Running head: THE MCGURK EFFECT IN CHILDREN

The McGurk Effect in Children with Autism and Asperger Syndrome

James M. Bebko, Jessica H. Schroeder, & Jonathan A. Weiss

York University, Toronto, Ontario, Canada

Correspondence:

Dr. James M. Bebko

York University, Department of Psychology

4700 Keele St., Toronto, Ontario, M3J 1P3

Phone: (416) 736-2100, ext. 66250

Fax: (416) 736-5814

Email: jbebko@yorku.ca

This is the pre-peer reviewed version of the following article: Bebko, J., Schroeder, J., & Weiss, J. A. (2014). The McGurk effect in children with Autism and Asperger Syndrome. *Autism Research, 7*(1), 50-59, which has been published in final form at http://onlinelibrary.wiley.com/doi/10.1002/aur.1343/full. 

## Lay Abstract

Children with autism may have difficulties putting together what they see and hear during speech, which has been linked to understanding of speech and language development. However, little has been done to examine children with Asperger Syndrome as a group on tasks assessing integration of what is seen and heard during speech, despite this group's often greater language skills. Samples of children with Autism, Asperger Syndrome, and Down Syndrome, as well as a typically developing sample, were presented with an auditory-only condition, a speech-reading condition, and an audiovisual condition that involved mismatching auditory and visual signals. Children with Autism demonstrated auditory-only and speech-reading performance at the same level as the other groups, yet showed a lower performance on the audiovisual condition compared to the Asperger, Down and typical samples. These results suggest that children with Autism may have unique difficulties integrating what is seen and heard during speech perception that may be linked to how they mentally representation speech sounds.

# Scientific Abstract

Children with autism may have difficulties in audiovisual speech perception, which has been linked to speech perception and language development. However, little has been done to examine children with Asperger Syndrome as a group on tasks assessing audiovisual speech perception, despite this group's often greater language skills. Samples of children with Autism, Asperger Syndrome, and Down Syndrome, as well as a typically developing sample, were presented with an auditory-only condition, a speech-reading condition, and an audiovisual condition designed to elicit the McGurk effect. Children with Autism demonstrated unimodal performance at the same level as the other groups, yet showed a lower rate of the McGurk effect compared to the Asperger, Down and typical samples. These results suggest that children with Autism may have unique intermodal speech perception difficulties linked to their representations of speech sounds.

Keywords: Autism, Intermodal Perception, Asperger Syndrome, Intellectual Disability, Speech

Individuals with Autism are impaired in many aspects of communication, both verbal and nonverbal, and lack an important preference for speech sounds over non-speech sounds (Klin, 1991). There have also been perceptual abnormalities identified, such as deficits in intermodal perception (IMP), and specifically in speech processing (e.g. Bebko, Weiss, Demark, & Gomez, 2006; Mongillo, Irwin, Whalen, Klaiman, Carter, & Schultz, 2008). Although the sensory impairment literature is fraught with methodological cautions, especially with reference to multisensory processing (see Rogers & Ozonoff, 2005, for a recent review), it is nonetheless important to determine the linkages between basic perceptual and cognitive processes and the development of higher order linguistic competence. As a part of that research, it is critical to examine whether there are differences in the basic processing of audiovisual signals among children in the different diagnostic subgroupings that comprise the "autism spectrum disorders" (ASDs).

In the present study we investigated the ability of children with Autism to process auditory and visual speech information as a unified percept, a process called intermodal speech perception, compared to peers with Asperger Syndrome (AS), Down Syndrome (DS) and typical development (TD). We examined susceptibility to the "McGurk effect," which occurs when auditory and visual information in speech is put into conflict by presenting one syllable acoustically (e.g., /ba/) and another syllable visually (e.g., /ga/), resulting in a "fused" perception of a different syllable (e.g., /da/; McGurk & MacDonald, 1976). The McGurk effect and related visual capture effects (e.g., Desjardins, Rogers, & Werker, 1997) demonstrate that speech perception is an intermodal process and seems to be "integrated by perceptual mechanisms to form a phenomenally unified phonetic percept (Kuhl & Meltzoff, 1984, p. 363)."

Three studies have demonstrated that individuals with autism-related disorders show little visual influence on speech perception. de Gelder, Vroomen and van der Heide (1991) found that while those with Autism and those with TD performed equally well during auditory-only and visual-only conditions, the Autism group demonstrated significantly less visual influence on their perceptions during the audiovisual McGurk condition. Consistent results were found in studies by Mongillo and colleagues (2008) and Irwin, Tornatore, Brancazio, and Whalen (2011), both of which included several intermodal tasks, including the McGurk task. Children with ASD performed equally well on audiovisual tasks involving non-human stimuli, but had lower intermodal performance on tasks involving human faces, including lower reported McGurk rates relative to TD peers.

Several other studies yielded contradictory results. In one (Williams, Massaro, Peel, Bosseler, & Suddendorf, 2004), children with ASD performed worse than TD peers in both uni- and bimodal conditions. After statistically controlling for speech-reading ability (unimodal visual condition), the group difference in the bimodal condition was lost, suggesting that the difference in the McGurk was best explained by poor speech-reading ability in children with ASD. Iarocci, Rombough, Yager, Weeks and Chua (2010) and Keane, Rosenthal, Chun and Shams (2010) also reported no group difference in the bimodal conditions after controlling for unimodal differences, although the Keane et al. study was with adults.

As Williams and colleagues (2004) acknowledged, one issue that confounds interpretation is the heterogeneity of the ASD group. The ASD term combines three of the DSM-IV subcategories within Pervasive Developmental Disorders (American Psychiatric Association, 2000): Autism, AS and PDD-NOS. However, AS may show characteristically different patterns of cognitive functioning and different language abilities than individuals with "classic" Autism.

