

SCIENTIFIC REPORTS



OPEN

Enhanced segregation of concurrent sounds with similar spectral uncertainties in individuals with autism spectrum disorder

Received: 17 November 2014

Accepted: 16 April 2015

Published: 22 May 2015

I-Fan Lin¹, Takashi Yamada^{2,3}, Yoko Komine², Nobumasa Kato² & Makio Kashino^{1,4,5}

When acoustic signals from different sound sources are mixed upon arrival at the ears, the auditory system organizes these acoustic elements by their features. This study shows that individuals with autism spectrum disorder (ASD) performed better in terms of hearing a target sequence among distractors that had similar spectral spectral uncertainties. Their superior performance in this task indicates an enhanced discrimination between auditory streams with the same spectral uncertainties but different spectro-temporal details. The enhanced discrimination of acoustic components may be related to the absence of the automatic grouping of acoustic components with the same features, which results in difficulties in speech perception in a noisy environment. On the other hand, the ASD group and the control group had similar performance in hearing a target sequence among distractors that had different spatial cues defined by interaural intensity differences.

Sounds help us know our environment, and the ability to differentiate sounds generated by different sources is not trivial. In a social environment, we usually encounter difficulties when trying to separate target speech from competitive speech. Autism spectrum disorder (ASD) is a neurodevelopmental disorder, and it has been suggested that the impaired social interaction and communication of individuals with ASD can be partly explained by the difficulties they experience in understanding target speech in a noisy environment^{1,2}. Previous studies have shown that their difficulties in segregating the auditory target from distractors may be due to their deficit in hearing the target in the temporal dips of the distractors^{1,2} and in segregating auditory streams based on frequency separation³ and different spatial cues^{4–6}. Here we investigated whether individuals with ASD had difficulty in integrating acoustic elements with similar features or in segregating acoustic elements with different features, and the features examined in this study were spectral uncertainties (i.e., jittered frequencies vs. non-jittered frequencies) and spatial cues (specifically, interaural intensity differences). In the auditory system, unwanted or ‘masking’ sounds compete with desired or ‘target’ sounds at various levels. The terms ‘energetic masking’ and ‘informational masking’ have been coined to describe the disruption of the auditory presentation of target sounds caused by masking sounds⁷. Whereas ‘energetic masking’ refers to the disruption in the auditory periphery, ‘informational masking’ refers to the disruption in any of several stages of processing beyond the auditory periphery. This study focused on auditory processes above the peripheral level to avoid energetic masking but below language-related processes to avoid the different orientation behaviors toward speech

¹NTT Communication Science Laboratories, NTT Corporation, Atsugi, Kanagawa, Japan. ²Medical Institute of Developmental Disabilities Research, Showa University, Tokyo, Japan. ³ATR Brain Information Communication Research Laboratory Group, Kyoto, Japan. ⁴Department of Information Processing, Interdisciplinary Graduate School of Science and Engineering, Tokyo Institute of Technology, Yokohama, Kanagawa, Japan. ⁵CREST, JST, Atsugi, Kanagawa, Japan. Correspondence and requests for materials should be addressed to I.F.L. (email: lin.ifan@lab.ntt.co.jp) or M.K. (email: kashino.makio@lab.ntt.co.jp)

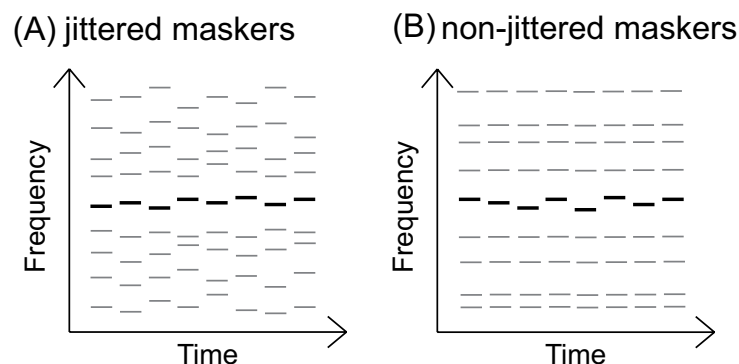


Figure 1. The auditory stimuli contained one target sequence (black lines) and eight masker sequences (gray lines). The target sequence always had jittered frequencies within a fixed protected region. **(A)** The masker sequences had jittered frequencies outside the protected region for the jittered conditions. **(B)** The masker sequences had fixed frequencies outside the protected region for the non-jittered conditions.

