Can gaze-cueing be helpful for detecting sound in autism spectrum disorder?
（自閉症スペクトラムにおいて視線手掛かりは聴覚的注意を促進するだろうか？）

趙　朔

1. Introduction

By following others people's eye gaze, we can recognize their focus of attention. The ability to coordinate attention to an object or event with others is a phenomenon called joint attention (Mundy et al., 1986). Although joint attention is thought to be built upon basic neural mechanisms detecting the gaze direction, it also extends to understanding pointing or other social cues by the interactive partner or sharing an awareness of the object or event with the partner; therefore, it is considered to be a uniquely human ability underlying social communication. It is widely known that people attend to each other (by eye contact, smiling, mimicking facial expressions and gestures etc.) to augment shared understanding, and it plays a crucial role in verbal communication. Joint attention is important for at least two reasons: 1) it allows us to learn about the other person's inner state (such as interest, emotion and intentions), 2) it informs us about what the speaker is talking about and where the object of interest is in the environment.

Autism spectrum disorders (ASD), including autistic and Asperger's disorders, are characterized by qualitative impairments in social interaction (American Psychiatric Association [APA], 2000). One of the important features of these social impairments is thought to be a deficit of joint attention (APA, 2000; Mundy et al., 1986). A lack of joint attention has been well documented in the clinical literature and recently it has drawn attention as an early marker of ASD (APA, 2000); however, contrary to the clinical findings, experimental studies generally have reported intact joint attention in ASD (Chawarska, Klin, & Volkmar, 2003; Kylliäinen & Hietanen, 2004; Rutherford & Krysko, 2008; Senju et al., 2004; Swettenham et al., 2003; Vlamings et al., 2005; see Nation et al., 2008 for review). These studies applied Posner's cueing paradigm (Posner, 1980), in which uni-modal (i.e. visual) cue-target pairs were used. In the paradigm, subjects first saw a gaze cue (directed toward right or left), followed by a target, a dot or a letter, which appeared either on the right or left of the display screen. The subjects were asked to locate the target and respond as quickly and accurately as possible, and the reaction time (RT) was measured. Most of the previous studies observed that individuals with ASD, as well as typically developing subjects, responded faster when targets appeared in the same direction as gaze cues than when they appeared in the opposite direction. These results indicate intact joint attention in ASD.

To investigate gaze-triggered joint attention in ASD, previous studies have used visual cues and targets; however, in real life, there are various environmental stimuli, such as sounds and objects. We constantly need to use joint and shared attention with others by recognizing cues and targets that belong to different modalities; therefore, it is important to investigate the effect of joint attention under cross-modal conditions. Previous studies have examined joint attention in typically developing individuals under cross-modal conditions (Borjon et al., 2010; Newport & Howarth, 2009). Both studies presented visual cues and auditory targets and showed significant gaze-triggered joint attention under cross-modal conditions, as in a uni-modal paradigm. Based on the finding that cross-modal processing, such as attention-switching (Reed & McCarthy, 2011) and audio-visual integration (Charbonneau et al., 2013), was impaired in children with ASD and that social communication has a cross-modal aspect in many circumstances, it is speculated that individuals with ASD may fail to show joint attention under cross-modal conditions.

Previous studies have reported orientation to auditory stimuli, social voice and non-social tone in ASD. Individuals with ASD showed poor responses to social (i.e., calling the child’s name) compared with non-social (i.e., phone ringing) stimuli in 3–4-year-old children with ASD (Dawson et al., 2004). Furthermore, atypical activation in response to social sounds in adult ASD was also observed when participants were requested to simply distinguish sounds of voice and tone (Èeponienë et al., 2003; Gervais et al., 2004; Whitehouse & Bishop, 2008), although another study did not suggest impairment of the auditory cortex while listening to sounds intentionally (i.e., voice, tone or story) (Funabiki, Murai, & Toichi, 2012). Based on these findings, we manipulated social (voice) and non-social (tone) sounds to examine cross-modal joint attention without any hypothesis about the impairment of gaze-triggered attention to a specific target.

In the present study, we examined joint attention under visual-auditory cross-modal conditions in individuals with high-functioning ASD and age-matched typically developing controls. In addition, we manipulated two sounds as targets, i.e. social voice and non-social tone, to refer the relationship between the cue and target. The subjects were first presented with a neutral eye gaze as a cue and subsequently asked to identify the direction of the following auditory target as accurately and rapidly as possible. The aims of the study were: 1) to investigate whether visual-auditory cross-modal joint attention is impaired in ASD, and 2) to examine whether cross-modal joint attention is affected by the type of auditory target in ASD.