This is the first study to directly investigate IMP in individuals with AS compared to children with Autism. Although the subject of ongoing debate, the current DSM-IV diagnostic criteria (APA, 2000) differentiate AS and Autism in part on the absence of clinically significant early and current speech and language impairments, although the developmental courses can be variable (e.g., Bennett et al., 2008), and typically less severe symptoms of Autism. Preschool-aged children with ASD show a deficit at the neurological level in the discrimination of auditory syllables (Kuhl, Coffey-Corina, Padden, & Dawson, 2005), a deficit that can impact on speech and language development, such as by providing imperfect prototypes against which to compare auditory information in a task like the McGurk. Kuhl et al also found that degree of deficit was positively correlated with Autism severity. Thus individuals with AS, identified by minimally impaired early language impairments, and typically with less severe Autism symptoms, represent a potential test of the language-specific IMP deficit hypothesis for Autism (e.g, Bebko et al., 2006). Early language prototypes may have been less affected, leading to less cumulative impact for those with AS. Moreover, between 40% and 70% of individuals with a PDD have an intellectual disability (i.e., IQ < 70; Fombonne, 2005), and it is important to distinguish whether impaired IMP is associated with Autism, or is more associated with intellectual disability (Rogers & Ozonoff, 2005). In the present study, we tested a sample of children with DS. Little is known about how children with DS process audiovisual speech similar to TD peers. Bebko and colleagues (2006) found that children with intellectual impairments (including DS) performed similarly to those with typical development on a preferential looking audiovisual speech task.

Following from our earlier findings in Bebko et al. (2006), we hypothesized that a deficit in audiovisual intermodal speech perception would be unique to Autism, and expected that children with Autism would report less McGurk effect than the AS group, who in turn would be similar to the DS and TD groups. At the same time, we expected that children with Autism would be similar to the comparison groups in the unimodal auditory-only and visual-only conditions.

## Method

### *Participants*

Four groups of children ranging in age from 6 to 17 years (see Table 1) were recruited through schools, organizations for Developmental Disabilities, ASD agencies, and a research registry maintained for the Autism Spectrum Disorders - Canadian American Research Consortium (ASD-CARC; www.autismresearch.ca). All participants were free of any identified additional impairment (e.g., hearing, visual impairment), as reported by parents

------------------------------------

Insert Table 1 here

------------------------------------

The Autism group were15 children with a primary diagnosis of Autism with or without cognitive impairment. All had been diagnosed by a registered psychologist or psychiatrist and diagnoses were verified with the Autism Diagnostic Interview-Revised scores (ADI-R; Lord, Rutter, & LeCouteur, 1994; see Measures section), with classifications determined using the Autism Genetic Resource Exchange (AGRE, n.d.) affected status categories. Fourteen participants were in the Autism range. No score was available for one child; however the psychological report indicated a formal diagnosis of Autistic Disorder.

The AS group consisted of 15 children with a primary diagnosis of AS, confirmed through a review of psychological reports using DSM-IV-TR criteria (including absence of clinically significant early language development) and by scores on the Krug Asperger's Disorder Index (KADI; Krug, & Arick, 2003; see Measures section). KADI standard scores ranged from 74-114, comparable to the normative sample of individuals with AS ($M = 100.27$, $SD = 10.47$).

The DS group was 15 children with a diagnosis of Down Syndrome. The Typically Developing Group (TD) consisted of 19 children with no intellectual disability, ASD, or known psychiatric disorder .

A multivariate analysis of variance (MANOVA) compared age, language, Full Scale IQ (FSIQ), and Autism severity between groups (Table 2). No group main effect was found for age, $F(3, 60) = 1.51$, $p = .221$, and post hoc analyses revealed no significant differences in age between any two groups (all $p > .10$). As expected, significant differences between groups were found in FSIQ, $F(3, 60) = 42.74$, $p < .001$, $\eta^2 = .68$, receptive vocabulary, $F(3, 60) = 19.99$, $p < .001$, $\eta^2 = .49$, and expressive vocabulary, $F(3, 60) = 15.74$, $p < .001$, $\eta^2 = .44$. Post hoc tests revealed that the AS group had higher FSIQ scores, $\Delta M = 41.1$, $p < .001$, receptive $\Delta M = 35.3$, $p < .001$, and expressive vocabulary standard scores, $\Delta M = 25.2$, $p = .001$, than the Autism group. The TD group also had higher FSIQ scores, $\Delta M = 22.5$, $p < .001$, and receptive vocabulary standard scores, $\Delta M = 26.1$, $p$

= .005, than the Autism group. There were no significant differences in measured language or intellectual abilities between the Autism and DS groups (all $p > .10$).

------------------------------------

Insert Table 2 here

------------------------------------

Among the comparison groups, the TD group had significantly higher FSIQ scores, $\Delta M = 33.8$, $p < .001$, receptive, $\Delta M = 43.6$, $p < .001$, and expressive scores, $\Delta M = 28.1$, $p < .001$, than the DS group. The AS group had significantly higher FSIQ than the TD group, $\Delta M = 18.6$, $p = .002$, and significantly higher FSIQ scores, $\Delta M = 52.4$, $p < .001$, receptive, $\Delta M = 52.8$, $p < .001$, and expressive scores, $\Delta M = 42.3$, $p < .001$, than the DS group.

