AUDITORY AND AUDIOVISUAL SPEECH PERCEPTION IN CHILDREN WITH AUTISM SPECTRUM DISORDERS

Julia R. Irwin
Haskins Laboratories
New Haven, Connecticut 06511

Autism Spectrum Disorders (ASD) are neurodevelopmental disorders along a continuum of severity that are generally characterized by marked deficits in social and communicative functioning (American Psychiatric Association, 2000). Whether due to increased incidence or more effective detection, the diagnosis of young children with autism is increasing (Fombonne, 2003 a, b). The number of children with autism spectrum disorders is an emerging public health crisis, with the Centers for Disease Control and Prevention reporting that one in 150 children who are born this year will be affected, an increase of 172% in diagnosis over the past decade (Centers for Disease Control and Prevention, 2007).

"Results from this area of inquiry will move the field forward significantly by providing basic data to guide focused interventions related to speech perception and processing in the population of children with ASD."

Autism diagnostic criteria

Of the autism spectrum disorders, the most prominent is autism, which is characterized by the presence of restricted or repetitive behaviors in addition to the social and communicative deficits mentioned above. This constellation of behaviors must emerge prior to three years of age for a child to receive a diagnosis of autism (American Psychiatric Association, 2000).

A core deficit of autism is a significant deficit in social reciprocity. These social difficulties manifest themselves behaviorally in a number of ways: poor modulation of non-verbal social behaviors, such as eye-to-eye gaze for the purpose of exchanging social affect and intent, lack of social and emotional reciprocity, and an absence of shared enjoyment with others (Lord *et al.*, 1989). This profile of behaviors results in significant problems in the development of peer relationships (American Psychiatric Association, 2000).

From the earliest points in development, typically developing children show a clear interest in those around them. In contrast, deficits in early social responsivity are present in infants and toddlers who go on to receive a diagnosis of autism. For example, failure to orient when one's name is called is an early "red flag" behavior associated with a later diagnosis (Landa, 2007). This lack of response to such a salient social signal is a powerful example of early deficits in social reciprocity. Accordingly, later in childhood and through adulthood, individuals with autism are less likely to seek out and engage in social play with peers than typically developing children (Lord and Magill-Evans, 1995).

Individuals with autism also exhibit significant delays in the development of communication in both non-verbal and verbal communicative behavior. Deficits in non-verbal communication include a reduced amount of both manual gestures, e.g., in the use of pointing to bring one's attention to either a proximal or distant object (Baranek, 1999) and in pre-linguistic vocalization, such as babbling and early vocal play (Landa, 2007). Moreover, the production of spoken language is strongly affected in this population. An estimated 25% of children with autism never develop functional language skills (Klinger *et al.*, 2002). Even those affected children who develop the ability to communicate verbally exhibit delays in both language acquisition and use such that they begin to speak later and at a significantly slower rate than their typically developing peers (Lord and Paul, 1997). For those individuals with autism who develop spoken language, the initiation of and ability to maintain conversation is typically severely restricted. Other common features of the language of individuals with autism are stereotyped or idiosyncratic patterns. For example, the speech of individuals with autism may be high in pitch, uninflected and robot-like or sing-song in nature (Rapin, 2001). Echolalic speech is repetition of the speech of others that sometimes appears to be used communicatively (Loveland and Tunali-Kotoski, 1997). Stereotyped speech refers to highly repetitive, highly specific language that is often centered on inappropriately formal and arbitrary topics (Lord *et al.*, 1989). Stereotyped speech can interfere with the child's overall quality of social interactions, not only because the topic is unusual and highly specific to an individual, but also because it is typically not well integrated into conversation (Lord *et al.*, 1989).

Finally, individuals with autism display restricted, repetitive or stereotyped behaviors. These behaviors can be manifested in a number of ways, such as intense preoccupations with unusual interests or parts of objects. For example, an affected child may focus exclusively on spinning the wheels of a toy car, or on the way one part of the car feels or smells, as opposed to playing with it as it was intended (Lord *et al.*, 1989). Affected individuals also may engage in repetitive, ritualized or routinized behaviors that are not required to complete the task at hand. Often, the child will feel anxious if this ritual is disrupted or blocked from being carried out. For example, a child may insist on reciting the names of his/her classmates every time they report about their day at school (Lord *et al.*, 1989). Finally, individuals with autism may exhibit stereotyped motor mannerisms, e.g., hand flapping or finger flicking or twisting (Lord *et al.*, 1989).

Although the set of core behaviors described above are associated with a diagnosis of autism (American Psychiatric Association, 2000), Lord and colleagues (2000) note that profiles of behavior will vary for affected individuals, and even within an individual, with the behavioral repertoire changing as a child develops.

Pervasive developmental disorders

In addition to autism, there is a range of related disorders that share the primary deficits in social reciprocity and communication. The disorders that fall under this general umbrella of Pervasive Developmental Disorders (PDD) (American Psychiatric Association, 2000) include Childhood Disintegrative Disorder, Asperger syndrome, Rett syndrome and Pervasive Developmental Disorder—Not Otherwise Specified (Lord *et al.*, 2000).

