Psychoacoustic abilities as predictors of vocal emotion recognition in autism

Eitan Globerson[1, 2], *Noam Amir*[3], *Michal Lavidor*[1,4], *Liat Kishon-Rabin*[3], *Ofer Golan*[4]

[1]Gonda Multidisciplinary Brain Research Center, Bar-Ilan University, Ramat-Gan, Israel
[2]Jerusalem Academy of Music and Dance, Israel
[3]Department of Communication Disorders, The Sackler Faculty of Medicine, Tel-Aviv University, Israel
[4]Department of Psychology, Bar-Ilan University, Ramat-Gan, Israel

Abstract

Individuals with Autism Spectrum Disorders (ASD) show deficiencies in prosodic abilities, both pragmatic and affective. Their deficiencies in affective prosody have been mostly related to cognitive difficulties in emotion recognition. The current study tested an alternative hypothesis, linking vocal emotion recognition difficulties in ASD to lower level auditory perceptual deficiencies. Twenty one high functioning male adults with ASD and 32 male adults from the general population, matched on age and verbal abilities, and screened for normal hearing thresholds, undertook a battery of auditory tasks. Results demonstrated that individuals with ASD scored significantly lower than controls on vocal emotion recognition. Psychoacoustic abilities were strong predictors of vocal emotion recognition in both the ASD and control groups. Psychoacoustic abilities explained 48.1% of the variance of vocal emotion recognition scores in the ASD group, and 28.0% of the variance in the general population group. These results highlight the importance of lower level psychoacoustic factors in the perception of prosody in autism.
Index Terms: prosody, autism, psychoacoustic abilities

1. Introduction

Prosodic impairments have been reported extensively in Autism Spectrum Disorders (ASD). Most of these reports focused on prosodic production [1], referring to a number of abnormal features in ASD, including monotonic intonation, uncontrolled loudness of speech, deficiencies in voice quality, and abnormal stress patterns [2]. Fewer studies focused on prosodic **perception** in ASD [3]. Such studies showed impairments in the perception of both affective [4] and pragmatic prosody [2] in ASD. It has therefore been commonly argued that affective prosody recognition difficulties in ASD are related to a more general emotion-recognition deficit [5]. According to this approach, a hampered higher-cognitive mechanism affects all modalities of emotion recognition, including the auditory domain. Supportive of this approach, are findings demonstrating deficiencies in different modalities of emotion recognition [6, 7]. However, although social and communicative abnormalities have been shown to correlate with prosodic deficiencies in ASD, this correlation is weak and cannot fully account for prosodic deficiencies [8]. Hence, it is essential to search for additional factors underlying deficient prosody in ASD. The search for such a mechanism may encompass both sensory and cognitive domains. The sensory domain would include auditory perceptual abilities involved in the perception of prosody. Abnormalities in such sensory mechanisms could influence the ability of individuals with ASD to perceive prosodic cues, thus affecting their ability to interpret them. Supportive of this approach are findings indicating abnormal auditory brain mechanisms in autism. Gomot, Giard, Adrien, Barthelemy, and Bruneau (2002) reported abnormal pattern of brain responses to deviant auditory stimuli in ASD. Using the mismatch negativity (MMN) paradigm, they demonstrated that the neural response to a deviant frequency (1010 Hz tones, as deviant stimuli in a 1000 Hz tone series) exhibits an earlier peak in ASD participants, compared with controls. Furthermore, they found an atypical left frontal activity in children with ASD in response to deviant auditory stimuli [9]. This difference in brain responses may imply that individuals with ASD utilize unique brain mechanisms in analyzing pitch differences. Due to the major role of pitch as cue to interpreting affective and pragmatic prosody recognition [10, 11], it is possible that abnormal brain responses to pitch changes in ASD, could hamper the ability of this population to utilize prosody as a communicative tool.

Recently, we reported a strong association between various pitch perception abilities and the perception of pragmatic and affective prosody in the general population[12]. More specifically, it was found that thresholds for pitch direction recognition (high/low) tasks are stronger predictors of vocal emotion recognition than tasks which only require pitch discrimination (identical/different). It would therefore be interesting to examine whether the same associations are found in ASD. The current study aimed to explore these associations. A test battery of two psychoacoustic tasks and a vocal emotion recognition task were employed. The psychoacoustic tests included a discrimination task, and a pitch direction (high/low) recognition task. Our purpose in constructing this experimental design was to examine the roles of both pitch discrimination and pitch direction recognition in vocal emotion recognition in ASD.