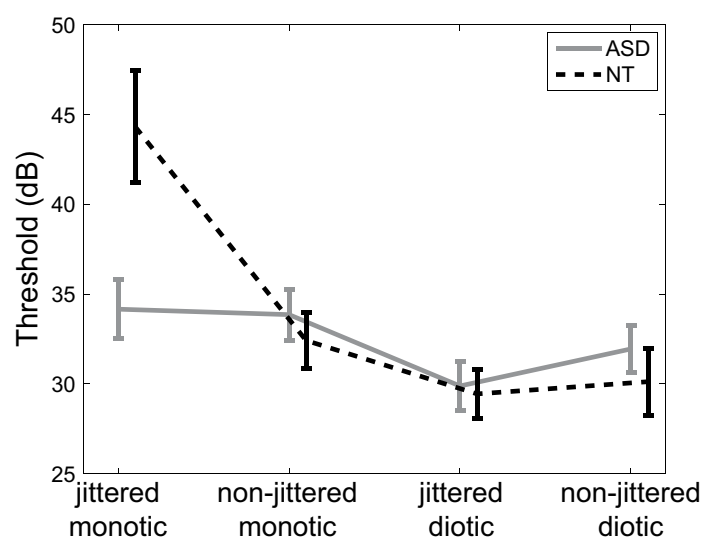


Figure 2. Target detection thresholds in the ASD and control groups for the four conditions, indicated as mean $\pm$ standard error. The maskers were sent to the right ear for the monotonic conditions or to both ears for the diotic conditions.

found in ASD^{8,9}. The target sounds and interfering sounds (maskers) used in the experiment were composed of spectrally separated pure tones.

In the experiment, the threshold for detecting the target among the maskers was measured by using an adaptive staircase procedure with a 3-down 1-up rule. The target sequence always had jittered frequencies and was sent only to the right ear. The masking sequences either had jittered frequencies (i.e., the jittered conditions shown in Fig. 1A) or had the same frequencies (i.e., the non-jittered conditions shown in Fig. 1B). For the monotic conditions, the maskers were sent to the right ear, and the target and maskers shared the same spatial cues. For the diotic conditions, the maskers were sent to both ears, and the target and maskers had different spatial cues. Listeners should perform better under the non-jittered conditions than under the jittered conditions if they can segregate acoustic signals carrying different spectral uncertainties, and they should perform better under the diotic conditions than under the monotic conditions if they can segregate acoustic signals carrying different spatial cues. On the other hand, if the participants with ASD did not group acoustic signals with similar spectral uncertainties automatically, their performance for the jittered condition should be better than that of the control group.

Results

The cross-subject thresholds in the ASD and control groups for the four conditions are shown in Fig. 2. To investigate whether target-masker similarities in spectral uncertainty and the difference between

spatial cues influenced the ASD subjects and neurotypical (NT) subjects in the same way, a three-way mixed-design ANOVA, with the between-subject factor as *Group* (ASD and NT) and the within-subject factor as *Spectral Uncertainty* (jittered and non-jittered frequencies) and *Spatial Cues* (monotic and diotic), was performed for the averaged thresholds in each condition. The ANOVA analysis was conducted with SPSS v.19 (IBM, USA).