Childhood Disintegrative Disorder (CDD) differs from autism in that affected children exhibit a serious regression in multiple areas of functioning following normal development for at least two years. The profile of behavior shown by children with CDD includes a significant loss of interest in the environment and lack of social and emotional reciprocity and language skills. In addition, repetitive motor movements such as hand flapping and finger flicking emerge in this population. Children also exhibit restricted interests and activities (Lord, *et al.*, 2000). CDD is usually associated with mental retardation in childhood consistent with the significant, progressive loss of skills in all areas of functioning (American Psychiatric Association, 2000).

Another disorder under the PDD umbrella, Asperger syndrome, is closely related to autism. While individuals affected with Asperger syndrome share deficits in social reciprocity and interaction, cognitive and linguistic skill is relatively spared, with no evidence of significant delays in these areas in early development (American Psychiatric Association, 2000). Although individuals affected with Asperger syndrome are interested in the social environment, they are impaired in the ability to read social cues and show great difficulty with pragmatic rules when communicating, e.g., by speaking in a particularly loud voice or discussing socially inappropriate topics.

Rett syndrome, a pervasive developmental disorder which primarily affects girls, shares common characteristics with autism early in development, including deficits in social, cognitive, and language skills, as well as stereotyped behaviors (Glaze, 2004). However, there are several symptoms specific to Rett syndrome, including loss of hand skills, problems with gait, deceleration of head growth, and breathing and motor problems that allow it to be differentiated as a distinct disorder from autism (Lord *et al.*, 2000; Mount *et al.*, 2003).

Finally, Pervasive Developmental Disorder—Not Otherwise Specified is a category for those with deficits in the areas of social interaction, communication, or restricted/repetitive interests, but who do not meet the criteria for autism because of atypical time of onset, profile or severity of symptoms (American Psychiatric Association, 2000).

Autism and language

The range of functioning for individuals with autism and related disorders varies from significantly impaired non-verbal individuals with mental retardation to individuals with above-average cognitive functioning and fluent linguistic skill. To describe the level of functioning within this range, researchers often characterize an individual's behavior as low- or high-functioning (e.g., Jansiewicz *et al.*, 2006). High functioning children with autism have IQ's of 80 or above and the ability to speak (Williams *et al.*, 2006). By extension, an individual who is low functioning will be in the mentally retarded range cognitively, without the ability to produce speech. Because the ability to speak is a significant skill in typical development, one important difference between low- and high-functioning status in individuals with autism appears to be whether or not an individual has productive speech. Moreover, the development of productive speech prior to the age of 5 is associated with better prognosis with respect to symptomatology (Rutter, 1983).

Auditory speech perception

The ability to perceive and produce speech is foundational in human communication from the earliest points in development. Thus, fundamental, basic problems in the processing of speech sounds create significant deficits in language and social functioning throughout development. One likely source of the marked impairment in language across the range of individuals with autism is a deficit in speech perception; however, the role of perceptual processing in this population has not been extensively examined. The available evidence indicates that individuals with ASD show atypical responses to speech and vocal sounds, (Gervais *et al.*, 2004). In addition, affected individuals show deficits in understanding (receptive language functioning, Koning and Magill-Evans, 2001) and producing language (expressive language functioning, Howlin, 2003). Further, children with ASD have atypical asymmetry in brain regions associated with language, with greater volume in the right hemisphere, rather than the left (Herbert *et al.*, 2002). Finally, visual influence on heard speech appears to be impaired in children with ASD (DeGelder *et al.*, 1991; Massaro and Boesseler, 2003; Williams *et al.*, 2004). In particular, because of their severe limitations in productive language, non-verbal individuals with autism have been vastly under-represented in perceptual studies of speech. Because language ability is the key prognostic factor for long-term outcomes among children and adults with ASD (Lord and Venter, 1992), it is critical that we examine speech perception and processing in this population so that we may better understand the mechanisms that underlie delays and deficits in language acquisition.

Neurobiological techniques for studying speech perception in children with ASD

A number of informative neurobiological techniques have recently been applied to the study of individuals with autism and related disorders, making the study of even severely affected individuals possible. Event-related potentials (ERP) provide a direct measure of neural processing that can help reveal any potential underlying processing differences in affected and unaffected individuals. ERP studies of speech and tone processing have identified a characteristic auditory waveform. Generally, in research focusing on ERP it has been found that verbal children with autism spectrum disorders are less attentive to speech than typically developing controls, exhibiting differences in the P3, a brain component that reflects attention to stimuli in the environment (Courchesne *et al.*, 1985; Dawson *et al.*, 1988). Kuhl *et al*. (2005) used mismatch negativity (MMN), an ERP component that reflects discrimination of a stimulus change. They reported a correlation between degree of social deficits, expressive language skill and ability to discriminate a consonant-vowel syllable contrast in preschool-aged children with ASD. Further, Lepistö and colleagues (2005, 2006) have examined speech perception in individuals with ASD using ERP techniques and discovered poorer discrimination of duration changes of speech sounds in affected individuals compared to typically developing controls, even in those with Asperger syndrome, in whom language development is relatively spared.

