Vol. 6 No.2 (2015) 9-12.
LEBS
Letters on Evolutionary Behavioral Science

# The McGurk Effect and Autistic Traits: An Analogue Perspective

**Yuta Ujiie[1,2,*], Tomohisa Asai[3], Akihiro Tanaka[4], and Akio Wakabayashi[5]**

[1] Graduate School of Advanced Integration Science, Chiba University, 1-33 Yayoi-cho, Inage, Chiba 263-8522, Japan

[2] Japan Society for the Promotion of Science, 5-3-1 Kojimachi, Chiyoda-ku, Tokyo 102-0083, Japan

[3] NTT Communication Science Laboratories, NTT Corporation, 3-1 Morinosato-Wakamiya, Atsugi, Kanagawa 243-0198, Japan

[4] Department of Psychology, Tokyo Woman's Christian University, 2-6-1 Zempukuji, Suginami-ku, Tokyo 167-8585, Japan

[5] Faculty of Letters, Chiba University, 1-33 Yayoi-cho, Inage, Chiba 263-8522, Japan

[*]Author for correspondence (chiba_psyc_individual@yahoo.co.jp)

The McGurk effect is a perceptual phenomenon that an observer would perceive the intermediate phoneme when presented with a speaking movie dubbed with an incongruent voice. Several studies on autism spectrum disorders (ASD) revealed that individuals with ASD showed weak influence of visual speech in the McGurk effect. Other studies, however, reported negative results. This inconsistency might be caused by the clinical group's heterogeneity. This study examined the relationship between autistic traits and the McGurk effect among 51 healthy university students based on the dimensional model of ASD (Frith, 1991). Results revealed that for the McGurk stimuli, autistic traits negatively correlated with the rate of visual response and positively correlated with that of fused response; no correlation was observed with accuracy in perceiving the audiovisual congruent stimuli. This indicates that autistic traits might predict the weak influence of visual speech in the McGurk effect.

**Keywords**
autism spectrum disorder, autism spectrum quotient, the McGurk effect, individual differences.

## Introduction

Autism spectrum disorder (ASD) is defined in terms of difficulties in social interaction and communication, repetitive behavior, and atypical input of sensory information (American Psychiatric Association, 2013). Indeed, considerable evidence has proved that individuals with ASD showed atypical processing in not only theory of mind tasks but also various perceptual and cognitive tasks (Happé & Frith, 2006). Specifically, audiovisual speech perception in people with ASD has received focus, as their disturbed social communication skills might result from an atypical multisensory perceptual integration (Massaro & Bosseler, 2006).

The McGurk effect (McGurk & MacDonald, 1976) is a well-known illusion in speech perception where an observer perceives the intermediate phoneme when a speaking movie dubbed with an incongruent voice is presented, indicating a multisensory integration in normal population. This illusion might reflect an adaptive function that visual information improves the auditory speech perception in noisy conditions (de Gelder & Bertelson, 2003). However, several studies on ASD discovered an atypical integration in the McGurk effect (de Gelder, Vroomen, & van der Heide, 1991; Saalasti et al., 2012; Stevenson et al., 2014b). Though these researches reported that individuals with ASD showed a weak influence of visual speech on perceiving voice, other studies reported no significant difference (e.g., Taylor, Isaac, & Milne, 2010; Woynaroski et al., 2013).

This inconsistency might be caused by the clinical group's heterogeneity (Walter, Dassonville, & Bochsler, 2009) in, for example, severity of symptoms, multiple diagnosis (i.e., ASD plus ADHD), or intelligence quotient. In this situation, an analogue design to study ASD symptoms in the general population would help control potentially confounding factors (Reed, Lowe, & Everett, 2011), as the dimensional model of ASD assumes that ASD symptoms could be distributed on a continuum over the clinical as well as general population who are free from factors such as multiple diagnosis (Frith, 1991).

This study examined the relationship between the McGurk effect and autistic traits in the general population. We hypothesized that autistic traits should predict the weak influence of visual speech in the McGurk effect as suggested by some previous clinical studies of ASD (Saalasti et al., 2012). The results would be discussed in terms of the dimensional model of ASD as well as multisensory integration in people with ASD and autistic traits.

## Methods

### *Participants*

Fifty-one healthy students (12 males and 39 females) participated in the experiment. Participants' mean age was 19.8 years (*SD* = 1.41). People who reported normal or corrected vision and hearing were all native Japanese speakers. Participants provided written informed consent before the experiment.