In the ANOVA analysis, a significant effect of *Spectral Uncertainty* indicates masking release from target-masker dissimilarity with lower thresholds for the non-jittered condition compared to the jittered condition. On the other hand, a significant effect of *Spatial Cues* indicates masking release from the difference between spatial cues with lower thresholds for the diotic condition compared to the monotic condition. The ANOVA revealed a significant effect of *Spectral Uncertainty* ($F(1,24) = 5.787$, $p = 0.024$) and *Spatial Cues* ($F(1,24) = 40.508$, $p < 0.001$) but not *Group* ($F(1,24) = 0.658$, $p = 0.425$), and significant interactions between *Spectral Uncertainty* and *Group* ($F(1,24) = 10.898$, $p = 0.003$), *Spatial Cues* and *Group* ($F(1,24) = 8.994$, $p = 0.006$), and *Spectral Uncertainty* and *Spatial Cues* ($F(1,24) = 21.549$, $p < 0.001$) and among *Spectral Uncertainty* and *Spatial Cues* and *Group* ($F(1,24) = 10.073$, $p = 0.004$). The follow-up ANOVAs showed that the factor *Spectral Uncertainty* and the interaction between *Spectral Uncertainty* and *Group* was significant for the monotic conditions but not for the diotic conditions (while *Group* was not a significant factor for either monotic or diotic conditions), and the factors *Spatial Cues* and *Group* and the interaction between them were all significant for the jittered conditions but the factor *Spatial Cues* was the only significant one for the non-jittered conditions.

Driven by the hypothesis that the significant interaction between *Spectral Uncertainty* and *Group* for the monotic conditions and the significant interaction between *Spatial Cues* and *Group* for the jittered conditions were caused by the significantly different thresholds between groups for the jittered monotic condition, 4 between-group t-tests were conducted for the four conditions. These results confirmed that the significant difference between the performance observed in the ASD group and NT group mainly attributed to their significantly different thresholds for the jittered monotic condition ($p = 0.018$) but not other conditions ($p > 0.05$).

Discussion

This study showed that both the ASD group and the control group were capable of segregating the target and the masker that carried different spatial information. On the other hand, for the monotic conditions, the significant difference between the jittered and non-jittered thresholds observed in the control group is consistent with the theory that informational masking is eliminated when the target and maskers have different spectral uncertainties¹⁰, but this difference was not observed in the ASD group. The fact that similar thresholds for the non-jittered monotic condition were observed in these two groups indicates that individuals with ASD had no difficulty in segregating the target and the masker that carried different spectral uncertainties. Nevertheless, the performance under the jittered monotic condition was significantly better in the ASD group than in the control group. This observation indicates that for individuals with ASD, the targets and maskers with similar spectral uncertainties were not mandatorily grouped together.

Previous studies show that individuals with ASD perform normally as regards sound localization in the horizontal plane but not in the vertical plane with background noise¹¹. On the other hand, their spatial attention is affected when the target and distractors are presented from different directions at the same time⁶. These two previous studies did not differentiate the binaural cues used in horizontal sound localization. Another two studies that investigated sound segregation based on interaural time difference reported that individuals with ASD have reduced auditory evoked potentials related to pre-attentive target-masker segregation^{4,5}. The combination of these studies with our findings suggests that individuals with ASD are likely to have a deficit in processing interaural time differences but not in processing interaural intensity differences. Nevertheless, the difference could also be explained by the single burst presented in the previous studies and the multiple bursts presented in this study because auditory object continuity (in the multiple-burst case) is known to increase selective auditory attention¹².

On the other hand, the superior performance observed for the jittered monotic condition in the ASD group indicates that they were analytic listeners¹³, and this observation cannot be attributed to developmental delay because children are more susceptible to informational masking than adults¹⁴. Their superior performance in auditory stream segregation observed in this study might be attributed to (1) their enhanced discrimination between the target and masker, (2) their enhanced top-down target excitation, and (3) their enhanced distractor inhibition. The previous visual search study found that the superior performance of the ASD group is related to their superior discrimination between the target and distractors¹⁵. In this study, the target and maskers had the same spectral uncertainty under the jittered monotic condition and were grouped together in the control group, so their performance was disrupted by informational masking in this condition. However, if individuals with ASD could segregate the target and masker by the difference between their exact spectro-temporal structures, they could focus on the fixed spectral region for the target and had superior performance (as their performance in the non-jittered monotic condition). On the other hand, several studies indicate diminished top-down modulation in individuals with ASD^{16–18}, and individuals with ASD have difficulty in filtering out the distractors^{3,19,20}.