Magnetoencephalography (MEG), like ERP, is a non-invasive technique used to assess perceptual processing in individuals with ASD. MEG makes use of magnetic fields produced by electrical activity in the brain. Using MEG, researchers found that children with ASD have demonstrated a significantly delayed electrophysiological response to a change in both speech (vowels) and non-speech stimuli as compared to typically developing controls (Oram Cardy *et al.*, 2005). Overall, these studies reveal that children with autism spectrum disorders exhibit difficulty in discriminating and/or recognizing sounds that can be detected at a basic level of processing.

Another method employed in speech perception research with individuals with autism and related disorders is functional magnetic resonance imaging (fMRI). Using volumetric measurements of blood flow, fMRI can provide information about activation in areas of the brain implicated in processing various types of stimuli. Gervais and colleagues (2004) reported an atypical pattern of response to vocal sounds in verbal adult males with autism as measured with fMRI. As a group, the affected individuals showed significantly less activation of the superior temporal sulcus, an area associated with perception of speech, and voices in particular (Belin *et al.*, 2000). Using fMRI, Bigler *et al.* (2007) report a dissociation between language skill and superior temporal gyrus size in affected individuals as compared to typically developing controls. Further, fMRI scans of affected individuals studied by Herbert and colleagues (Herbert *et al.*, 2002) revealed abnormal asymmetry in individuals with autism in brain regions associated with language.

Audiovisual speech perception

A major focus of research has been on detecting the sounds of speech in the development of language. This research studies language use in a non-communicative setting, involving a single language user—the subject. However, a great deal of our daily communication takes place in a face-to-face context. Accordingly, it is not surprising that visual information about speech has been shown to influence what typical listeners hear, assisting not just in the recognition of speech in noise (Sumby and Pollack, 1954), but in the perception of unambiguous speech as well (Desjardins *et al.*, 1997; Reisberg *et al.*, 1987). One powerful demonstration of the influence of visual information on what is heard is perceptual integration of mismatched audiovisual (AV) speech. McGurk and MacDonald (1976) first demonstrated this by presenting mismatching audio and video consonant-vowel (CVCV) tokens to perceivers. Perceivers watching these dubbed productions sometimes reported hearing consonants that combined the places of articulation of the visual and auditory tokens (e.g., visual /ba/ + auditory /ga/ heard as /bga/), or "fused" the two places (e.g., visual /ga/ + auditory /ba/ heard as /da/), or reflected the visual place information alone (visual /va/ + auditory /ba/ heard as /va/). This visual influence on mismatched speech, called the "McGurk effect" has been described as compelling for those perceivers who get the effect, occurring even when a perceiver is aware of how the stimuli have been manipulated (Massaro, 1987). In addition, the McGurk effect is robust, and has been demonstrated in the context of a number of manipulations, including asynchronous auditory and visual signals (Munhall *et al.*, 1996), non-frontal views of the speaker's face (Massaro, 1998), size reduction of visual stimuli (Jordan and Sergeant, 1998), point-light displays of the articulators (Johnson and Rosenblum, 1996) and presentation of very brief visual stimuli (Irwin *et al.*, 2006). In addition to natural speech, synthetic speakers have been developed to allow for precise manipulation of the auditory and visual speech signals, e.g., Massaro, 1998; Rubin and Vatikiotis-Bateson, 1998). The ability to integrate audiovisual speech is thought to be present at birth (Meltzoff and Kuhl, 1994). Visual influence on heard speech has been demonstrated in infants as young as 5 months of age (Rosenblum *et al.,* 1997).

The robustness of the visual influence on heard speech and the early age at which it occurs, e.g., Rosenblum *et al.*, 1997, suggests that the use of visual speech information is a central part of typical perceptual development. In typical perceivers, sensitivity to visual speech information is evident in infancy (Rosenblum *et al.*, 1997) and is thought to foster native language acquisition (Legerstee, 1990). In contrast, children with ASD show reduced social gaze to others' faces, when speech is produced (Hobson *et al.*, 1988). This reduction in social gaze may contribute to atypical language development in children with ASD. The tendency to avert gaze from the faces of others may lead to impoverished experience with speaking faces that could contribute to a reduction in detection of visual speech information. Moreover, a few studies (DeGelder *et al.*, 1991; Massaro and Bosseler, 2003; Williams *et al.*, 2004) suggest that children with ASD are less influenced by visual information for speech than are typically developing controls. Although these studies provide evidence that individuals with ASD are less influenced by some types of visible speech information, if perceivers fail to gaze at the speaker during production of speech it is impossible to determine whether perceivers with ASD actually integrate visual and auditory information less than controls do, or whether they are simply not gazing at the talking face. In particular, the tendency of individuals with ASD to avert gaze from the faces of others means that attenuated visual influence on heard speech in ASD may reflect reduced fixation on the face of a speaker (Pelphrey *et al.*, 2002). Alternatively, van der Geest and colleagues (2001a, 2001b, 2002) report similar patterns of gaze to still images of faces and social scenes in ASD and control perceivers, suggesting that the reduction in AV speech integration is not due to a lack of gaze to the speaker's face. Either the speaker's face may not hold the same information for a perceiver with ASD as for a typical perceiver, with affected individuals showing difficulty in extracting phonetic information from the face, or may be due to a more fundamental deficit in the capacity for AV integration.