A one-way ANOVA comparing CARS scores among the AS, DS, and Autism groups yielded expected significant group effects, $F(2, 42) = 26.75$, $p < .001$, $\eta^2 = .56$. The Autism group had higher CARS scores than the AS, $\Delta M = 8.6$, $p < .001$, and the DS groups, $\Delta M = 10.4$, $p < .001$. No difference was found between the AS and DS groups ($p > .10$).

*Measures*

The *Wechsler Abbreviated Scale of Intelligence* (WASI; The Psychological Corporation, 1999) is a brief test of intelligence that yields a Verbal IQ, Performance IQ, and Full Scale IQ score.

*The Peabody Picture Vocabulary Test –III* (PPVT-III; Dunn & Dunn, 1997) is a screening test of single-word listening comprehension with excellent test-retest reliability (*Mdn* $r = .89$) and internal consistency (*Mdn* $r = .94$) (Sattler, 2001). Although only measuring one component of language ability for reasons of time, the PPVT has good concurrent validity with more global indices, such as the OWLS (.69 to .75) and the CELF Core Language (.72), as reported in the PPVT manual.

The *Expressive One-Word Picture Vocabulary Test–2000 Edition* (EOWPVT-2000; Brownell, 2000) requires children to provide names for a series of easy-to-interpret pictures. The manual reports split-half reliability of $r$ = .70 and test-retest reliability of $r$ = .80. The EOWPVT-2000 has good concurrent validity with more global indices, such as the OWLS and the CELF, with correlations with Total Language scores ranging .71-.85.

The *Childhood Autism Rating Scale* (CARS; Schopler, Reichler, & Renner, 1992) is a well-researched instrument to establish the level of severity of symptoms of Autism. The CARS was used with the DS group to confirm no presence of Autism, and with the AS and Autism groups to evaluate potential relations among severity and testing performance, although CARS applicability to those with AS is limited.

*The Autism Diagnostic Interview- Revised* (ADI-R; Lord, Rutter & LeCouteur, 1994) is a standardized, semi-structured clinical interview for caregivers of children and adults. It was administered only to parents of children with Autism to provide diagnostic information.

The *Krug Asperger's Disorder Index* (KADI; Krug, & Arick, 2003) is a parent-report scale that indicates the presence of behaviours indicative of AS, and is meant to discriminate between AS and high functioning Autism; KADI scores are significantly higher in those with AS relative to those with high-functioning autism (Krug & Arick, 2003). In a review of five diagnostic tools for AS (Campbell, 2005), the KADI had the strongest psychometric properties. It was used only with the AS group to corroborate AS diagnosis.

*Design*

Participants were tested either in a mobile lab or a land-based lab. In both cases, participants were seated in proximity to the screen to ensure that a consistent visual viewing angle was subtended. For the land-based lab, stimuli were digitally projected onto a 21 inches wide by 34 inches high screen, and sound was presented using the projector's centrally located speaker. In the mobile lab, stimuli from the same DVD recording were centered on a 27-inch LCD television screen. No effect of set-up was found across test conditions (all $p > .10$). Children were video-recorded during testing, and stimuli were only presented when they were looking at the screen.

*Stimuli:*

Tasks used were similar to the Facial speech test (FSP) used by de Gelder and colleagues (1991). Three conditions were used: auditory-only, visual-only, and audiovisual. A woman's face was recorded pronouncing a series of VCV syllables consisting of one of six plosive stops /p/, /b/, /t/, /d/, /va/, /ka/ or nasal /m/ or /n/ sounds, in between the vowel /a/ (e.g., /apa/ or /ana/). In the audiovisual condition, five discordant, yet temporally synchronous, auditory-visual combinations were created (e.g., the auditory */apa/* presented with the visual */ata/*). The audiovisual targets presented were (auditory-visual): /*apa*/-/*ata*/; /*aba*/-/*ada*/; /*ama*/-/*ana*/; /*ava*/-/*ada*/; and /*apa*/-/*aka*/. These five pairs were shown to cause visual capture or fusion at least 85% of the time during pilot testing with a nonclinical sample. In the auditory-only condition, the original auditory syllables (/ava/, /aba/, /ama/, /ada/) were combined with a stationary face with a neutral, closed mouth. In the visual-only condition, only visual articulations were presented, with no auditory signal. Speech-reading is a challenging task, so the articulations used were limited to the coronal consonants among the full set of stimuli (i.e., /ada/, /ata/, /ana/). Tone of voice used and the duration of clip were consistent across trials. In the audiovisual task, the tracks were synchronized such that the release of the consonant occurred simultaneously across the audio and visual tracks.

*Condition set*. Sets were created to ensure an evenly distributed order of presentation. In the audiovisual condition, each set presented the 5 auditory-visual pairs three times in a row (resulting in 15 presentations per set) to ensure that children with difficulty maintaining attention had the opportunity to see the stimuli before moving on. Three different audiovisual condition sets were developed, resulting in a total of 45 presentations of the McGurk stimuli in the experimental session. Similar condition sets were created for the auditory-only and visual-only conditions. These sets each contained four sets of three presentations (resulting in 12 presentations per set), and 36 presentations per experimental session. Therefore, each participant saw 117 presentations in total (45 + 36 + 36).