2. Methods and participants

2.1 Participants

Two groups took part in the study:

1. <u>ASD group</u>: A group of 23 male participants, aged 20-40, diagnosed with ASD in specialist centers, using established criteria [13]. Participants with ASD were recruited from national residential and occupational programs for high functioning individuals with ASD. Participants filled out the Autism Spectrum Quotient (AQ) [14] to assess their self reported level of autistic traits. Eight participants did not meet the adapted cutoff level of 26 [15]. For these participants diagnosis was confirmed using the Autism Diagnostic Observation Schedule (ADOS) [16]. Using the ADOS, seven out of eight participants met criteria for ASD on social interaction, communication, and the combination of the two. One participant did not meet criteria for ASD on the ADOS. Another participant did not complete all the required assignments. The data of these two participants was therefore excluded from the analysis.
2. <u>Typical Development (TD) group</u>: a group of 32 male participants, aged 23-39,with no reported neurological or psychiatric history, volunteered to

take part in the study. They were screened out for ASD using the AQ. None of the participants reached cutoff.

The two groups were matched on chronological age. In addition, groups were matched on verbal abilities using the vocabulary task from the third edition of the Wechsler Adult Intelligence Scale [17].

2.2 Experimental design

Psychoacoustic tasks

Two psychoacoustic tasks and a three-block Vocal Emotion Recognition (VER) task were presented to the participants binaurally with Sennheiser HD-201 linear headphones and a Line 6 Ux1 external sound card. The order of the assignments was randomly modified between subjects to avoid an order effect. Participants were allowed to take as many breaks as needed between the various tasks. On average, participants took 60 minutes to complete all tasks.

The following psychoacoustic tasks were employed:

a. Pitch Direction Recognition task (PDR): In this task, participants were presented with two constant frequency pure tones (PT) in each trial. One of the tones was a 1 kHz PT, and the other had a larger frequency value (i.e., higher in pitch). Participants had to indicate the tone that was higher. The initial frequency difference between the tones was 200 Hz.

b. Pitch Discrimination (PD) task: In this task, a three-interval, 2-alternative choice paradigm was employed. A fixed reference PT of 1 kHz was followed by two other tones, one of which was a repetition of the reference and the other was higher in pitch. Participants had to indicate which tone (2nd or 3rd) was different from the reference tone. The initial frequency difference between the deviant stimulus and the other two stimuli was 200 Hz.

The thresholds of the two psychoacoustic tasks were obtained using a 2 down 1 up adaptive staircase procedure, converging at a performance level of 70.7%[18]. The initial step size was 40 Hz. The step size was divided by 2 after each reversal, until a final step size of 1 Hz was reached. Assessment was terminated after 10 reversals with the final step size. Thresholds were calculated using the arithmetic mean of the last 8 reversals.

All stimuli were 300 ms in duration, (including rise and fall ramps) with a 500 ms inter-stimulus gap.

All tests were preceded by a training session. Participants had to achieve 5 subsequent successes in the relevant tasks in order to begin the experiment.

Vocal emotion recognition (VER) task

The VER task tested for the recognition of four basic emotions: happiness, sadness, anger and fear. A "neutral" option was given to test the sensitivity of the participants to emotional utterances. The task employed in the current study is a further development of a similar task which was used in our previous experiment with the general population [12]. Following the previous experiment, we re-inspected the scores for the individual stimuli obtained by the 60 participants who took part in that experiment. In order to make the task balanced in the number of emotions per stimulus category, we eliminated the stimuli which obtained the highest number of votes from the participants in the prior experiment, thus reaching an equal number of 16 stimuli for each emotion, in each stimulus category (words, non-words, sentences). By eliminating the highest scored stimuli, the task was rendered more demanding, thus enhancing its ability to discriminate between the ASD and TD groups. The new task therefore comprised 48 stimuli for each emotion, reaching a total number of 192 stimuli. The task was divided into blocks, characterized by the stimulus category. The order of the blocks was randomized between subjects, as well as the order of the stimuli within the blocks.