2. Materials and Methods

2.1 Ethics Statement

All subjects older than 18 years of age and the parents of those younger than 18 years of age gave written informed consent to participate in this study, the principle of which was applied to the participants in the ASD group since they all had normal IQ. This study was approved by the local ethics committee of Kyoto University Graduate School and Faculty of Medicine.

2.2 Participants

Eighteen individuals with ASD and 20 controls participated in this study. The ASD and control groups were matched for chronological age (ASD group: $M = 25.6$, $SEM = 2.1$; Control: $M = 22.6$, $SEM = 0.93$, independent t-test, $t(36) = 1.34$, $p = 0.189$) and gender (13 men and 7 women in the control group and 15 men and 3 women in the ASD group, Fisher's exact test, $p = 0.278$). All of the controls were recruited from Kyoto University students. The verbal and performance IQ in the ASD group was measured using the Japanese version of the Wechsler Adult Intelligence Scale – revised (Shinagawa 1990) or the Wechsler Intelligence Scale for Children – revised (Kodama 1982). The IQs of all participants in the ASD group were in the normal range (full-scale IQ: M = 109.2, SEM = 4.1; verbal IQ: M = 108.1, SEM = 4.8; performance IQ: M = 106.9, SEM = 3.4). All subjects were right-handed, as assessed by the Edinburgh Handedness Inventory (Oldfield, 1971), and had normal or corrected-to-normal visual and auditory acuity.

Fourteen of the ASD group had been diagnosed with Asperger's disorder, and four with pervasive developmental disorder not otherwise specified (PDD-NOS) by two child psychiatrists using DSM-IV-TR (APA, 2000). The diagnosis was based on an interview with the subjects, information from their parents, teachers, or professional counselors, and clinical records during childhood. The participants in the ASD group were outpatients who had been referred to Kyoto University Hospital or the Faculty of Human Health Science of Kyoto University Graduate School of Medicine for consultation. They were all free of neurological or psychiatric problems other than those associated with ASD and they were not taking any medication.

2.3 Design

The experiment was constructed as a four-factorial mixed randomized-repeated design, with Group (ASD or control) as the randomized factor, and auditory target conditions (voice sound or tone sound), validity conditions (valid, invalid), and SOA (200ms, 800ms) as the repeated factors.

2.4 Stimuli

We selected the cue stimuli from Ekman and Friesen (Ekman 1967). Photographs of a female model with a neutral face were selected. The gaze direction was then manipulated. The irises and pupils of the eyes were cut from the original photographs and pasted to fit over 18 pixels on the right or left side of the eyes using Photoshop 5.0 (Adobe). We cropped the photographs in an ellipse 8.3° wide and 12.1° high to exclude the hair and background.

Two types of auditory stimuli were selected. One was sampled from a native Japanese woman; an /i/ voice sound (F0 frequency of 300Hz, 80dB SPL) which is similar to the /iy/ sound in English. The other was a pure tone of similar frequency to F0 of a voice (300Hz, 80dB SPL), which was produced with the Audacity V1.3.13 (AudacityStore.com). The auditory stimuli were presented for 150ms.

2.5 Apparatus

Stimulus presentation and data acquisition were controlled by Presentation (Neurobehavioral Systems) on a Windows computer. The temporal resolution of Stimuli the Presentation system is within the range of 5-50 msec. Stimuli were presented on a 19-inch monitor (Dell: screen resolution 1024 × 768 pixels; refresh rate 100Hz). The distance between the monitor and the participants was fixed at approximately 57cm using a headrest. All of the auditory stimuli were presented through headphones. The response time (RT) measurements were based on a response button.