Theories of perceptual deficits in autism

A number of theories of the perceptual deficits in individuals with ASD can account for a weakness in AV integration in this population. According to the central coherence theory (Frith and Happe, 1994) there is an impaired ability to perceive central coherence from individual features of a stimulus, such as a face (van der Geest, 2001a). The executive function theory (Ozonoff *et al.*, 1991) explains autistic symptomatology as deficits in planning, inhibition, flexibility and search behaviors mediated by the frontal lobes. Problems with executive functioning may lead to atypical patterns of gaze and thus impoverished perceptual processing of visual information. An additional theoretical account, derived from the perceptual learning theory of Gibson (1969), proposes that lack of attention to a speaker's face deprives a child with ASD of the experience necessary to develop typical sensitivity to visual speech information. Experience with auditory speech has been found to be crucial in developing perceptual sensitivities in early development (Bergeson and Pisoni, 2004). Accordingly, there is evidence that the production of speech differs in blind individuals (Brieland, 1950; Lezak and Starbuck, 1964). Because of their tendency to avert gaze from others' faces, individuals with ASD would have significantly limited access to visible speech information (Hobson *et al.*, 1988; Volkmar and Mayes, 1990; Volkmar *et al.*, 1989). Both seeing and listening to speech are essential for the development or maintenance of AV integration (Bergeson and

Pisoni, 2004). According to this account, impoverished experience with the faces of speakers diminishes the ability of children with ASD to detect structure in visible speech information, leading to a reduction in AV integration. Consistent with this hypothesis, in the context of AV mismatch, perceivers with ASD have been shown to use visual speech information less than auditory information (Massaro and Bosseler, 2003; Williams *et al*., 2004). Regardless of the theory, each makes the same general prediction, of reduced AV integration in children with ASD.

Recent research on audiovisual speech processing in ASD

My current program of research, conducted at Haskins Laboratories, is designed to refine these existing theories and provide data that will allow for more fine-grained accounts of audiovisual speech integration in this population. The principal goal of this program of research is to examine sensitivity to visual speech information in children with ASD when they are fixated on the face of a speaker, which has not been done in previous studies of audiovisual speech perception in the population of individuals with autism spectrum disorders. By employing visual tracking methodology, we can evaluate the degree to which children with ASD integrate audiovisual speech when fixated on the face of a speaker, as compared to typically developing controls. This application of visual tracking methodology allows us to adjudicate between two possible underlying causes of atypical audiovisual integration of speech in ASD: that affected children show reduced audiovisual integration because of gaze aversion to the face of a speaker and that children with ASD have an underlying weakness in integration of AV speech.

Converging evidence about integration of AV speech is currently being obtained by examining ASD and controlling perceivers' sensitivity to three types of AV speech processing: audiovisual integration, detection of audiovisual asynchrony (which, in typical perceivers, is related to audiovisual integration) and perception of audiovisual speech in the context of auditory noise (which, in typical perceivers leads to increased reliance on the visual speech information). Typical perceivers are influenced by visual speech information even when the auditory signal is unambiguous (the McGurk effect). Furthermore, for typical perceivers, the speaking face assists in recognition of auditory speech in noise (Sumby and Pollack, 1954). By examining the influence of visual speech information when the auditory signal is degraded, such as in the context of auditory noise we can assess whether, when pressed, the affected children's perceptual processing of AV speech can parallel typical processing.

At this point, preliminary data are available from this project for mismatched audiovisual (McGurk) speech stimuli. These pilot data support the hypothesis that children with ASD show less AV integration than typically developing (TD) children even when fixated on the face of a speaker. Two verbally fluent boys with autism (mean age 9.25 years, age range 9-9.5 years) were compared to 3 TD children (2 girls, 1 boy, mean age 9.3 years, range 7.5-10.5 years) on degree of visual influence of seen speech on heard speech. All of the children were native speakers of American English, and were reported by their parents to have normal hearing and normal or normal-to-corrected vision (one TD child wore corrective lenses during the testing procedure). The children with autism had received a clinical diagnosis of autism, and met criteria for autism on the Autism Diagnostic Observation Schedule (Lord *et al.*, 1996), an instrument for directly assessing in an individual's behaviors associated with autism and on the Autism Diagnostic Interview-revised (Lord *et al.*, 1994), a semi-structured interview for caregivers of children and adults for whom autism or pervasive developmental disorders is a possible diagnosis.

Eye gaze data were collected by superimposing a cursor on an image from a remote mounted scene camera that shows the participant's field of view, thus the system is able to measure gaze (Applied Science Laboratories, 2004). To optimize the accuracy of the pupil coordinates obtained by the optical camera, a magnetic head tracking unit in the form of a small sensor was attached to the head of the participant (an 8 millimeter sensor was attached to a slim wire placed on the head with a headband, see Figure 1).