An increased prevalence of absolute pitch^{21,22} has been observed in individuals with ASD. In contrast, individuals with ASD have difficulty with processing complicated auditory stimuli²³. There are two models that explain the seemingly contradictory auditory perception in autism: the weak central coherence theory and the enhanced perceptual functioning theory^{24,25}. These two theories together describe the enhanced perception of local features and the diminished perception of global features. Nevertheless, the way to define 'global' features in the auditory domain is under debate. For example, the pitch contour is defined as a 'global' features in one study²⁶ but a 'local' feature in another²⁷, and in yet another study there are 'global' pitch contour and 'local' pitch contour²⁸. The results of this study show that individuals with ASD were less affected by 'global' processes such as grouping auditory elements with similar spectral uncertainties. Since global processes are important for speech perception, the absence of these automatic processes in those with ASD might explain the observation that their superior performance in local processes seems to hinder their performance in speech perception²⁹.

There have been few studies exploring why individuals with ASD have difficulty in isolating target speech from interfering sounds in a social environment. Our observation suggests that the absence of automatic grouping processes, which may be related to enhanced discrimination of acoustic components with similar features, may explain their difficulty in segregating target speech and background sounds. Speech requires the integration of acoustic elements across frequencies and time, especially when it is disrupted by background noise. If those with ASD have encountered difficulties in hearing target speech in a social environment since childhood because of the absence of automatic grouping processes, they may also encounter difficulties in communication and social interaction. In a clinical setting, these findings provide a possible early detection and intervention approach with respect to the difficulties in social interaction and communication exhibited by those with ASD, especially for those suffering from hearing difficulties without a detectable deficit in their peripheral auditory system.

Methods

Participants. Thirteen high-functioning adults with autism (3 females) and 13 NT subjects (5 females) were matched by age (mean $\pm$ SD ASD group: 28 ± 7.47 , NT group: 26.85 ± 5.23) and IQ (mean $\pm$ SD ASD group: 108.62 ± 10.45 , NT group: 114.31 ± 11.06) for both VIQ (mean $\pm$ SD ASD group: 112.69 ± 10.7 , NT group: 117.62 ± 11.74) and PIQ (mean $\pm$ SD ASD group: 101.85 ± 12.16 , NT group: 106.92 ± 11.89). All participants were evaluated regarding their autistic traits by autism spectrum quotient (mean $\pm$ SD ASD group: 37.15 ± 4.32 , NT group: 18.54 ± 5.19). All the participants performed normally in an audiometric test (hearing levels <20 dB from 0.25 to 8 kHz). All of the participants were naive as to the purposes of the study. All procedures were conducted in accordance with the Declaration of Helsinki and approved by the ethics committee of the NTT Communication Science Laboratories. The participants were paid for their time and gave their informed consent prior to their participation.

ASD participants were recruited from outpatient units of the Karasuyama Hospital, Tokyo, Japan. The diagnosis of ASD was based on a consensus reached by three experienced psychiatrists and one psychologist according to the criteria of Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR), fourth edition³⁰, after two detailed interviews conducted independently by a psychiatrist and a clinical psychologist belonging to the team at the hospital that included the participant's developmental history, present illness, past history, and family history. They also confirmed that none of the participants with ASD met the DSM-IV criteria for any other psychiatric disorders (e.g., mood disorders, schizophrenia, anxiety disorders, or substance-related disorders). In addition, the diagnosis was reconfirmed at least after two months later.

Auditory stimuli. The experiment was conducted in a sound-insulated booth. Auditory stimuli were generated by MATLAB (7.10.0) at a 20-kHz sampling rate and low-pass filtered at 7.5 kHz. The auditory stimuli were then processed by an audio interface (M-AUDIO FAST-TRACK PRO) and then sent to headphones (Sennheiser HDA200).