2.6 Procedure

The sequence of stimulus presentation is shown in Figure 1. For each trial, a fixation cross point was first presented for 600ms in the center of the screen. A neutral face with a straight gaze was then presented at this location as a background. After 500ms, a neutral facial cue with the eye gaze directed right or left was presented in the center of the screen. The stimulus onset asynchrony (SOA) between the auditory target and gaze cue was manipulated for 200/800ms. The SOA condition was presented by randomizing each auditory target condition to exclude an effect specific to a sequence of SOA conditions. Subsequently, an auditory stimuli target (voice sound or tone sound) was presented in the left or right ear for 150ms through headphones. The participants were asked to respond as quickly and exactly as possible whether the target appeared on the left or right side of the headphone by pressing the corresponding key on the switch key using the index or middle figure of their dominant hand, respectively. RT was measured in each trial. The gaze cue remained on the screen until the trial was finished. Total time of a trial was manipulated to be 3600ms. The targets appeared randomly on the same or opposite side to the gaze direction when the eyes looked left or right. If participants could not respond in a trial, the data were excluded as an incorrect result. The target appeared at the cued location in 50% of the trials. The participants were told that the cues did not predict the target location and were instructed to fix on the center of the screen in each trial.

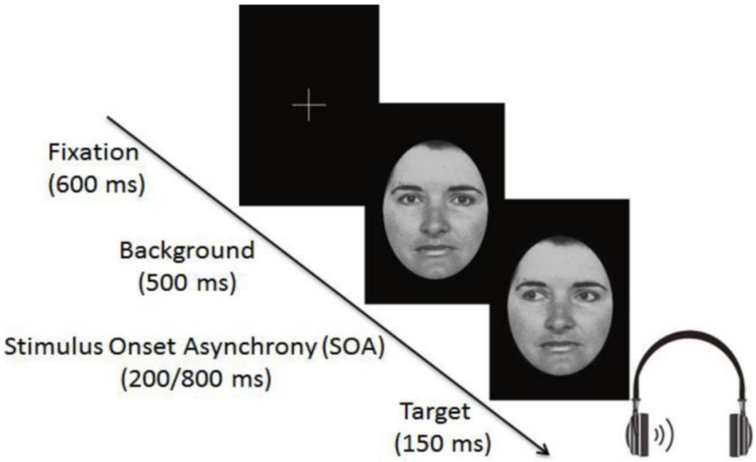

Figure 1. Illustrations of stimulus presentations. Image of the sequence of stimulus presentation is shown. We selected the face cue stimuli from Ekman and Friesen (1967). The auditory stimuli were represented by voice sound or tone sound. Speeded response time was measured from the onset of the target in each trial.

The experiment consisted of eight blocks of 28 trials, including 8 × 4 catch trials in which the target did not appear. Each condition was presented in pseudorandom order. Participants were allowed to rest between blocks. Twenty-eight practice trials preceded the experimental trials.

2.7 Date Analysis

The data were analyzed using SPSS 10.0. Incorrect responses and responses of <150ms or >1000ms were excluded from the RT analysis. The mean RT under each condition was calculated for each participant. First, the mean RT was analyzed using 4-way mixed measures analysis of variance (ANOVA) with Validity conditions (valid, invalid), auditory target conditions (voice, tone) and the SOA condition (200ms, 800ms) as within-participant factors, and Group (ASD and control) as the between-participant factor. For significant interaction, if present, the data were analyzed separately with two 3-way ANOVAs under target conditions. To examine whether two 3-way ANOVAs were significant for interaction, if present, follow-up simple-effect analyses were conducted.

3. Results

There was no significant difference between the ASD and Control groups in error rates under all conditions, indicating no speed-accuracy trade-off (all $p > 0.05$).

Mean (with SEM) RTs for all subjects under each condition are shown in Table 1. We conducted a 2 Group (ASD, Control) × 2 Target (voice, tone) × 2 Validity (valid, invalid) × 2 SOA (200, 800ms SOA) mixed-model, repeated measures ANOVA (analysis of variance) on these RTs. A significant main effect of Validity ($F (1, 36) = 14.274$, $p < 0.001$) and SOA ($F (1, 36) = 40.146$, $p < 0.001$) was found, but there were no significant main effects of Target ($F (1, 36) = 0.132$, $p > 0.05$) or Group ($F (1, 36) = 3.957$, $p > 0.05$). There was a significant interaction of Group × Target × Validity × SOA ($F (1, 36) = 6.104$, $p < 0.05$), Target × Validity × SOA ($F (1, 36) = 15.134$, $p < 0.001$), Validity × SOA ($F (1, 36) = 28.781$, $p < 0.001$), and Target × Validity ($F (1, 36) = 5.208$, $p < 0.05$).