*Procedure*

Two master sequences were created to counterbalance order. Prior to beginning, participants were presented orally with a sound that did not occur during the experiment, and were asked to repeat it to ensure that children could repeat speech sounds and to familiarize them with the task. After each presentation in the experiment, the display was paused and the participant was asked what the woman had said. For many participants no instruction was required after the first trial.[1] Responses were recorded and the

[1] Due to experimental error the participants in the AS group were provided with varied instructions. Five participants in the AS group were asked what they had heard, and eight were given a hybrid instruction of "Tell me what you heard; what did the woman say?" While a "heard" instruction can sometimes decrease audio-visual fusion by drawing more attention to the auditory information (e.g., Summerfield and McGrath, 1984), any effect was minimal, as the AS group was not significantly different from the other comparison

experimenter initiated the next sound. Only responses that occurred while participants were looking at the screen during stimulus presentation were deemed as eligible trials. Trials were considered ineligible if participants looked away from the screen during the trial, if they were out of their seat, or if they were speaking during the video presentation.

## Results

Separate one-way ANOVAs were calculated to test whether groups differed in the number of eligible trials per condition (Table 3). There were no significant group main effects in the auditory-only condition, $F(3, 63) = 1.63$, $p = .193$, the visual-only condition, $F(3, 63) = 1.68$, $p = .182$, or the audiovisual condition, $F(3, 63) = 2.27$, $p = .090$.

------------------------------------

Insert Table 3 here

------------------------------------

Analyses of responses were based on percent of correct responses of the total number of eligible trials (i.e., accuracy rates) for visual-only and auditory-only conditions. Visual-only accuracy was measured by correct place of articulation, because visual information alone is often insufficient to discriminate phonemes *within* a category that varies in manner of articulation or voicing. Therefore, if respondents nominated any coronal consonant (e.g., /z, s, t, d, n/) it was considered correct. For the audiovisual condition, percent of responses representing visual capture or fusion responses out of the total number of eligible trials was used.

Mean response rates by group (see Table 4) were analysed in a 3 X 4 repeated measures ANOVA, with condition as a within-subject factor (auditory-only, visual-only,

---

groups (TD & DS) in reported fusion responses (see Results, Table 3). Further, there were no significant differences in performance between those in the AS group who received the "heard" instruction, and those who did not.

and audiovisual) and group as a four-level between-subject factor, using Greenhouse-Geiser epsilon adjustments. A significant group X condition interaction emerged, $F(5.44, 108.78) = 3.06$, $p = .010$, $\eta^2 = .13$. There were also significant main effects for group, $F(3, 60) = 5.21$, $p = .003$, $\eta^2 = .21$, and condition, $F(1.81, 108.78) = 13.20$, $p < .001$, $\eta^2 = .18$. Post hoc contrasts with Sidak-Bonferroni adjustment verified that the Autism group, as hypothesized, showed significantly less McGurk effect than the AS group, $\Delta M = -36.5\%$, $p = .005$, the DS group, $\Delta M = -41.8\%$, $p = .001$, and TD group, $\Delta M = -44.9\%$, $p < .001$. There were no other significant differences in McGurk performance between groups (all $p > .95$). For the unimodal conditions, differences were not significant across groups (all $p > .16$).

-------------------------------

Insert Table 4 here

-------------------------------

Overall the pattern of group differences remains consistent across audio-visual pairs, with the Autism group experiencing lower rates of McGurk effect than the other three groups. The ABA/ADA combination produced the highest McGurk response rate across all groups. There was more variability across groups regarding the pair that elicited the lowest McGurk response rate.

-------------------------------

Insert Table 5 here

-------------------------------

The significant difference in McGurk effect between DS and Autism groups suggests that intellectual functioning was not the reason for the differential in the McGurk condition, particularly since IQ scores for the DS group were lower than the Autism group, yet the DS group's performance was more similar to that of the TD group. However, since the Autism group had a wider IQ range than the DS group, we further examined intellectual functioning within the Autism group by dividing the 15 participants based on a median split of IQ scores, with participants with FSIQ greater than or equal to 82 termed a higher-functioning Autism subgroup (HFA, n = 8) and those with FSIQ less than 82 a lower-functioning Autism subgroup (LFA, n = 7). Responses were analysed in a 2 X 3 repeated measures ANOVA (group X condition) where group was now a two-level between-subject factor (LFA, HFA). No main effect of group, $F(1, 13) = 1.54$, $p = .240$, or interaction effect, $F(1.78, 23.10) = 3.02$, $p = .073$, was found, indicating similar performance for both subgroups. A significant main effect of condition type was found, $F(1.78, 23.10) = 29.55$, $p < .001$.

Within-group Pearson product moment correlation coefficients were calculated to determine the relations among responses in the unimodal and audiovisual conditions (see Table 6). Correlations among auditory-only and audiovisual conditions were strong and in the negative direction, with significance levels for the TD and the AS groups at $p <.05$ and marginal for the DS and Autism groups ($p = .097$ and $p = .075$, respectively). For each group, the shared variance ranged from 20-36% between the auditory-only and audiovisual conditions. Correlations were in the positive direction between visual-only accuracy and audiovisual conditions. They were significant for the DS group ($p<.01$) and nearly so for the Autism group ($p = .051$), indicating shared variance of 50% and 26%, respectively. Correlations were more moderate and not significant, for the TD and AS groups (both $p >$ .10), although shared variance was approximately 13%.. Across group Pearson product moment correlation coefficients were also calculated to determine the relations between McGurk performance and a number of sample characteristics. Non-significant positive correlations were found between audiovisual performance and age, IQ, receptive and expressive vocabulary (all $r(64) < .14$, all $p > .257$). CARS scores were significantly negatively correlated with McGurk performance, $r(45) = .49$, $p$<001); that is, the greater the severity score on the CARS, the less the McGurk effect was reported.