The target sequence consisted of eight 60-ms pure tones (with 10-ms ramps) that were chosen randomly on a logarithmic frequency scale from 848 to 1180 Hz. Each of the masker sequences also consisted of eight 60-ms pure tones (with 10-ms ramps) that were chosen randomly on a logarithmic frequency scale from 200 to 5000 Hz with the exclusion of the protected spectral region from 800 to 1250 Hz. In total the duration of the auditory stimuli in each interval was 480 ms. Each component in the masker was 57 dB SPL, and the total sound level was around 66 dB SPL.

The frequencies of the target and masker components were randomly selected for each interval. By selecting the frequencies of the masker components outside a protected spectral region, although energetic masking was not completely avoided, it was minimized. For 'holistic' listeners, Durlach *et al.* shows that subjects are severely distracted by the masker and find it difficult to discern the target even though there is little masker energy in the frequency region of the target¹⁰.

Procedure. Each trial consisted of two intervals, separated by a 1-s silent gap. In one of the two intervals, chosen at random to be the first or the second with equal probability, the target was presented with the masker; in the other interval, only the masker was presented. Listeners were instructed to select the interval in which the target was presented. Once listeners had given their answers, feedback was provided in the form of a message displayed on the computer screen ('correct' or 'incorrect').

The threshold of target detection was measured by using an adaptive staircase procedure with a 3-down 1-up rule, which tracks the point corresponding to 79.4% correct on the psychometric function³¹. A total of four adaptive tracks, with a minimum of 50 trials and 12 reversals each, were obtained for every condition. In the beginning of each adaptive track, the sound level of the components in the target was set to be 57 dB SPL. The sound level of the target was decreased or increased by 8 dB before the second reversal, and then the sound level was decreased or increased by 4 dB before the fourth reversal. After the fourth reversal, the sound level of the target was decreased or increased by 2 dB. The sound level of the target in the last six reversals was averaged as the threshold in that adaptive track. While the first adaptive track was served as practice, the average of the thresholds measured in the last three adaptive tracks in each condition was used for further analysis.