Table.1

Mean (with SEM) response times (RTs) (ms) at tone (a) or voice (b)

(a) Mean (with SEM) RTs (ms) at tone

	200ms		800ms	
Auditory target	Valid	Invalid	Valid	Invalid
Control	402.1 (17.5)	419.2 (16.3)	368.9 (12.0)	387.4 (13.8)
ASD	473.2 (31.4)	480.6 (31.2)	427.5 (20.5)	422.8 (21.4)

(b) Mean (with SEM) RTs (ms) at voice

	200ms		800ms	
Auditory target	Valid	Invalid	Valid	Invalid
Control	389.1 (15.7)	436.8 (17.3)	378.0 (14.6)	368.7 (13.2)

ASD	458.6 (31.9)	486.4 (30.2)	426.2 (22.5)	428.9 (21.4)

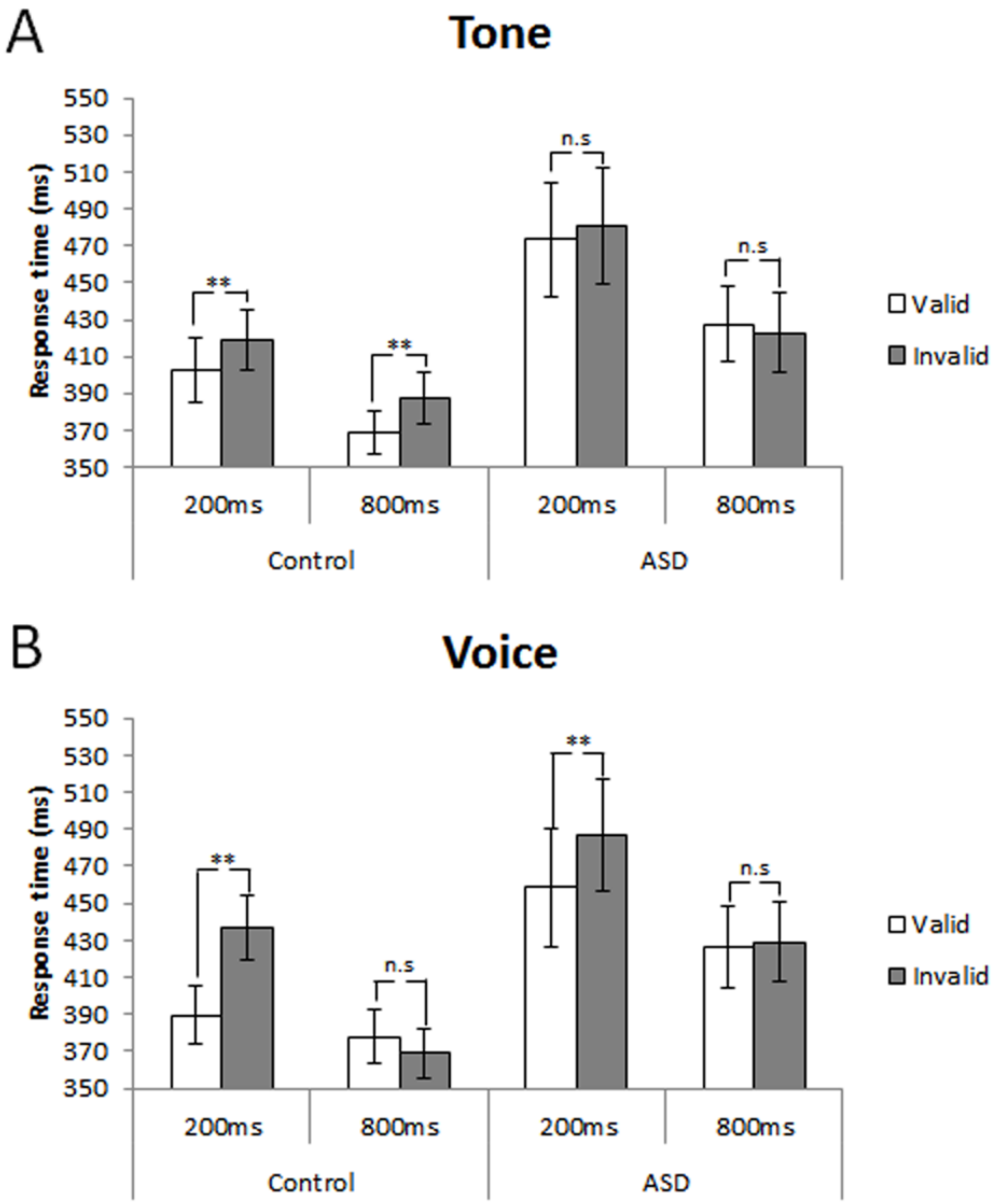

Figure 2. Effects of gaze direction on response times (RTs). Mean RTs using tone (A) as targets at different stimulus onset asynchrony (SOA) in typically developing subjects and individuals with ASD. Mean RTs using voice (B) as targets at different