3. Results

3.1 Calculated scores

A total of four scores was obtained for each participant, as follows:

1. The Vocal Emotion Recognition (VER) task score (in percent correct): the overall percentage of correct responses in the VER task. Any incorrect emotion labeling, including the "Neutral" responses, were counted as errors.
2. The VER task "neutral score": the percentage of stimuli rated as "neutral" by the participant in the voice emotion recognition task.
3. Pitch Discrimination (PD) task threshold.
4. Pitch Direction Recognition task (PDR) threshold.

Thresholds for both psychoacoustic tasks were log-transformed for analysis. The employment of a logarithmic transformation of frequency discrimination thresholds is common practice in psychoacoustic literature [19, 20]

3.2 Group comparison

Psychoacoustic tasks

In order to check for group differences on psychoacoustic tasks a multivariate analysis of variance was conducted, with the two psychoacoustic task scores as dependent measures. The multivariate analysis ($F[2,50]=.149$, $p>.05$.) as well as each measure's between-group analyses showed no significant differences between the ASD and TD groups. Table 1 summarizes group descriptive statistics, as well as the results of each dependent psychoacoustic measure.

Table 1: Descriptive statistics (mean, SD, and range), and F values for psychoacoustic tasks:

Task	TD	ASD	F(1,51)
PDR(Hz)	45.4(70.0),3-200	42.8(49.9),1.5-175.7	.01
PD(Hz)	21.5(41.4),3.8-192.6	24.6(31.1),2.1-109	.05

Vocal Emotion recognition scores

An independent-samples t-test was conducted to compare VER task scores and the VER neutral scores between the TD and ASD groups. There was a significant difference in the VER scores, but not for the VER neutral scores (Table 2)."

Table 2: Descriptive statistics (mean, SD, and range), and F values for the VER task scores and neutral scores

Task	TD	ASD	t(1,52)
VER	76.6(10.0)55.2-93.2	68.2(10.9)46.3-80.2	2.87**
VER NS	12.7(7.2), 0-26.6	15.2(5.2), 4.2-28.1	-1.490

**p<.01

Abbreviations:

VER: vocal emotion recognition

VER NS: vocal emotion recognition neutral score

Recognition of particular emotions was examined through a multivariate analysis of variance, with scores for individual emotions (i.e., happiness, sadness, anger and fear) as dependent measures and group as the independent factor. The overall analysis yielded significant results ($F[4,48]=2.99$, $p<.05$.), with the TD group performing better than the ASD group on the overall analysis. Univariate ANOVAs yielded significant group differences for the recognition of happiness and sadness. It is noteworthy that both groups demonstrated the same order of recognition rates for the expressed emotions(Table 3), implying emotion recognition in the ASD group was systematic and not random.

Table 3 : scores for individual emotion recognition in the vocal emotion recognition task(mean, SD, and range), and F values.

Emotion	TD	ASD	F(1,52)
Happiness	78.0 (15.9)29.2-95.8	66.4(16.3)27.1-89.6	6.72*
Sadness	75.4(13.0)43.8-90.8	64.2(17.4)25.0-93.7	7.2*
Fear	82.9(13.7)35.4-100	78.7(17.6)29.2-97.9	.98
Anger	69.9(15.1)42.0-100	63.6(14.8)25.0-83.0	2.3

P<.05

3.3 Bivariate correlations

Table 5 illustrates the correlation between the scores for the VER task and psychoacoustic thresholds. All correlation coefficients had a negative value, indicating that better psychoacoustic abilities correspond with better VER abilities. A significant correlation was found between the pitch direction recognition (PDR) task and vocal emotion recognition abilities in both groups. This correlation coefficient had a larger value for the ASD group. A significant correlation between thresholds obtained in the pitch discrimination task (PD) and vocal emotion recognition was found **only** in the ASD group.