-------------------------------

Insert Table 6 here

-------------------------------

To control for individual differences in unimodal performance, which had linear relationships to performance in the audiovisual condition, an analysis of covariance was conducted, with unimodal performance as covariate. As expected based on the correlations summarized in Table 6, auditory-only accuracy was found to be a significant covariate with rate of McGurk effect, $F(1, 58) = 22.38$, $p < .001$, $\eta^2 = .28$, as was visual-only accuracy, $F(1, 58) = 15.42$, $p < .001$, $\eta^2 = .21$. Nonetheless, after controlling for the covariates, group differences in rate of McGurk effect remained, $F(3, 58) = 7.91$, $p < .001$, $\eta^2 = .29$. Post hoc contrasts with adjusted means verified that the Autism group showed significantly less McGurk effect than the DS group, $\Delta M$ = -31.4%, $p < .001$, TD group, $\Delta M$ = -34.7%, $p < .001$, and AS group, $\Delta M$ = -24.8%, $p$ =.004. No significant differences were found among the AS, TD, and DS groups (all p > .10).

## Discussion

This is the first study to directly investigate intermodal speech processing abilities in individuals with Asperger Syndrome compared to Autism, as differentiated in the DSM-

IV-TR. The hypothesis that children with Autism would demonstrate lower rates of intersensory fusion (McGurk effect) than children with AS and other comparison groups was supported: When presented with discordant auditory and visual speech signals, the Autism group reported significantly less visual capture or fusion of the auditory and visual inputs than the AS group, even after controlling for differences in unimodal performance. We also compared the children with AS to typically developing peers, and found that they showed similar rates of McGurk effect.

Performance in the audiovisual condition was related to unimodal conditions in predictable ways. The more accurate participants were in the auditory-only condition, the less they perceived the McGurk effect. Conversely, the better they were at speech reading, the more they perceived the McGurk effect. This auditory-audiovisual relation is consistent with research with TD adults, where decreasing an auditory signal's intelligibility during an audiovisual speech event leads to increased reporting of the McGurk, as participants rely more on visual information (Colin, Radeau, & Deltenre, 2005). It is also consistent with the notion of visual capture for similar conflicting bimodal events (Desjardins et al., 1997).

Performance differences in unimodal conditions could not entirely account for the reduced perception of the McGurk effect. When differences between the groups on the unimodal tasks were covaried, the Autism group continued to show significantly lower rates of the McGurk effect compared to all other groups. The difference for the Autism group also could not be attributed to differential rate of orienting to the stimuli, as there was no difference in eligible trials (that is, those where participants were orienting to the stimuli versus looking away) between the Autism and other groups. Nonetheless, we took an additionally conservative approach in the analyses, basing them only on eligible trials as an additional control for rate of orienting to the stimuli.

We found that in the auditory unimodal condition, children with Autism were equally accurate as the other groups, in contrast with Williams and colleagues' (2004) findings of poorer performance. In both studies the visual-only condition yielded lower scores for those with Autism. In contrast to Williams et al., when we controlled for unimodal performance, the Autism group continued to show significant differences in the bimodal condition, indicating that integration of audio and visual inputs in speech processing is impaired in Autism. Similar decreased bimodal results were reported in Irwin et al. (2011), although they failed to evaluate or covary auditory unimodal performance.

The difference compared to Williams et al. (2004) may be related to the more tightly controlled sample characteristics in the present study or to the instructions given. Williams and colleagues included children described as being on the Autism spectrum, and their verbal IQ estimates suggest that high functioning ASD comprised most of the sample. This combination of a potentially wider range of diagnoses, and with higher IQs, would allow for a high proportion in their sample of participants with AS, who we found to have similar McGurk performance to the comparison groups. Their broader ASD group could then have led to a reduction in differences between their ASD and TD groups. This emphasizes the importance of well-defined diagnostic groups, and the corresponding importance of appropriate control groups.

A study by Iarocci et al. (2010) used more tightly controlled samples and found results similar to Williams et al. (2004). However, a key difference between those studies and the present one is that in both cases, there were explicit instructions in the bimodal condition to report what was *heard*. Summerfield and McGrath (1984) demonstrated that instructions to report auditory information in bimodal tasks decrease the influence of visual fusion by directing attention away from the visual input, compared to more neutral instruction to report what the speaker said. In the present study, participants were given a more neutral instruction, except for a subset of those in the AS group only. This experimental error may have been serendipitous, in that it provides indirect corroboration for the Summerfield and McGrath findings. The overall AS group data contained approximately a third of participants who received the "heard" instruction only, and their data indicates slightly less audio-visual fusion that the other non-Autism groups, but not significantly so. In both the previous studies, the ASD samples may have reported less visual fusion for the bimodal stimuli due to the instruction for all participants to report what was heard. The lowered response levels would then be further reduced when response rate in the visual-only condition was controlled. This combination could result in the masking of differences that may have been larger with more neutral instructions. Irwin et al. (2011) also gave instruction to focus on what was *heard* but did show significantly less visual influence for the ASD sample in the bimodal condition, similar to the present study; however, they did not covary unimodal contributions, and their sample was a combination of Autism and ASDs. Therefore, differences in samples and instructions may well have contributed to the different findings among the studies, and reinforces the importance of well-defined samples and appropriate instructions.

There are a number of factors that may contribute to the observed differences in the bimodal condition between the Autism and AS groups in the present study. These include the early history of more significant language delays that has been used in DSM-IV-TR to help differentiate these groups diagnostically, differences in basic cognitive or perceptual processing, or other characteristics not associated with the language or cognitive/perceptual differences measured in this study.