*stimulus onset asynchrony (SOA) in typically developing subjects and individuals with ASD. Error bars represent standard errors (SEMs). **: p <0.01.*

Since there was a significant interaction among the 4 factors, the data using voice and tone targets were analyzed separately. In tone targets (Figure. 2A), 3-way ANOVA of the RT data indicated a significant main effect of Validity ($F (1, 36) = 7.472$, $p < 0.05$) and SOA ($F (1, 36) = 30.257$, $p < 0.001$), but there were no significant main effects of Group ($F (1, 36) = 3.974$, $p > 0.05$). There was a significant interaction of Group × Validity ($F (1, 36) = 5.528$, $p < 0.05$). A post-hoc test yielded a significant Validity effect in the control group ($p < 0.01$), but not in the ASD group ($p > 0.1$), indicating impaired joint attention to the tone target in the ASD group.

In the voice targets (Figure. 2B), a three-way ANOVA on the RT data indicated a significant main effect of Validity ($F (1, 36) = 16.007$, $p < 0.001$) and SOA (F (1, 36)) = 43.929, $p < 0.001$), but there were no significant main effects of Group ($F (1, 36) = 3.853$, $p > 0.05$). There was a significant interaction of Group × Validity × SOA ($F (1, 36) = 6.01$, $p < 0.05$). Post-hoc tests yielded a significant Group effect at invalid with 800ms SOA ($p < 0.05$), but not at others (all $p > 0.05$). In participants, there was a significant Validity in both control and ASD groups at 200ms SOA (both $p < 0.01$), but not at 800ms SOA ($p > 0.05$), indicating comparable joint attention to the voice target between control and ASD groups.

4. Discussion

The present study investigated for the first time the effect of joint attention in ASD under visual-auditory cross-modal conditions with visual cues (eye gaze) and sound targets (voice or tone). It was found that only controls responded to intact joint attention to the tone when it was presented in the same direction as the cue, but not in ASD, although both groups showed joint attention to a voice. This result suggests that, in ASD, cross-modal joint attention is impaired when responding to a specific stimulus.

The current study clearly revealed the impairment of gaze-triggered joint attention under cross-modal cue and target conditions. In contrast, all previous experimental studies reporting intact gaze-triggered joint attention in ASD used a uni-modal but not cross-modal cue and target conditions (Chawarska et al., 2003; Kylliäinen & Hietanen, 2004; Rutherford & Krysko, 2008; Senju et al., 2004; Swettenham et al., 2003; Vlamings et al., 2005). Given that cross-modal processing was impaired in ASD (Charbonneau et al., 2013; Reed & McCarthy, 2011), cross-modal cue and target conditions might have higher sensitivity to impaired gaze-triggered joint attention in ASD. Furthermore, this could explain the discrepancy between the experimental setting and real-life communication in individuals with ASD, as natural settings frequently have a cross-modal aspect.

However, impaired joint attention was not found in ASD when a voice was used as the target, although it was impaired when tone was the target. A similar magnitude of joint attention was found under two SOA conditions between typically development controls and individuals with ASD. Given that previous studies showed impaired processing of voice stimuli, this is a counterintuitive finding; however, the results could be explained by the effect of the combination of cue and target. Previous studies investigating joint attention in typically developing subjects have suggested the importance of contextual effects, which improve the effect of joint attention (Bayliss et al., 2007; Bayliss, Schuch, & Tipper, 2010; Bayliss & Tipper, 2005; Fichtenholtz et al., 2007; Friesen, Halvorson, & Graham, 2011; Kuhn & Tipples, 2011). Knowing the focus of another person in a particular context provides important cues about the other person’s interest in objects and events, and can elicit complementary effects in the observer (Sebanz, Bekkering, & Knoblich, 2006). This effect was enhanced when cues and targets were in congruent contexts. For example, Bayliss et al. (2005) found a greater orienting effect of attention when congruent contexts (i.e., a social gaze cue - a social face target) were compared to incongruent contexts (i.e., a social gaze cue - a non-social scrambled face target). In fact, a greater joint attention effect was found for voice than tone as target at shorter SOA of 200msec. This finding was confirmed with a Target × SOA ANOVA under Validity in controls (significant interaction of Target × SOA ($F(1, 19) = 16.575$, $p < 0.01$, a post-hoc test yielded a significant larger voice than tone at 200ms SOA ($p < 0.01$), in contrast to 800ms SOA ($p < 0.01$)). The influence of contextual effects in typically developing controls might be speculated to play a role in the response of individuals with ASD to intact joint attention when voice was the target. Gaze-triggered joint attention in ASD might be enhanced in a congruent context (social gaze cue - social voice) but not in an incongruent context (social gaze – non-social tone), although they showed reduced gaze-triggered joint attention under cross-modal conditions, in general.