### *Stimuli*

We used the autism-spectrum quotient (AQ) Japanese version (Wakabayashi, Baron-Cohen, Wheelwright, & Tojo, 2006) to assess the degree of autistic traits. This questionnaire contained 50 items with a four-point Likert scale (from "agree" to "disagree"). Each item was scored as 0 or 1, with possible ranges of total scores from 0 to 50.



doi: 10.5178/lebs.2015.34.
*Received 15 May 2015.*
*Accepted 27 June 2015.*
*Published online 12 August 2015.*

As the McGurk stimuli, we used the combination of a voice (/pa/) and a movie (/ka/) (Sekiyama, 1994). The audiovisual congruent and incongruent stimuli were created from simultaneous audio and video recordings of six Japanese speakers' utterances (three females) for three syllables (/pa/, /ta/, and /ka/). The videos (640 × 480 pixels, Cinepak Codec video compression, 29.97 frames/s) were speakers' faces recorded using a digital video camera (HDR-A1J, Sony). The voices (digitized at 48 kHz, with a 16-bit quantization resolution) were collected using a condenser microphone (ECM-77B, Sony). These were combined and synchronized using Adobe Premiere Pro 2.0. A pink noise was added to the voices (the signal-noise ratio was 0 dB) to enhance the influence of visual speech on voice perception. The mean duration of the audiovisual stimuli was 933 ms.

There were 18 stimuli—three syllables derived from six speakers—in the congruent condition and six stimuli—one combination (audio /pa/, visual /ka/) created for every six speakers—in the incongruent condition. A combination of a video (/pa/) and a voice (/ka/) was excluded, because the percept (/pka/) caused by this combination does not exist in Japanese native syllables. Therefore, incongruent stimuli were presented three times per block to match the number of congruent stimuli.

### Procedure

The experiment was conducted using Hot Soup Processor Ver. 3.3 (Onion software). Participants were seated at a distance of approximately 50 cm from the CRT monitor (E193FPp, Dell), with a 800 × 600 pixel screen resolution. A voice and pink noise were presented from a headphone (QuietComfort 2, BOSE) at approximately 70 dB SPL adjusted using a mixing console (MW8CX, YAMAHA). On each trial, a fixation point was displayed for 1,000 ms at the center of the monitor followed by either the congruent or incongruent stimulus. Then, a blank display was presented until participants indicated what they heard by pressing a key (/pa/, /ta/, or /ka/). The order of trials was randomized in each block. All participants completed the main session of three blocks after a practice session of six trials; subsequently, they completed the AQ.

## Results

The mean accuracy in the congruent condition and mean response rate in the incongruent condition were summarized in Table 1. The mean accuracy in the congruent condition ($M$ = .77) was higher than in the incongruent condition [$M$ = .02, $t(50) = 69.7$, $p$ = .00]. To examine the degree of visual influence on voice perception, we conducted a chi-square test for each response rate in the incongruent condition. Possible responses in the incongruent condition were of three types: audio response (/pa/ response), fused response (/ta/ response), and visual response (/ka/ response). Consequently, audio response ($M$ = .02) was observed less than fused response [$M$ = .56; $\chi 2$ = 37.36, $p < .01$] and visual response [$M$ = .43; $\chi 2$ = 32.33, $p < .01$]. This indicates that the perception of an audio input was disrupted by a visual input under the incongruent audiovisual speech condition.

Then, to examine the relationship between task performances and AQ scores, we calculated Pearson's correlation coefficients (Table 1). The scores on the AQ ranged from 9 to 35 with a mean of 20.4 ($SD$ = 6.59). No significant correlation was observed for the congruent condition [$r(51) = -.17$, $n. s.$]. Conversely, for the incongruent condition, a significant positive correlation between AQ scores and fused responses [$r(51) = .55$, $p < .01$] and a significant negative correlation between AQ scores and visual responses [$r(51) = -.53$, $p < .01$] were found (Figure 1). These correlations suggest that individuals with high AQ showed less visual response and more fused response than individuals with low AQ in the incongruent condition.