The visual stimuli were presented on a computer monitor in front of the participant. The auditory speech stimuli were presented from a centrally located computer speaker placed directly below the monitor. A videotaped record of the participant was taken to allow for coding of verbal responses. A male, native speaker of English was videotaped producing the consonant-vowel syllables (CV) /ma/, /na/, and /ga/. These videotaped syllables were digitally edited with Adobe Premiere™ software to create the audio and video stimuli. Audio tokens were either /ma/ or /na/. Video tokens were either matching, cross-spliced tokens of /ma/ or /na/ (i.e., a different token of auditory /ma/ + visual /ma/) or mismatched AV tokens consisting of auditory /ma/ + visual /ga/that leads to a percept of /na/ when visual influence occurs.

Each subject was placed in a chair 25 inches from a computer monitor. The headband with the magnetic head tracking sensor was placed on each participant's head (see Fig. 1). The participant's pupil coordinates were calibrated with the eye-tracker system by asking the participant to look at col-

Fig. 1. A participant views a speaker using an eye tracker.

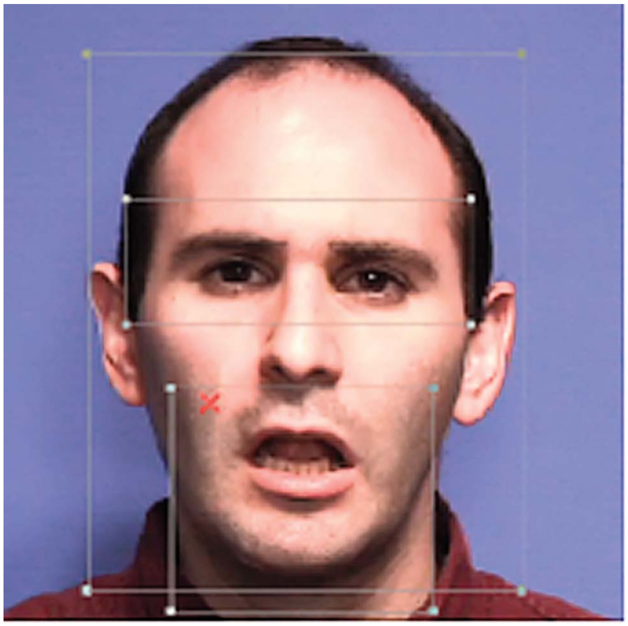

Fig. 2. "Look Zones" superimposed on a speaker's face.

ored dots that appeared on screen. Each participant was first presented with auditory stimuli consisting of the consonant-vowel (CV) tokens /ma/ and /na/. Each participant's verbal responses were coded for each trial from videotape by two separate coders, who were at 100% agreement. Participants reported hearing either /ma/ or /na/ for all trials. Both groups easily identified the matched /ma/ and /na/ stimuli (i.e., visual /ma/ + auditory /ma/), with only a single error across participants, with one child with ASD misidentifying one /ma/ as /na/.

To compare the amount of visual influence between groups on the AV speech trials, only those trials on which participants' gazes were on the face of the speaker during consonantal closure (which gives the critical visual information for what consonant is being produced) were assessed. Two independent coders examined videos of the crosshair that indicates participant's gaze superimposed on the face of the speaker (see the red cross on Figure 2). For those video frames displaying consonantal closure, trials where coders were in agreement that the crosshair was on the face of the speaker (see the largest outlined area containing the speaker's face, (Fig. 2) were included in analyses (gaze on the face of a speaker is sufficient for visual influence to occur, Paré *et al.*, 2003).

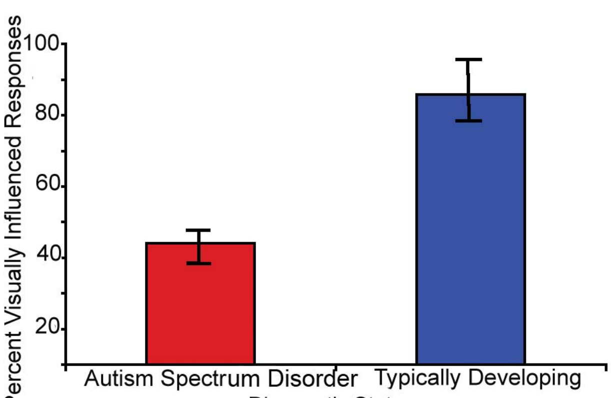

Fig. 3. Visually influenced responses when fixated on the speaker's face.

There were no differences between the groups in time spent gazing off the face of the speaker. For the children with ASD, one trial each had to be dropped from analyses because the children were not looking at the face at the time of consonantal closure. For the typically developing children, the mean number of looks off-face at time of consonantal closure was one (range of 0-2). For the AV integration stimuli, a response was considered to be visually influenced if the participant reported hearing /na/ for the mismatched (visual /ga/ + auditory /ma/) AV trials. A one-way analysis of variance (ANOVA) was run for the trials on which children were attending to the face of the speaker during consonantal closure. There was a significant mean difference in visual influence between the groups, $F(1,3) = 17.8$, $p<.02$. The mean visual influence for the children with ASD was 44% (range 41-47%) and 85.7% for the typically developing children (range 75-100%), suggesting as hypothesized, that the children with ASD were less visually influenced than their typically developing peers, even when fixated on the face of the speaker (see Fig. 3).