The findings that the effects of gaze-triggered joint attention between voice and tone targets, which is reversed under 800ms SOA conditions in typically developing individuals (i.e., Inhibition of Return: IOR) might be also explained in terms of the same contextual effects. IOR refers to the finding that targets at valid locations are responded to slowly, but those at invalid locations are responded to rapidly during the time course of SOA, which can be modulated by contextual effects (Taylor & Therrien, 2008). In this study, IOR occurred earlier in a congruent context (i.e., non-social symbolic cue - non-social scrambled face target) than in an incongruent context (i.e., non-social symbolic cue – social face target) during the time course of SOA. In the current study we manipulated two types of sounds as targets (i.e., voice and tone), referring to the weak relationship with a cue target in an incongruent context between social gaze and non-social tone and to the strong relationship with a cue target in a congruent context between social gaze and social voice. When the previous result (Taylor & Therrien, 2008) was replicated in our study, an earlier IOR occurred in a congruent context (i.e., social gaze - social voice) at longer SOA of 800msec, but not in an incongruent context (i.e., social gaze - non-social tone). Although the current study did not address this issue, the results suggest the contextual modulation of IOR gaze-triggered joint attention.

The current results have some clinical implications for social adaptation in individuals with ASD. The results showed that individuals with ASD have impaired gaze-triggered joint attention to a specific target (tone). In contrast with typically developing individuals showing gaze cueing irrespective of the target stimulus, this might reduce the chances of forming joint attention with others in real life; however, the results might provide a clue to enhance joint attention in ASD because they show intact gaze-triggered joint attention to a specific target (voice). A previous study showed that the focus of eye gaze normalized when the topic of conversation was interesting to individuals with ASD (Nadig et al, 2010). As the participants with ASD in the current study had higher or average verbal ability, voice rather than tone targets might be more interesting for them. The current results suggest that the use of interesting targets may facilitate adaptive behavior, including joint attention in individuals with ASD.

Researchers have proposed that the neural mechanism for gaze-triggered joint attention includes the STS (Akiyama et al., 2006) and the amygdala (Okada et al., 2008), which have reciprocal connections in the processing of social stimuli (Adolphs, 1999; Klein, Shepherd, & Platt, 2009). In addition, previous studies have demonstrated that cerebral activation in response to social stimuli was modulated by contextual components (such as congruent and incongruent action) in STS (Pelphrey, Morris, & McCarthy, 2004; Vander Wyk, Voos, & Pelphrey, 2012; see Zilbovicius et al., 2006 for review) and the amygdala (Kim et al., 2004) in typically developing people. Based on these findings, the greater cross-modal joint attention in congruent contexts compared to incongruent contexts in both ASD and control groups may be explained by the enhancement of STS/amygdala functions due to the strong relevance of gaze-voice pairs.

Two limitations of this study are that we tested joint attention with two types of targets (i.e., social and non-social targets) under only visual-auditory cross-modal conditions in individuals with ASD. In future research, joint attention conditions should also be examined with two types of targets under visual-visual un-modal conditions, in which joint attention has been reported to be intact. Second, this study included only individuals with high-functioning ASD. Further studies are needed to investigate whether low-functioning people with ASD also show gaze-triggered joint attention to voice targets.

5. Conclusions

This is the first study to investigate visual-auditory cross-modal joint attention in ASD. We found impaired joint attention to cue-target pairs in an incongruent context, whereas intact joint attention to these pairs in a congruent context, suggesting the narrow focus of joint attention in ASD. Investigation of cognitive function under natural conditions may improve our understanding of the social behaviors of individuals with ASD in the real world.