The language issue is a complex one as language differences between AS and Autism may be more difficult to identify in some individuals after early development, as illustrated in the ranges of language scores in our samples, and in the ongoing discussions about the early language criterion for the pending DSM-5 (see www.dsm5.org/ProposedRevisions). It is not simply a language delay that is affecting the Autism group's performance. The DS sample had even greater delays, yet their McGurk scores were similar to the other comparison groups. The language skills measured were limited to expressive and receptive vocabulary skills to keep testing sessions as compact as possible. More sensitive or global measures may have identified other differences, although the measures used show excellent concurrent validity with more comprehensive tests. Rather than a simple delay, if language variables are associated, it seems that it is aspects of the language disorder in Autism, such as the socio-pragmatic-communication impairment, that is associated with the reduced McGurk effect. Individuals with AS also often have socio-pragmatic-communication impairments although typically to a lesser degree, so it may be the degree of such impairments that affects performance on the McGurk task. Results from Kuhl et al (2005) are consistent with this view: They found strong correlations between their measure of speech discrimination and a social measure in a group of preschoolers with ASD, and the degree of atypicality in the speech processing scores was strongly associated with severity of Autism.

The second variable suggested as contributing to the observed Autism-AS effect was differences in basic cognitive or perceptual processing. We tested a sample of children with DS, and found that despite having an intellectual disability and lower vocabulary scores, the DS group‘s performance on McGurk and unimodal tasks was relatively unaffected, being similar to the AS and TD groups. These similarities between the DS and TD comparison groups are consistent with the previous findings of Bebko et al., (2006) using the preferential looking methodology with a mixed sample of children with intellectual disabilities. To further examine the role of intellectual skills, when the Autism group was median split into high and low IQ subgroups, no significant differences were found on the McGurk task. Therefore, it is unlikely that perceptual fusion is being affected by general cognitive impairment in the children with Autism.

Other perceptual or cognitive factors may play a role. When presented with a discordant auditory speech signal and visual articulation (e.g., /ada/ and /aga/), each source of information is assumed to be evaluated against prototypes in memory and a match is made (Massaro, 1998; Williams, et al., 2004). Features matching each prototype are then integrated and, based on the relative degree of support for each possibility, a perceptual decision is made. If children with Autism have difficulty evaluating auditory and visual speech signals against co-registered prototypes, then audiovisual speech perception would be impacted. Irwin et al. (2011) concluded that children with ASD exhibit particular difficulty with processing of audio-visual phonetic information, which is consistent with this view. On the auditory side, it is known that preschool-aged children with ASD show a deficit at the neurological level in the discrimination of auditory syllables, with degree of difficulty positively correlated with Autism severity (e.g. Kuhl, Coffey-Corina, Padden, &

Dawson, 2005). Such a deficit could lead to impaired speech representations, thus providing imperfect prototypes against which to compare auditory information in the McGurk task. On the visual side, children with Autism are known to be deficient in discriminating and recognizing faces (Boucher, Lewis, & Collis, 2000). These differences in discriminating auditory and visual information, and relating them to stored prototypes, are consistent with the view of Autism as characterized by a weak central coherence tendency: that is, a weak tendency to combine stimulus details or features into coherent wholes, and relate them to previous experience (Happe & Frith, 2006). The investigation of speech prototype development (e.g., Williams et al., 2004) may be useful starting points in further determining where in the process of audiovisual integration children with Autism have difficulty.

The current methodology was an adaptation of de Gelder and colleagues' (1991) work and related paradigms examining the McGurk effect (e.g., Colin et al., 2005; Desjardins et al., 1997). Other researchers (e.g., Massaro & Cohen, 2000) have recommended the addition of a confusion matrix design to help isolate audiovisual integration ability from unimodal performance in McGurk tasks. Although the current approach yielded a clear picture of McGurk performance in children with Autism, AS, and DS, a confusion matrix design could be useful as a next step to examine potential mechanisms underlying the different performance in Autism. Similarly, using a wider range of phonemes may help clarify the inconsistent contribution of the visual-only (speech-reading) findings in the various studies.

A comprehensive understanding of audiovisual integration difficulties in speech can open the door to new interventions. For example, a computer-based training program has been used with children with ASD and increased the rate of the McGurk effect reported (Williams et al., 2004; Massaro & Bosseler, 2006). Furthermore, the McGurk effect has been detected in infants as young as 4.5 months (Burnham & Dodd, 2004). If a corresponding early detection of an intermodal speech deficit proves reliable, then early behavioral interventions may be helpful in increasing the child's attention to meaningful social stimuli, including faces and speech, and increase the child's motivation to engage in face-to-face interactions. This refocusing of attention could in turn influence the neural plasticity of the face and speech processing systems.

The results of the present study, using an entirely different research paradigm, corroborate the findings by Bebko et al. (2006) of an audio-visual integration deficit specific to speech stimuli in autism. The difficulties for the autism group do not appear to be shared by children with intellectual disabilities, including those with Down syndrome, or typically developing children. Moreover the integration skills involved in processing speech appear to be intact in the Asperger sample tested in the present study, suggesting that the deficit may not characterize the entire autism spectrum, but be limited to those with characteristics associated with our autism sample (e.g., symptomatology associated with elevated ADI-R and CARS scores). More research is needed to better understand the perceptual and cognitive processes underlying intermodal speech perception involved in the McGurk effect vs single modality perception which, in turn, may help us understand *why* children with Autism differ from other children in speech-related intermodal perception.

