay, K.; Kraus, N.; McGee, T.; Ponton, C.; Otis, B. Central auditory plasticity: Changes in the N1-P2 complex after speech-sound training. *Ear Hear.* **2001**, *22*, 79–90. [[CrossRef](#)] [[PubMed](#)]
85. Tremblay, K.L.; Kraus, N. Auditory training induces asymmetrical changes in cortical neural activity. *J. Speech Lang. Hear. Res.* **2002**, *45*, 564–572. [[CrossRef](#)]
86. Kraus, N.; McGee, T.; Carrell, T.; Sharma, A. Neurophysiologic bases of speech discrimination. *Ear Hear.* **1995**, *16*, 19–37. [[CrossRef](#)]
87. Tong, Y.; Melara, R.D.; Rao, A. P2 enhancement from auditory discrimination training is associated with improved reaction times. *Brain Res.* **2009**, *1297*, 80–88. [[CrossRef](#)] [[PubMed](#)]
88. Tremblay, K.L.; Ross, B.; Inoue, K.; McClannahan, K.; Collet, G. Is the auditory evoked P2 response a biomarker of learning? *Front. Syst. Neurosci.* **2014**, *8*, 1–13. [[CrossRef](#)] [[PubMed](#)]
89. Banse, R.; Scherer, K.R. Acoustic profiles in vocal emotion expression. *J. Personal. Soc. Psychol.* **1996**, *70*, 614–636. [[CrossRef](#)]
90. Yildirim, S.; Bulut, M.; Lee, C.M.; Kazemzadeh, A.; Deng, Z.; Lee, S.; Narayanan, S.; Busso, C. An acoustic study of emotions expressed in speech. In Proceedings of the Eighth International Conference on Spoken Language Processing, Jeju Island, Korea, 4–8 October 2004; pp. 1–4.
91. Belyk, M.; Brown, S. The acoustic correlates of valence depend on emotion family. *J. Voice* **2014**, *28*, 523.e9–523.e18. [[CrossRef](#)]
92. Peppé, S.J.E.; McCann, J.; Gibbon, F.; O'Hare, A.; Rutherford, M. Receptive and expressive prosodic ability in children with high-functioning Autism. *J. Speech Lang. Hear. Res.* **2007**, *50*, 1015–1028. [[CrossRef](#)]
93. Cisler, J.; Bacon, A.; Williams, N. Phenomenological characteristics of attentional biases towards threat: A critical review. *Cogn. Ther. Res.* **2009**, *33*, 221–234. [[CrossRef](#)] [[PubMed](#)]
94. Green, H.L.; Shuffrey, L.C.; Levinson, L.; Shen, G.; Avery, T.; Randazzo-Wagner, M.; Sepulveda, D.M.; Garcia, P.; Maddox, C.; Garcia, F.; et al. Evaluation of mismatch negativity as a marker for language impairment in Autism Spectrum Disorder. *J. Commun. Disord.* **2020**, *87*, 1–15. [[CrossRef](#)] [[PubMed](#)]
95. Baranek, G.T.; David, F.J.; Poe, M.D.; Stone, W.L.; Watson, L.R. Sensory Experiences Questionnaire: Discriminating sensory features in young children with autism, developmental delays, and typical development. *J. Child Psychol. Psychiatry* **2006**, *47*, 591–601. [[CrossRef](#)]
96. Harlor, A.D.B.; Bower, C. Hearing assessment in infants and children: Recommendations beyond neonatal screening. *Pediatrics* **2009**, *124*, 1252–1263. [[CrossRef](#)]
97. Cupples, L.; Ching, T.Y.C.; Crowe, K.; Seeto, M.; Leigh, G.; Street, L.; Day, J.; Marnane, V.; Thomson, J. Outcomes of 3-year-old children with hearing loss and different types of additional disabilities. *J. Deaf Stud. Deaf Educ.* **2013**, *19*, 20–39. [[CrossRef](#)]
98. Demopoulos, C.; Lewine, J.D. Audiometric profiles in Autism Spectrum Disorders: Does subclinical hearing loss impact communication? *Autism Res.* **2016**, *9*, 107–120. [[CrossRef](#)] [[PubMed](#)]
99. Peppé, S.J.E.; McCann, J.; Gibbon, F.; O'Hare, A.; Rutherford, M. Assessing prosodic and pragmatic ability in children with high-functioning Autism. *J. Pragmat.* **2006**, *38*, 1776–1791. [[CrossRef](#)]