Acknowledgements

We are grateful to the subjects of this study. This study was technically supported by an incorporated non-profit organization called Organization for promoting Developmental Disorder Research, Japan.

References

Adolphs, R. (1999). Social cognition and the human brain. *Trends in Cognitive Sciences, 3*(12), 469-479.

Akiyama, T., Kato, M., Muramatsu, T., Saito, F., Umeda, S., & Kashima, H. (2006). Gaze but not arrows: A dissociative impairment after right superior temporal gyrus damage. *Neuropsychologia*, *44*(10), 1804-1810.

American Psychiatric Association [APA]. (2000). Diagnostic and statistical manual for mental disorders (DSM-IV-TR). *Washington, DC: APA.*

Bayliss, A. P., Frischen, A., Fenske, M. J., & Tipper, S. P. (2007). Affective evaluations of objects are influenced by observed gaze direction and emotional expression. *Cognition*, *104*(3), 644-653.

Bayliss, A. P., Schuch, S., & Tipper, S. P. (2010). Gaze cueing elicited by emotional faces is influenced by affective context. *Visual Cognition, 18 (8)*(1350-6285), 1214-1232.

Bayliss, A. P., & Tipper, S. P. (2005). Gaze and arrow cueing of attention reveals individual differences along the autism spectrum as a function of target context. *British Journal of Psychology, 96*(1), 95-114.

Borjon, J. I., Shepherd, S. V., Todorov, A., & Ghazanfar, A. A. (2010). Eye-gaze and arrow cues influence elementary sound perception. *Proceedings of the Royal Society B: Biological Sciences* 287 (1714),1997-2004.

Čeponienė, R., Lepistö, T., Shestakova, A., Vanhala, R., Alku, P., Näätänen, R., & Yaguchi, K. (2003). Speech–sound-selective auditory impairment in children with autism: They can perceive but do not attend. *Proceedings of the National Academy of Sciences*, *100*(9), 5567-5572.

Charbonneau, G., Bertone, A., Lepore, F., Nassim, M., Lassonde, M., Mottron, L., & Collignon, O. (2013). Multilevel deficits in the processing of audio–visual emotion expressions in autism spectrum disorders. *Neuropsychologia* 51(5) 1002-1010.

Chawarska, K., Klin, A., & Volkmar, F. (2003). Automatic Attention Cueing Through Eye Movement in 2-Year-Old Children With Autism. *Child Development, 74*(4), 1108-1122.

Dawson, Geraldine, TOTH, Karen, ABBOTT, Robert, OSTERLING, Julie, MUNSON, Jeff, ESTES, Annette, LIAW, & Jane. (2004). *Early social attention impairments in autism: Social orienting, joint attention, and attention to distress*. Developmental Psychology, 40(2), 271-283.

Ekman P, & Friesen WV (1967) *Pictures of facial affect.* Palo Alto, CA: Consulting Psychologists Press.

Fichtenholtz, H. M., Hopfinger, J. B., Graham, R., Detwiler, J. M., & LaBar, K. S. (2007). Happy and fearful emotion in cues and targets modulate event-related potential indices of gaze-directed attentional orienting. *Social Cognitive and*

Affective Neuroscience, 2(4), 323-333.

Friesen, C. K., Halvorson, K., & Graham, R. (2011). Emotionally meaningful targets enhance orienting triggered by a fearful gazing face. *Cognition and Emotion, 25*(1), 73-88.

Funabiki, Y., Murai, T., & Toichi, M. (2012). Cortical activation during attention to sound in autism spectrum disorders. *Research in Developmental Disabilities, 33*(2), 518-524.

Gervais, H., Belin, P., Boddaert, N., Leboyer, M., Coez, A., Sfaello, I., Barthelemy, C., Brunelle, F., Samson, Y., & Zilbovicius, M. (2004). Abnormal cortical voice processing in autism. *Nat Neurosci, 7*(8), 801-802.

Kim, H., Somerville, L. H., Johnstone, T., Polis, S., Alexander, A. L., Shin, L. M., & Whalen, P. J. (2004). Contextual Modulation of Amygdala Responsivity to Surprised Faces. *Journal of Cognitive Neuroscience, 16*(10), 1730-1745.

Klein, J. T., Shepherd, S. V., & Platt, M. L. (2009). Social Attention and the Brain. *Current biology : CB, 19*(20), R958-R962.