Table 5: Correlation between vocal emotion recognition scores and psychoacoustic thresholds, illustrated separately for the ASD and TD groups

	Vocal emotion recognition	PDR	PD
Vocal emotion recognition		-.529*	-.348
PDR	-.693**		.599***
PD	-.655**	.660**	

ASD
TD

* Correlation is significant at the 0.05 level (Bonferroni adjustment for multiple correlations, critical p level=0.025)

** Correlation is significant at the 0.01 level (Bonferroni adjustment for multiple correlations, critical p level=0.005)

*** Correlation is significant at the 0.001 level (Bonferroni adjustment for multiple correlations, critical p level=0.0005)

Abbreviations:

PDR: Pitch direction recognition task

PD: pitch discrimination task

3.4 Linear regression

In order to predict the VER task scores, a hierarchic linear regression was performed. Group was entered into the regression in the first block, followed by Pitch direction recognition (PDR), and Pitch discrimination (PD), which were entered in the second block, using stepwise method. Finally, group by PDR, and group by PD (products of z-scores of group and scores for the relevant tasks) were entered into the regression, using stepwise method. In the joint group (TD+ASD), psychoacoustic abilities explained 31% of the variance in the data. It is important to note that since group interactions did not significantly predict vocal emotion recognition, any difference in the associations between vocal emotion recognition and psychoacoustic abilities found between the groups does not significantly support an assumption of between group differences. However, it may point in such a direction, which will have to be supported by further data obtained with additional participants.

4. Discussion

The current study explored the association between psychoacoustic abilities and vocal emotion recognition in a group of individuals with ASD and a matched group of individuals with typical development (TD). Compared with the TD group, performance in the ASD group was significantly poorer in vocal emotion recognition. This supports prior findings indicating deficient emotion recognition in ASD [21]. Our findings demonstrate a significant association between auditory perceptual abilities and the ability to perceive affective messages in vocal communication, in both groups. This association was not affected by group, indicating no significant differences in the association between prosodic perception and psychoacoustic abilities between the TD and the ASD groups. Taken together, results may indicate that the group differences in vocal emotion recognition abilities derive from deficits in higher level emotion recognition mechanisms, rather than sensory abnormalities. Nevertheless, the association between pitch discrimination abilities and vocal emotion recognition appears to be more pronounced in the ASD group than in the TD group. This finding may suggest that psychoacoustic abilities play a prominent role in the ability of individuals with ASD to interpret affective prosody.

The correlations between psychoacoustic thresholds in the TD group showed the same pattern as in our previous study [12], namely, that the thresholds for the task in which pitch direction recognition (high/low) was required, correlated significantly with prosody recognition, while the thresholds for the purely discriminative task did not demonstrate such a pronounced association. This pattern of association can be explained by the importance of pitch direction recognition to the perception of prosody. For example, while fear is associated with an increase in the fundamental frequency (F0) of the speech signal, sadness is associated with a decrease in F0 [22]. Therefore, pitch discrimination abilities alone, will not suffice in differentiating between these emotions. It is a combined ability of pitch discrimination and pitch direction recognition which enables the recognition and differentiation of emotions in vocal communication. In the ASD group, however, a significant correlation between psychoacoustic thresholds was found for **both** psychoacoustic tasks, requiring either only pitch discrimination, or both discrimination and pitch direction recognition. The different pattern of associations found in the ASD group could indicate that in the case of ASD, psychoacoustic factors contributing to prosodic deficiencies are more general. It is plausible, for example, that auditory attentional factors play an important role in prosody perception in ASD. Deficient auditory attention may disrupt the ability of individuals with ASD to fully perceive and process certain acoustic cues, thus obstructing vocal emotion recognition. This view is supported by prior studies indicating general auditory attention deficits in autism [23].

5. Conclusion

The current study presents a first attempt to examine the association between psychoacoustic abilities and vocal emotion recognition in ASD. A significant association was found between psychoacoustic abilities and vocal emotion recognition in a group of individuals with typical development and individuals with ASD. This association was more pronounced in the ASD group.

6. Acknowledgements

We are grateful to the National Institute of Psychobiology in Israel for their support of this study. We thank the participants in the research for their time and effort.

7. References

[1] Fay, W.H. and A.L. Schuler, *Emerging language in autistic children*. 1980, London: Edward Arnold. xiii,216p.

[2] Paul, R., et al., *Perception and production of prosody by speakers with autism spectrum disorders.* J Autism Dev Disord, 2005. **35**(2): p. 205-20.