References

- Alcantara, J. I., Weisblatt, E. J., Moore, B. C. & Bolton, P. F. Speech-in-noise perception in high-functioning individuals with autism or Asperger's syndrome. *J Child Psychol Psychiatry* **45**, 1107–1114 (2004).
- Groen, W. B., *et al.* Intact spectral but abnormal temporal processing of auditory stimuli in autism. *J Autism Dev Disord* **39**, 742–750 (2009).
- Lepisto, T., *et al.* Auditory stream segregation in children with Asperger syndrome. *Biol Psychol* **82**, 301–307 (2009).
- Brock, J., Bzishvili, S., Reid, M., Hautus, M. & Johnson, B. W. Brief report: atypical neuromagnetic responses to illusory auditory pitch in children with autism spectrum disorders. *J Autism Dev Disord* **43**, 2726–2731 (2013).
- Lodhia, V., Brock, J., Johnson, B. W. & Hautus, M. Reduced object related negativity response indicates impaired auditory scene analysis in adults with autistic spectrum disorder. *PeerJ* **2**; <https://dx.doi.org/10.7717/peerj.261> (2014).
- Teder-Salejari, W. A., Pierce, K. L., Courchesne, E. & Hillyard, S. A. Auditory spatial localization and attention deficits in autistic adults. *Brain Res Cogn Brain Res* **23**, 221–234 (2005).
- Pollack, I. Auditory informational masking. *J Acoust Soc Am* **57**, S5; <http://dx.doi.org/10.1121/1.1995329> (1975).
- Ceponiene, R., *et al.* Speech-sound-selective auditory impairment in children with autism: they can perceive but do not attend. *Proc Natl Acad Sci U S A* **100**, 5567–5572 (2003).
- Whitehouse, A. J. & Bishop, D. V. Do children with autism 'switch off' to speech sounds? An investigation using event-related potentials. *Dev Sci* **11**, 516–524 (2008).
- Durlach, N. I., *et al.* Informational masking: counteracting the effects of stimulus uncertainty by decreasing target-masker similarity. *J Acoust Soc Am* **114**, 368–379 (2003).
- Visser, E., *et al.* Atypical vertical sound localization and sound-onset sensitivity in people with autism spectrum disorders. *J Psychiatry Neurosci* **38**, 398–406 (2013).
- Best, V., Ozmeral, E. J., Kopco, N. & Shinn-Cunningham, B. G. Object continuity enhances selective auditory attention. *Proc Natl Acad Sci USA* **105**, 13174–13178 (2008).
- Neff, D. L., Dethlefs, T. M. & Jesteadt, W. Informational masking for multicomponent maskers with spectral gaps. *J Acoust Soc Am* **94**, 3112–3126 (1993).
- Hall, J. W., 3rd, Buss, E. & Grose, J. H. Informational masking release in children and adults. *J Acoust Soc Am* **118**, 1605–1613 (2005).
- O'Riordan, M. & Plaisted, K. Enhanced discrimination in autism. *Q J Exp Psychol A* **54**, 961–979 (2001).
- Greenaway, R. & Plaisted, K. Top-down attentional modulation in autistic spectrum disorders is stimulus-specific. *Psychol Sci* **16**, 987–994 (2005).
- Ropar, D. & Mitchell, P. Shape constancy in autism: the role of prior knowledge and perspective cues. *J Child Psychol Psychiatry* **43**, 647–653 (2002).
- Soulieres, I., Mottron, L., Saumier, D. & Larochelle, S. Atypical categorical perception in autism: autonomy of discrimination? *J Autism Dev Disord* **37**, 481–490 (2007).
- Burack, J. A. Selective attention deficits in persons with autism: preliminary evidence of an inefficient attentional lens. *J Abnorm Psychol* **103**, 535–543 (1994).
- Remington, A., Swettenham, J., Campbell, R. & Coleman, M. Selective attention and perceptual load in autism spectrum disorder. *Psychol Sci* **20**, 1388–1393 (2009).
- Miller, L. K. The savant syndrome: intellectual impairment and exceptional skill. *Psychol Bull* **125**, 31–46 (1999).
- Takeuchi, A. H. & Hulse, S. H. Absolute pitch. *Psychol Bull* **113**, 345–361 (1993).
- Samson, F., *et al.* Atypical processing of auditory temporal complexity in autistics. *Neuropsychologia* **49**, 546–555 (2011).
- Happe, F. & Frith, U. The weak coherence account: detail-focused cognitive style in autism spectrum disorders. *J Autism Dev Disord* **36**, 5–25 (2006).
- Mottron, L. & Burack, J. *The development of autism: Perspectives from theory and research* (Hillsdale, NJ: Erlbaum, 2001).
- Mottron, L., Peretz, I. & Menard, E. Local and global processing of music in high-functioning persons with autism: beyond central coherence? *J Child Psychol Psychiatry* **41**, 1057–1065 (2000).
- Foxton, J. M., *et al.* Absence of auditory 'global interference' in autism. *Brain* **126**, 2703–2709 (2003).
- Bouvet, L., Simard-Meilleur, A. A., Paignon, A., Mottron, L. & Donnadieu, S. Auditory local bias and reduced global interference in autism. *Cognition* **131**, 367–372 (2014).
- O'Connor, K. Auditory processing in autism spectrum disorder: a review. *Neurosci Biobehav Rev* **36**, 836–854 (2012).
- APA. *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.) (American Psychiatric Association, Washington, DC, 2000).
- Levitt, H. Transformed up-down methods in psychoacoustics. *J Acoust Soc Am* **49**, Suppl 2:467–477 (1971).

Acknowledgements

I.F.L. was supported by MEXT KAKENHI "Constructive Developmental Science" (24119006). The authors thank the technical assistance from Dr. Masaharu Kato.

Author Contributions

I.F.L. and M.K. designed the experiment. I.F.L., T.Y., Y.K. and N.K. conducted the experiment. I.F.L. conducted the analysis and wrote the manuscript. M.K. revised the manuscript. All authors reviewed the manuscript.

Additional Information

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Lin, I. F. *et al.* Enhanced segregation of concurrent sounds with similar spectral uncertainties in individuals with autism spectrum disorder. *Sci. Rep.* **5**, 10524; doi: 10.1038/srep10524 (2015).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>