## Acknowledgements

We would like to thank Nicole Aliya Rahim, Lisa Hancock, Carly McMorris, Maggie Slusarczyk, and Ksusha Blacklock for help in data collection. Special thanks to all the parents and children involved in this study, as well as the organizations who assisted in recruitment. A special thank you to Melissa Hudson, and the late Jeanette Holden for their generous assistance with the ASD-CARC Research Registry. This research was supported by Canadian Graduate Scholarship Master's and Doctoral Awards to the second author from the Canadian Institutes of Health Research (CIHR), a Research Studentship to the third author from Autism Ontario and the Ontario Mental Health Foundation, and a CIHR Interdisciplinary Health Research Team grant (RT-43820) to the second author and third author through the ASD-CARC consortium (JJAH, principal investigator; www.autismresearch.ca).

References

AGRE Affected Status Categories. (n.d.). Retrieved February 1, 2009, from http://www.agre.org/agrecatalog/algorithm.cfm

American Psychiatric Association. (2000). *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision*. Arlington, VA: American Psychiatric Publishing, Inc.

Bebko, J. M., Weiss, J. A., Demark, J. L., & Gomez, P. (2006). Discrimination of temporal synchrony in intermodal events by children with Autism and children with developmental disabilities without Autism. *Journal of Child Psychology and Psychiatry, 47*, 88-98.

Bennett, T. A., Szatmari, P., Bryson, S. E., Volden, J., Zwaigenbaum, L., Vaccarella, L., Duku, E. & Boyle, M. H. (2008). Differentiating autism and Asperger syndrome on the basis of language delay or impairment*. Journal of Autism and Developmental Disorders, 38,* 616-625.

Boucher, J., Lewis, V., & Collis, G. M. (2000). Voice processing abilities in children with Autism, children with specific language impairments, and young typically developing children. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 41,* 847-857.

Brownell, R. (2000). *Expressive one-word picture vocabulary test* (2000 ed.). Novato, CA: Academic Therapy Publications.

Campbell, J.M. (2005). Diagnostic assessment of Asperger's Disorder: A review of five third-party rating scales. *Journal of Autism and Developmental Disorders, 35(1)*, 25-35.

Colin, C., Radeau, M., & Deltenre, P. (2005). Top-down and bottom-up modulation of audiovisual. integration in speech. *European Journal of Cognitive Psychology, 17*, 541-560.

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.

Desjardins, R. N., Rogers, J., & Werker, J. F. (1997). An exploration of why preschoolers perform differently than do adults in audiovisual speech perception tasks. *Journal of Experimental Child Psychology, 66*, 85-110.

Dunn, L. M., & Dunn, L. M. (1997). *Peabody Picture Vocabulary Test – III*. Circle Pines, MN: American Guidance Services.

Fombonne E. (2005). Epidemiological studies of Autism and pervasive developmental disorders. In F. Volkmar (ed.). *Handbook of Autism and Pervasive Developmental Disorders.* 3rd ed. (pp. 42 – 69) New York, NY: Wiley & Sons.

Happe, 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.

Irwin, J.R., Brancazio, L.,Tornatore, L.A. & Whalen, D.H. (2011) Can Children With Autism Spectrum Disorders ''Hear'' a Speaking Face? *Child Development, 82*, 1397-1403.

Keane, B.P., Rosenthal, O., Chun, N.H. & Shams, L. (2010) Audiovisual integration in high functioning adults with autism. *Research in Autism Spectrum Disorders, 4* 276–289

Klin, A. (1991). Young autistic children's listening preferences in regard to speech: A possible characterisation of the symptom of social withdrawal. *Journal of Autism and Developmental Disorders, 21*, 29-42.

Krug, D.A., & Arick, J.R. (2003). *Krug Asperger's Disorder Index*. Austin, Texas: Pro-Ed Inc.

Kuhl, P. K., Coffey-Corina, S., Padden, D., & Dawson, G. (2005). Links between social and linguistic processing of speech in preschool children with Autism: behavioural and electrophysiological evidence. *Developmental Science, 8,* F1-F12.

Kuhl, P. K., & Meltzoff, A. N. (1984). The intermodal representation of speech in infants. *Infant Behavior and Development, 7*, 361-381.

Lord, C., Rutter, M., & LeCouteur, A. (1994). Autism Diagnostic Interview-Revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders, 24*, 659-685.

Massaro, D.W. (1998). *Perceiving Talking Faces: From Speech Perception to a Behavioural Principle*. Cambridge, MA: MIT Press.

Massaro, D.W., & Bosseler, A. (2006). Read my lips. *Autism, 10(5)*, 495-510.

Massaro, D.W., & Cohen, M.M. (2000). Tests of auditory-visual integration efficiency within the framework of fuzzy logical model of perception. *Journal of the Acoustical Society of America, 108(2)*, 784-789.

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature, 264,* 746-748.

Mongillo, E.A., Irwin, J.R., Whalen, D.H., Klaiman, C., Carter, A.S., & Schultz, R.T. (2008). Audiovisual processing in children with and without autism spectrum disorders. *Journal of Autism and Developmental Disorders, 38 (7)*, 1349 –1358.

Rogers, S. J., & Ozonoff, S. (2005). Annotation: What do we know about sensory dysfunction Autism? A critical review of the empirical evidence. *Journal of Child Psychology and Psychiatry 46*, 1255–1268.

Sattler, J. M. (2001). *Assessment of children: Cognitive applications*, 4th Ed. San Diego, CA: Jerome M. Sattler, Publisher, Inc.

Schopler, E., Reichler, R. J., & Renner, B. R. (1992). *The childhood Autism rating scale (CARS).* Los Angeles: Western Psychological Services.

Summerfield, Q., & McGrath, M. (1984). Detection and resolution of audio-visual incompatibility in the perception of vowels. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology 36*, 51–74.

The Psychological Corporation. (1999). *Weschler Abbreviated Scale of Intelligence*. San Antonio: Author.