Future directions

These intriguing preliminary findings suggest less visual influence on heard speech in children with autism as compared to typically developing controls. Ongoing research will closely examine ASD and control perceivers' sensitivity to audiovisual speech processing in a range of contexts, including audiovisual integration, detection of audiovisual asynchrony and perception of audiovisual speech in the context of auditory noise. This research will move the study of perceptual processing of audiovisual speech in children with autism forward by determining whether AV speech processing in children with ASD is fundamentally disrupted, or whether visual influence can be modulated depending on task conditions such as in the presence of auditory noise. This work will have implications for the identification and characterization of audiovisual integration deficits in children with ASD, including the potential for identifying novel subgroups within ASD, and the design of targeted interventions to improve sensitivity to visual speech information.

The continued examination of basic processes in speech perception in children with autism spectrum disorders will make it possible to characterize the nature of auditory and audiovisual speech perception in this population. Results from this area of inquiry will move the field forward significantly by providing basic data to guide focused interventions related to speech perception and processing in the population of children with ASD. Understanding the mechanisms that underlie delays and deficits in language acquisition is particularly important because language ability is a key prognostic factor for long-term outcomes among children and adults with ASD (e.g., Lord and Venter, 1992) and early intervention for individuals with language deficits is associated with long term improved developmental and cognitive outcomes (Lord, 1995; Vostanis *et al.*, 1994; Robins *et al.*, 2001).**AT**

References for further reading

Applied Science Laboratories (**2004**). *Technical report*. www.a-s-l.com

American Psychiatric Association (**2000**). *Diagnostic and Statistical Manual of Mental Disorders DSM-IV-TR (Text Revision)* (American Psychiatric Association, Washington, DC).

Baranek, G. T. (**1999**). "Autism during infancy: A retrospective video analysis of sensory-motor and social behaviors at 9-12 months of age." J. Autism and Developmental Disorders **29**, 213–224.

Belin, P., Zatorre, R. J., Lafaiile, P., Ahad, P., and Pike, B. (**2000**). "Voice-selective areas in human auditory cortex," Nature **403**, 309.

Bergeson, T. R., and Pisoni, D. B. (**2004**). "Audiovisual speech perception in deaf adults and children following cochlear implantation," in *The Handbook of Multisensory Processes*, edited by B. E. Stein, C. Spence, and G. Calvert (MIT Press; Cambridge, MA), pp. 749–771.

Bigler, E. D., Mortensen, S., Neeley, E. S., Ozonoff, S., Krasny, L., Johnson, M., Lu, J., Provencal, S. L., McMahon, W., and Lainhart, J. E. (**2007**). "Superior temporal gyrus, language function and autism," Developmental Neuropsychology **31**(2), 217–238.

Brieland, D. M. (**1950**). "A comparative study of the speech of blind and sighted children," Speech Monographs **17**, 98–103.

Centers for Disease Control and Prevention. Surveillance Summaries (**2007**). Morbidity and Mortality Weekly Report **56**, SS#1, February.

Courchesne, E., Lincoln, A., Kilman, B., and Galambos, R. (**1985**). "Event-related brain potential correlates of the processing of novel visual and auditory information in autism," J. Autism and Developmental Disorders **15**, 55–76.

Dawson, G., Finley, C., Phillips, S., Galpert, L., and Lewy, A. (**1988**). "Reduced P3 amplitude of the event-related brain potential: Its relationship to language ability in autism," J. Autism and Developmental Disorders **18**, 493–504.

De Gelder, B., Vroomen, J., Van Der Heide, L. (**1991**). "Face recognition and lip-reading in autism," European J. Cognitive Psychology **3**, 69–86.

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

Fombonne, E. (**2003a**). "The prevalence of autism," J. American Medical Association **289**, 87–89.

Fombonne, E. (**2003b**). "Epidemiological surveys of autism and other pervasive developmental disorders: An update," J. Autism and Developmental Disorders **33**, 365–382.

Frith, U., and Happe, F. (**1994**). "Autism: Beyond 'Theory of Mind,'" Cognition **50**, 115–132.

Gervais, H., Belin, P., Boddaert, N., Leboyer, M., Coez, A., Sfaello, I., Barthelemy, C., Brunelle, F., Samson, Y., and Zilbovicius, M. (**2004**). "Abnormal cortical voice processing in autism," Nature Neuroscience **7**, 801–802.

Gibson, E. J. (**1969**). *Principles of Perceptual Learning and Development* (Appleton, New York).

Glaze, D. G. (**2004**). "Rett syndrome: Of girls and mice–lessons for regression in autism," Mental Retardation and Developmental Disabilities Research Review **10**, 154–158.

Herbert, M. R., Harris, G. J., Adrien, K. T., Ziegler, D. A., Makris, N., Kennedy, D. N., Lange, N. T., Chabris, C. F, Bakardjiev, A., Hodgson, J., Takeoka, M., Tager-Flusberg, H., and Caviness, V. S. (**2002**). "Abnormal asymmetry in language association cortex in autism," Annals of Neurology **52**, 588–596.

Hobson, R. P., Ouston, J., and Lee, A. (**1988**). "What's in a face? The case of autism," British J. Psychology **79**, 441–453.