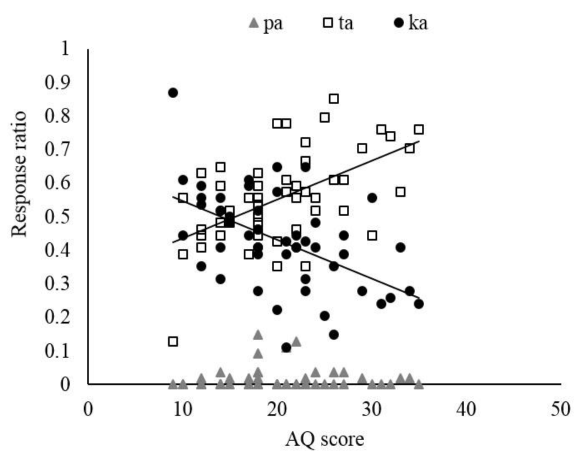


**Figure 1.** Relationships between the AQ scores and each response rate in the incongruent condition (N = 51).

**Table 1**. Mean, Standard deviations (SD), and correlations between the AQ scores and correct response rate in the congruent condition, and each response rate in the incongruent condition.

| | Mean | SD | Correlations with the AQ scores |
|---|---|---|---|
| Audiovisual congruent condition | | | |
| Correct response (/pa/,/ta/,/ka/) | .77 | .06 | −.17 |
| Audiovisual incongruent condition | | | |
| Audio response (/pa/) | .02 | .03 | .02 |
| Fused response (/ta/) | .56 | .14 | .55 ** |
| Visual response (/ka/) | .43 | .14 | −.53 ** |

Note: N = 51. AQ = Autism-spectrum Quotient. Correct response (/pa/, /ta/, /ka/) = Mean correct responses for all stimuli in the congruent condition.

## Discussion

In this study, we found that autistic traits negatively correlated with visual responses and positively correlated with fused responses for McGurk stimuli. This indicates that autistic traits might predict the weaker influence of visual information in audiovisual speech perception, as the relationship between visual and fused responses was a trade-off. Less visual responses could produce more fused responses in the current experiment where the background noise was played; therefore, audio response was rare (Saalasti, Tiippana, Kätsyri, & Sams, 2011; Saalasti et al., 2012). This finding supports the previous studies concerning children with ASD (de Gelder et al., 1991; Stevenson et al., 2014b) and adults with ASD (Saalasti et al., 2012). As such people tend to show maladaptive behavior in social interaction, an atypical processing in audiovisual speech integration might underlie their disturbed social communication skills as suggested by Massaro and Bosseler (2006). Our results would have implications for the dimensional model of ASD where individuals with high autistic traits should have similar tendencies toward people with ASD—at least in audiovisual speech integration.

These results might reflect an atypical integration in general multisensory inputs in ASD. Further, an atypical integration in ASD has been reported in audiovisual temporal processing (Donohue, Darling, & Mitroff, 2012; Stevenson et al., 2014a). Stevenson et al. (2014a) revealed that among people with ASD, poor temporal acuity across vision and audition was related to the weakness of binding of auditory and visual speech in the McGurk effect, suggesting that deficits in low-level sensory processing lead to higher-order domains. Additionally, a weak integration in ASD has been reported in other multisensory integration, for example, in the rubber hand illusion that elicits integration of visual, tactile, and proprioceptive input (Paton, Hohwy, & Entricott, 2012). Palmer, Paton, Hohwy, and Entricott (2013) suggested that people with high AQ showed reduced illusion because they depend less on visual input than other sensory inputs. This is consistent with the result of a previous clinical study (Paton et al., 2012). These results indicate that the degree of autistic traits could predict the weak reliance of visual input when visual and other sensory inputs are incongruent. Further studies should examine whether the weak reliance of visual input is generally observed in multisensory integration.

Furthermore, the effect of other factors should be considered. For example, the role of visual attention modulates the likelihood of the McGurk effect (Tiippana, Anderson, & Sams, 2003). A dysfunction to control attention is included in ASD symptoms (Wakabayashi et al., 2006) and has been reported in several cognitive tasks (Happé & Frith, 2006). Therefore, future studies should examine the intervention of visual attention in individual differences of the McGurk effect.

## Acknowledgements

We thank all the students for voluntarily participating in our experiments. This study was supported by a Grant-in-Aid from the Japan Society for the Promotion of Science Fellows (Grant No. 26-8144).

## References

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (5th ed.). Washington, DC: American Psychiatric Pub.

de Gelder, B., & Bertelson, P. (2003). Multisensory integration, perception, and ecological validity. Trends in Cognitive Sciences, 7, 460-467. (doi: 10.1016/j.tics.2003.08.014).

de Gelder, B., Vroomen, J., & van der Heide, L. (1991). Face recognition and lip-reading in autism. European Journal of Cognitive Psychology, 3, 69-86. (doi: 10.1080/09541449108406220).

Donohue, S. E., Darling, E. F., & Mitroff, S. R. (2012). Links between multisensory processing and autism. Experimental Brain Research, 222, 377-387. (doi: 10.1007/s00221-012-3223-4)

Frith, U. (1991). Autism and Asperger's syndrome. Cambridge: Cambridge University Press.