Williams, J. H. G., Massaro, D. W., Peel, N. J., Bosseler, A., & Suddendorf, T. (2004). Visual auditory integration during speech imitation in Autism. *Research in Developmental Disabilities, 25*, 559-575.

Table 1 Characteristics of samples.

| Group | Age (Years) | CARS | Full Scale IQ score | Receptive Vocabulary | Expressive Vocabulary |
|---|---|---|---|---|---|
| Autism (*n* = 15) | | | | | |
| *M* | 10.53 | 34.10 | 76.53 | 81.80 | 89.27 |
| *SD* | 2.46 | 6.08 | 19.68 | 30.31 | 27.43 |
| Range | 6.58 – 14.66 | 26-45 | 52 - 105 | 40 – 134 | 55 - 136 |
| AS (*n* = 15) | | | | | |
| *M* | 11.89 | 25.53 | 117.67 | 117.13 | 114.47 |
| *SD* | 2.65 | 2.72 | 15.34 | 19.95 | 12.26 |
| Range | 6.25 – 15.42 | 21-30 | 89 – 141 | 81 - 143 | 95 – 141 |
| TD (*n* = 19) | | | | | |
| *M* | 10.21 | N/A | 99.05 | 107.89 | 100.26 |
| *SD* | 2.74 | | 9.51 | 19.28 | 12.23 |
| Range | 6.00 – 15.67 | | 81 – 119 | 78 – 153 | 81 – 119 |
| DS (*n* = 15) | | | | | |
| *M* | 11.96 | 23.67 | 65.27 | 64.33 | 72.20 |
| *SD* | 3.96 | 2.77 | 9.72 | 11.94 | 14.81 |
| Range | 6.00 – 17.20 | 20-29 | 50 – 84 | 40-83 | 55 – 110 |

Table 2: Sample characteristics: Tests of group differences

| | *F(df)* | *p* | $\eta^2$ | *Bonferroni Post Hoc Group Differences* |
|---|---|---|---|---|
| Age | $F(3, 60) = 1.51$ | .221 | | Not significant |
| CARS | $F(2, 42) = 26.75$ | < .001 | .56 | AUT > AS, $p < .001$<br>AUT > DS, $p < .001$ |
| FSIQ | $F(3, 60) = 42.74$ | < .001 | .68 | TD > AUT, $p < .001$<br>TD > DS, $p < .001$<br>AS > AUT, $p < .001$<br>AS > DS, $p < .001$<br>AS > TD, $p = .002$ |
| Receptive Vocabulary | $F(3, 60) = 19.99$ | < .001 | .49 | TD > AUT, $p = .005$<br>TD > DS, $p < .001$<br>AS > AUT, $p < .001$<br>AS > DS, $p < .001$ |
| Expressive Vocabulary | $F(3, 60) = 15.74$ | < .001 | .44 | TD > DS, $p <.001$<br>AS > AUT, $p = .001$<br>AS > DS, $p <.001$ |

.

Table 3: Percentage of eligible trials by group.

| | Audiovisual Condition | Auditory-Only Condition | Visual-Only Condition |
|---|---|---|---|
| Autism | 87.11 | 97.03 | 92.11 |
| Asperger Syndrome | 96.16 | 97.78 | 91.85 |
| Typically Developing | 94.98 | 100 | 99.81 |
| Down Syndrome | 91.84 | 99.25 | 93.33 |

Table 4

<u>Rates of auditory-only and visual-only accuracy, and of visual capture in the in the audiovisual condition (McGurk effect) among participant groups (in percent).</u>

| | Audiovisual Condition *M (SD)* | Auditory-Only Condition *M (SD)* | Visual-Only Condition *M (SD)* |
|---|---|---|---|
| Autism | 24.49 (21.95) | 82.92 (11.77) | 31.54 (29.09) |
| Asperger Syndrome | 55.86 (30.35) | 78.97 (23.07) | 58.36 (25.56) |
| Typically Developing | 69.47 (33.24) | 77.92 (24.31) | 58.42 (28.64) |
| Down Syndrome | 62.33 (29.28) | 72.61 (19.99) | 41.02 (31.72) |

Table 5: Rates of visual capture (McGurk effect) of each audiovisual stimulus pairing among participant groups (in percent).

| | | Audio/Visual Stimulus Pair | | | | |
|---|---|---|---|---|---|---|
| | Total Rate | AVA/ ADA | APA/ ATA | ABA/ ADA | APA/ AKA | AMA/ ANA |
| Autism | 24.49 | 30.37 | 21.43 | 33.33 | 28.52 | 12.59 |
| Asperger Syndrome | 55.86 | 55.93 | 40.74 | 75.19 | 38.52 | 68.52 |
| Typically Developing | 69.47 | 54.24 | 68.52 | 80.70 | 75.15 | 68.71 |
| Down Syndrome | 62.33 | 62.96 | 49.63 | 82.59 | 52.22 | 61.48 |

Table 6 Correlations among conditions for each group.

| | | Auditory-only condition | Visual-only condition |
|---|---|---|---|
| Visual-only condition | Autism | *r*(15) = .09 | |
| | AS | *r*(15) = -.19 | |
| | TD | *r*(19) = -.24 | |
| | DS | *r*(15) = -.25 | |
| Audiovisual condition | Autism | *r*(15) = -.47^ | *r*(15) = .51^ |
| | AS | *r*(15) = -.61* | *r*(15) = .38 |
| | TD | *r*(19) = -.57* | *r*(19) = .36 |
| | DS | *r*(15) = -.45^ | *r*(15) = .71** |

* = p < .05, ** = p < .01, ^ = approaching significance