Howlin, P. (**2003**). "Outcome in high-functioning adults with autism with and without early language delays: Implications for the differentiation between autism and Aspberger syndrome," J. Autism and Developmental Disorders **33**, 3–13.

Irwin, J. R., Whalen, D. H., and Fowler, C. A. (**2006**). "A sex difference in audiovisual integration of speech," Perception and Psychophysics **68**(4) 582–592.

Jansiewicz, E. M., Goldberg, M. C., Newschaffer, C. J., Denckla, M. B., Landa, R., and Mostofsky, S. H. (**2006**). "Motor signs distinguish children with high functioning autism and Aspberger's Syndrome from controls," J. Autism and Developmental Disorders **36**, 613–621.

Johnson, J. A., and Rosenblum, L. D. (**1996**). "Hemispheric advantage for point-light McGurk effects," Brain and Cognition **32**, 167–170.

Jordan, T. R., and Sergeant, P. C. (**1998**). "Effects of facial image size on visual and audio-visual speech recognition," in *Hearing by eye II: Advances in the psychology of speechreading and auditory-visual speech*, edited by R. Campbell, B. Dodd and D. Burnham (Psychology Press, Hove), pp. 155–176.

Klinger, L., Dawson, G., and Renner, P. (**2002**). "Autistic disorder," in *Child Psychopathology* (2nd edition), edited by E. Mash and R. Barkley (Guilford Press, New York) 409–454.

Koning, C., and Magill-Evans, J. (**2001**). "Social and language skills in adolescent boys with Asperger syndrome," Autism b, 23–36.

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

Landa, R. (**2007**). "Early communication development and intervention for children with autism," Mental Retardation and Developmental Disabilities Research Reviews **13**, 16–25.

Legerstee, M. (**1990**). "Infants use multimodal information to imitate speech sounds," Infant Behavior and Development **13**, 343–354.

Lepistö, T., Kujala, T., Vanhala, R., Alku, P., Huotilainen, M., and Näätänen, R. (**2005**). "The discrimination of and orienting to speech and non-speech sounds in children with autism," Brain Research **1066**,147–157.

Lepistö, T., Silokallio, S., Nieminen-von Wendt, T., Alku, P., Näätänen, R., and Kujala, T. (**2006**). "Auditory perception and attention as reflected by the brain event-related potentials in children with Aspberger syndrome," Clinical Neurophysiology **117**, 2161–2171.

Lezak, R. J., and Starbuck, H. B. (**1964**). "Identification of children with speech disorders in a residential school for the blind," Int. J. Education of the Blind **14**, 8–12.

Lord, C., and Magill-Evans, J. (**1995**). "Peer interactions of autistic children and adolescents," Development and Psychopathology **7**, 611–626.

Lord, C., and Paul, R. (**1997**). "Language and communication in autism," in *Handbook of Autism and Pervasive Developmental Disorders* edited by D. J. Cohen and F. R. Volkmar (Wiley, New York), pp. 195–225.

Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Jr., Leventhal, B. L., DiLavore, P. C., Pickles, A., and Rutter, M. (**2000**). "The autism diagnostic observation schedule–Generic: A standard measure of social and communication deficits associated with the spectrum of autism," J. Autism and Developmental Disorders **30**(3), 205–223.

Lord, C. Rutter, M., and DeLavore, P. (**1996**). Autism Diagnostic Observation Schedule. Unpublished manuscript. University of