Happé, F., & Frith, U. (2006). The weak coherence account: Detail-focused cognitive style in autism spectrum disorders. Journal of Autism and Developmental Disorders, 36, 5-25. (doi: 10.1007/s10803-005-0039-0)

Massaro, D. W., & Bosseler, A. (2006). Read my lips: The importance of the face in a computer-animated tutor for vocabulary learning by children with autism. Autism, 10, 495-510. (doi: 10.1177/1362361306066599)

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. Nature, 264, 746-748. (doi: 10.1038/264746a0)

Palmer, C. J., Paton, B., Hohwy, J., & Entricott, P. G. (2013). Movement under uncertainty: The effects of the rubber-hand illusion vary along the nonclinical autism spectrum. Neuropsychologia, 51, 1942-1951. (doi: 10.1016/j.neuropsychologia.2013.06.020)

Paton, B., Hohwy, J., & Entricott, P. G. (2012). The rubber hand illusion reveals proprioceptive and sensorimotor differences in autism spectrum disorders. Journal of Autism and Developmental Disorders, 42, 1870-1883. (doi: 10.1007/s10803-011-1430-7)

Reed, P., Lowe, C., & Everett, R. (2011). Perceptual learning and perceptual search are altered in male university students with higher Autism Quotient scores. Personality and Individual Differences, 51, 732-736. (doi: 10.1016/j.paid.2011.06.016)

Saalasti, S., Kätsyri, J., Tiippana, K., Laine-Hernandez, M., von Wendt, L., & Sams, M. (2012). Audiovisual speech perception and eye gaze behavior of adults with Asperger syndrome. Journal of Autism and Developmental Disorders, 42, 1606-1615. (doi: 10.1007/s10803-011-1400-0)

Saalasti, S., Tiippana, K., Kätsyri, J., and Sams,M. (2011). The effect of visual spatial attention on audiovisual speech perception in adults with Asperger syndrome. Exp. Brain Res. 213, 283–290. (doi: 10.1007/s00221-011-2751-7)

Sekiyama, K. (1994). Difference in auditory-visual speech perception between Japanese and Americans: McGurk effect as a function of incompatibility. Journal of the Acoustical Society of Japan, 15, 143-158. (doi: 10.1250/ast.15.143)

Stevenson, R. A., Siemann, J. K., Schneider, B. C., Eberly, H. E., Woynaroski, T. G., Camarata, S. M.,

& Wallace, M. T. (2014a). Multisensory temporal integration in autism spectrum disorders. The Journal of Neuroscience, 34, 691-697. (doi: 10.1523/JNEUROSCI.3615-13.2014)

Stevenson, R. A., Siemann, J. K., Woynaroski, T. G., Schneider, B. C., Eberly, H. E., Camarata, S. M., & Wallace, M. T. (2014b). Brief report: Arrested development of audiovisual speech perception in autism spectrum disorders. Journal of Autism and Developmental Disorders, 44, 1470-1477. (doi: 10.1007/s10803-013-1992-7)

Taylor, N., Isaac, C., & Milne, E. (2010). A comparison of the development of audiovisual integration in children with autism spectrum disorders and typically developing children. Journal of Autism and Developmental Disorders, 40, 1403-1411. (doi: 10.1007/s10803-010-1000-4)

Tiippana, K., Andersen, T. S. & Sams, M. (2004). Visual attention modulates audiovisual speech perception. European Journal of Cognitive Psychology, 16, 457-472. (doi: 10.1080/09541440340000268)

Wakabayashi, A., Baron-Cohen, S., Wheelwright, S., & Tojo, Y. (2006). The Autism-Spectrum Quotient (AQ) in Japan: A cross-cultural comparison. Journal of Autism and Developmental Disorders, 36, 263-270. (doi: 10.1007/s10803-005-0061-2)

Walter, E., Dassonville, P., & Bochsler, T. M. (2009). A specific autistic trait that modulates visuospatial illusion susceptibility. Journal of Autism and Developmental Disorders, 39, 339-349. (doi: 10.1007/s10803-008-0630-2)

Woynaroski, T. G., Kwakye, L. D., Foss-Feig, J. H., Stevenson, R. A., Stone, W. L., & Wallace, M. T. (2013). Multisensory speech perception in children with autism spectrum disorders. Journal of Autism and Developmental Disorders, 43, 2891-2902. (doi: 10.1007/s10803-013-1836-5)