[3] McCann, J. and S. Peppe, *Prosody in autism spectrum disorders: a critical review.* Int J Lang Commun Disord, 2003. **38**(4): p. 325-50.

[4] Rutherford, M.D., S. Baron Cohen, and S. Wheelwright, *Reading the mind in the voice: a study with normal adults and adults with Asperger syndrome and high functioning autism.* J Autism Dev Disord, 2002. **32**(3): p. 189-94.

[5] Hobson, R.P., *The cradle of thought*. 2002, London: Macmillan.

[6] Golan, O., et al., *The 'Reading the Mind in the Voice' test - Revised: A study of complex emotion recognition in adults with and without Autism Spectrum Conditions.* Journal of Autism and Developmental Disorders, 2007. **37**(6): p. 1096-1106.

[7] Golan, O., et al., *The "reading the mind in films" task: complex emotion recognition in adults with and without autism spectrum conditions.* Soc Neurosci, 2006. **1**(2): p. 111-23.

[8] Paul, R., et al., *Brief report: relations between prosodic performance and communication and socialization ratings in high functioning speakers with autism spectrum disorders.* J Autism Dev Disord, 2005. **35**(6): p. 861-9.

[9] Gomot, M., et al., *Hypersensitivity to acoustic change in children with autism: electrophysiological evidence of left frontal cortex dysfunctioning.* Psychophysiology, 2002. **39**(5): p. 577-84.

[10] Amir, N., B.C. Almogi, and H. Mixdorff *A systematic framework for studying the contribution of F0 and duration to the perception of accented words in Hebrew.* Proceedings of the Speech Prosody 4th international conference, Campinas, Brazil, 6-9 May 2008, 2008: p. 285-288.

[11] Hammerschmidt, K. and U. Jurgens, *Acoustical correlates of affective prosody.* J Voice, 2007. **21**(5): p. 531-40.

[12] Globerson, E., et al. *Psychoacoustic Abilities Can Predict Vocal Emotion Recognition*. in *Speech Prosody*. 2010. Chicago, IL.

[13] APA, *Diagnostic and statistical manual of mental disorders : DSM-IV-TR.* 4th ed., text revision. ed. 2000, Washington, DC: APA. xxxvii, 943 p.

[14] Baron Cohen, S., et al., *The autism-spectrum quotient (AQ): evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians.* J Autism Dev Disord, 2001. **31**(1): p. 5-17.

[15] Woodbury Smith, M.R., et al., *Screening adults for Asperger Syndrome using the AQ: a preliminary study of its diagnostic validity in clinical practice.* J Autism Dev Disord, 2005. **35**(3): p. 331-5.

[16] Lord, C., et al., *The autism diagnostic observation schedule-generic: a standard measure of social and communication deficits associated with the spectrum of autism.* J Autism Dev Disord, 2000. **30**(3): p. 205-23.

[17] Wechsler, D., *Manual for the Wechsler Adult Intelligence Scale-Third Edition (WAIS-III)*. 1997, San Antonio, TX: The Psychological Corporation.

[18] Levitt, H., *Transformed up-down methods in psychoacoustics.* J Acoust Soc Am, 1971. **49**(2): p. Suppl 2:467+.

[19] Irvine, D.R., et al., *Specificity of perceptual learning in a frequency discrimination task.* J Acoust Soc Am, 2000. **108**(6): p. 2964-8.

[20] Micheyl, C., et al., *Influence of musical and psychoacoustical training on pitch discrimination.* Hear Res, 2006. **219**(1-2): p. 36-47.

[21] 21. Golan, O., S. Baron Cohen, and J. Hill, *The Cambridge Mindreading (CAM) Face-Voice Battery: Testing complex emotion recognition in adults with and without Asperger syndrome.* J Autism Dev Disord, 2006. **36**(2): p. 169-83.

[22] Banse, R. and K.R. Scherer, *Acoustic profiles in vocal emotion expression.* J Pers Soc Psychol, 1996. **70**(3): p. 614-36.

[23] Teder-Salejarvi, W.A., et al., *Auditory spatial localization and attention deficits in autistic adults.* Cognitive Brain Research, 2005. **23**(2-3): p. 221-34.