Kuhn, G., & Tipples, J. (2011). Increased gaze following for fearful faces. It depends on what you're looking for! *Psychonomic Bulletin & Review, 18*(1), 89-95.

Kodama, H., Shinagawa, F., & Motegi, M. (1982). Japanese Wechsler intelligence scale for Children. *Journal of Autism and Developmental Disorders, 20,* 87-100.

Kylliäinen, A., & Hietanen, J. K. (2004). Attention orienting by another's gaze direction in children with autism. *Journal of Child Psychology and Psychiatry, 45*(3), 435-444.

Mundy, P., Sigman, M., Ungerer, J., & Sherman, T. (1986). Defining the social deficits of autism: the contribution of non-verbal communication measures. *Journal of Child Psychology and Psychiatry, 27*(5), 657-669.

Nadig, A., Lee, I., Singh, L., Bosshart, K., & Ozonoff, S. (2010). How does the topic of conversation affect verbal exchange and eye gaze? A comparison between typical development and high-functioning autism. *Neuropsychologia, 48*, 2730-2739.

Nation, Kate, Penny, & Sophia. (2008). *Sensitivity to eye gaze in autism : Is it normal? Is it automatic? Is it social?* Development and Psychopathology, 20(1), 79-97.

Newport, R., & Howarth, S. (2009). Social gaze cueing to auditory locations. *The Quarterly Journal of Experimental Psychology, 62*(4), 625-634.

Okada, Takashi, Sato, Wataru, Kubota, Yasutaka, Usui, Keiko, Inoue, Yushi, Murai, Toshiya, Hayashi, Takuji, Toichi, & Motomi. (2008). Involvement of medial temporal structures in reflexive attention shift by gaze. *Social Cognitive and Affect Neuroscience,* 3(1), 80-88.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia, 9*(1), 97-113.

Pelphrey, K. A., Morris, J. P., & McCarthy, G. (2004). Grasping the Intentions of Others: The Perceived Intentionality of an Action Influences Activity in the Superior Temporal Sulcus during Social Perception. *Journal of Cognitive Neuroscience, 16*(10), 1706-1716.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32*, 3-25.

Reed, P., & McCarthy, J. (2011). Cross-Modal Attention-Switching is Impaired in Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders*, 1-7.

Rutherford, M., & Krysko, K. (2008). Eye Direction, Not Movement Direction, Predicts Attention Shifts in Those with Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders, 38*(10), 1958-1965.

Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: bodies and minds moving together. *Trends in Cognitive Sciences, 10*(2), 70-76.

Senju, A., Tojo, Y., Dairoku, H., & Hasegawa, T. (2004). Reflexive orienting in response to eye gaze and an arrow in children with and without autism. *Journal of Child Psychology and Psychiatry, 45*(3), 445-458.

Shinagawa, F., Kobayashi, S., Fujita, K., & Maekawa, H. (1990). Ja*panese Wechsler adult intelligence scale – Revised.* Tokyo: Nihon-Bunka-Kagaku-sha.

Swettenham, John, Condie, Samantha, Campbell, Ruth, Milne, Elizabeth, Coleman, & Mike. (2003). Does the perception of moving eyes trigger reflexive visual orienting in autism? *Philosophical Transcactions of the Royal Society of Lodon. Series B: Biological Science*, 358(1430), 325-334.

Taylor, T., & Therrien, M. (2008). Inhibition of return for the discrimination of faces. *Perception & Psychophysics, 70*(2), 279-290.

Vander Wyk, B. C., Voos, A., & Pelphrey, K. A. (2012). Action representation in the superior temporal sulcus in children and adults: An fMRI study. *Developmental Cognitive Neuroscience, 2*(4), 409-416.

Vlamings, P. H. J. M., Stauder, J. E. A., van Son, I. A. M., Mottron, & Laurent. (2005). Atypical Visual Orienting to Gaze- and Arrow-Cues in Adults with High Functioning Autism. *Journal of Autism and Developmental Disorders, 35*(3), 267-277.

Whitehouse, A. J. O., & Bishop, D. V. M. (2008). Do children with autism 'switch off' to speech sounds? An investigation using event-related potentials. *Developmental Science, 11*(4), 516-524.

Zilbovicius, M., Meresse, I., Chabane, N., Brunelle, F., Samson, Y., & Boddaert, N. (2006). Autism, the superior temporal sulcus and social perception. *Trends in Neurosciences, 29*(7), 359-366.