Chicago.
Lord, C., Rutter, M., and LeCouteur, A. (**1994**). "Autism diagnostic interview–Revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders," J. Autism and Development Disorders **24**, 659–685
Lord, C., Rutter, M., Goode, S., Heemsbergen, J., Jordan, H., Mawhood, L., and Schopler, E. (**1989**). "Autism diagnostic observation schedule: A standardized observation of communicative and social behavior," J. Autism and Developmental Disorders **19**(2), 185–212.
Lord, C., and Venter, A. (**1992**). "Outcome and follow-up studies of high-functioning autistic individuals," in *High-functioning individuals with autism*, edited by Eric Schopler and Gary B. Mesibov (Plenum Press, New York), pp. 187–199.
Loveland, K., and Tunali-Kotoski, B. (**1997**). "The school-aged child with autism," in *The Handbook of Autism and Pervasive Developmental Disorders* (2nd Edition), edited by D. Cohen and F. Volkmar (Wiley, New York), pp. 283–308.
Massaro, D. W., and Boesseler, A. (**2003**). "Perceiving speech by ear and eye: Multimodal integration by children with autism," J. Developmental and Learning Disorders **7**, 111–146.
Massaro, D. W. (**1998**). *Perceiving Talking Faces: From Speech Perception to a Behavioral Principle* (MIT Press, Cambridge, MA).
Massaro, D. W. (**1987**). *Speech Perception by Ear and Eye: A Paradigm for Psychological Enquiry* (Lawrence Erlbaum Associates, Hillsdale, N.J).
McGurk, H., and MacDonald, J. (**1976**). "Hearing lips and seeing voices," Nature **264**, 746–748.
Meltzoff, A. N., and Kuhl, P. K. (**1994**). "Faces and speech: Intermodal processing of biologically relevant signals in infants and adults," in *The Development of Intersensory Perception: Comparative Perspectives* edited by D. J. Lewkowitz and R. Lickliter (Earlbaum, Hillsdale, NJ), pp. 335–369.
Mount, R. H., Charman, T., Hastings, R. P., Reilly, S., and Cass, H. (**2003**). "Features of autism in Rett syndrome and severe mental retardation," J. Autism and Developmental Disorders **33**(4), 435–442.
Munhall, K.G., Gribble, P., Sacco, L., and Ward, M. (**1996**). "Temporal constraints on the McGurk effect," Perception and Psychophysics **58**, 351–362.
Oram Cardy, J. E., Flagg, E. J., Roberts, W., and Roberts, T. P. L. (**2005**). "Delayed mismatch field for speech and non-speech sounds in children with autism," NeuroReport **16**(5), 521–525.
Ozonoff, S., Pennington, B. F., and Rogers, S. J. (**1991**). "Executive function deficits in high-functioning individuals: Relationship to theory of mind," J. Child Psychology and Psychiatry **32**, 1081–1105.
Paré, M., Richler, R. C., ten Hove, M., and Munhall, K. G. (**2003**). "Gaze behavior in audiovisual speech perception: The influence of ocular fixations on the McGurk effect," Perception and Psychophysics **65**, 553–567.
Pelphrey, K. A., Sasson, N. J., Reznick, J. S., Paul, G., Goldman, B. D., and Piven, J. (**2002**). "Visual scanning of faces in autism," J. Autism and Developmental Disorders **32**, 249–261.
Rapin, I. (**2001**). "An 8-year-old boy with autism," J. American Medical Association **285**, 1751.
Reisberg, D., McLean, J., and Goldfield, A. (**1987**). "Easy to hear but hard to understand: A lip-reading advantage with intact auditory stimuli," in *Hearing by Eye: The Psychology of Lip-Reading,* edited by B. Dodd and R. Campbell (Lawrence Erlbaum Associates, Ltd., London), pp. 97–113.
Robins, D., Fein, D., Barton, M., and Green, J. (**2001**). "The modified-checklist for autism in toddlers M-CHAT): An individual investigation in the early detection of autism and pervasive development disorders," J. Autism and Developmental Disorders **31**(2), 131–144.
Rosenblum, L. D., Schmuckler, M. A., and Johnson, J. A. (**1997**). "The McGurk effect in infants," Perception and Psychophysics **59**, 347–357.
Rubin, P., and Vatikiotis-Bateson, E. (**1998**). "Talking Heads," in *Proceedings of AVSP'98: International Conference on Auditory-Visual Speech Processing,* edited by D. Burnham, J. Robert-Ribes, and E. Vatikiotis-Bateson (Ink Press, Sydney), pp. 231–235.
Rutter, M. (**1983**). "Cognitive deficits in the pathogenesis of autism," J. Child Psychology and Psychiatry **24**, 513–531.
Sumby, W. H., and Pollack, I. (**1954**). "Visual contribution to speech intelligibility in noise," J. Acoustical Society of America **26**, 212–215.
van der Geest, J. N. (**2001a**). "Looking into autism, eye movements and gaze behavior of autistic children," doctoral dissertation, Utrecht University.
van der Geest, J. N., Kemner, C., Camfferman, G., Verbaten, M. N., and van Engleland, H. (**2001b**). "Eye movements, visual attention, and autism: A saccadic reaction time study using the gap and overlap paradigm," Biological Psychiatry **50**, 614–619.
van der Geest, J. N., Kemner, C. G., Verbaten, M. N., and van Engleland, H. (**2002**). "Gaze behavior of children with pervasive developmental disorder toward human faces: A fixation time study," J. Child Psychology and Psychiatry **43**, 669–678.
Volkmar, F. R., and Mayes, L. C. (**1990**). "Gaze behavior in autism," Development and Psychopathology 2, 61–69.
Volkmar, F. R., Sparrow, S. S., Rende, R. D., and Cohen, D. J. (**1989**). "Facial perception in autism," J. Child Psychology and Psychiatry **30**, 591–598.
Vostanis, P., Smith, B., Chung, M. C., and Corbett, J. (**1994**). "Early detection of childhood autism: A review of screening instruments and rating scales," Child Care, Health and Development **20**, 167–177.
Williams, D. L., Goldstein, G, Minshew, N. J. (**2006**). "Neuropsychologic functioning in children with autism: Further evidence for disordered complex information-processing," Child Neuropsychology **12**, 279–290.
Williams, J. H. G., Massaro D. W., Peel, N. J., Bosseler A., and Suddendorf, T. (**2004**). "Visual-auditory integration during speech imitation in autism," Research in Developmental Disabilities **25**, 559–575.

Dr. Julia Irwin received a Ph.D. in Child and Developmental Psychology from the University of Connecticut and completed her postdoctoral training at the Yale University Child Study Center. Currently, she is a Senior Research Scientist at Haskins Laboratories in New Haven, CT, an independent community of researchers conducting basic research on spoken and written language.