Auditory Deficits in Audiovisual Speech Perception in Adult Asperger's Syndrome: fMRI Study

Fabian-Alexander Tietze[1], Laura Hundertmark[1], Mandy Roy[1,2], Michael Zerr[3], Christopher Sinke[1], Daniel Wiswede[4,5], Martin Walter[6], Thomas F. Münte[4,5] and Gregor R. Szycik[1]*

[1] Department of Psychiatry, Social Psychiatry and Psychotherapy, Hannover Medical School, Hanover, Germany, [2] Asklepios Clinic North – Ochsenzoll, Hamburg, Germany, [3] Department of Psychosomatic Medicine and Psychotherapy, Hannover Medical School, Hanover, Germany, [4] Institute of Psychology II, University of Lübeck, Lübeck, Germany, [5] Department of Neurology, University of Lübeck, Lübeck, Germany, [6] Department of Psychiatry and Psychotherapy, University of Tübingen, Tübingen, Germany

Audiovisual (AV) integration deficits have been proposed to underlie difficulties in speech perception in Asperger's syndrome (AS). It is not known, if the AV deficits are related to alterations in sensory processing at the level of unisensory processing or at levels of conjoint multisensory processing. Functional Magnetic-resonance images (MRI) was performed in 16 adult subjects with AS and 16 healthy controls (HC) matched for age, gender, and verbal IQ as they were exposed to disyllabic AV congruent and AV incongruent nouns. A simple semantic categorization task was used to ensure subjects' attention to the stimuli. The left auditory cortex (BA41) showed stronger activation in HC than in subjects with AS with no interaction regarding AV congruency. This suggests that alterations in auditory processing in unimodal low-level areas underlie AV speech perception deficits in AS. Whether this is signaling a difficulty in the deployment of attention remains to be demonstrated.

Keywords: Asperger's syndrome, autism, audiovisual, speech, multisensory, fMRI

INTRODUCTION

Speech perception under natural conditions relies on auditory and on visual information. In particular, speech lip movements impact speech perception as powerfully demonstrated by the so called McGurk-illusion (McGurk and MacDonald, 1976). But audiovisual (AV) perception does not only play a role in artificial illusions. Under noisy conditions AV information (in comparison to auditory only information) results in considerable improvements in intelligibility (Sumby and Pollack, 1954; Schwartz et al., 2004; Ross et al., 2007a). In complex auditory environments, such as the proverbial cocktail-party, visual cues may enhance selective listening (Driver, 1996). Combining multisensory information is crucial not only for perception of speech but also for perception of other environmental events (e.g., traffic noise in the city) and provides different evolutionary benefits, like rapid stimulus detection (Gondan et al., 2005; Diederich and Colonius, 2015) and

localization (Hairston et al., 2003). Thus intact AV processing is crucial for swift interaction with the environment and also serves important social functions.

Deficits in AV perception have been shown for diverse populations with neuropsychiatric disorders. Individuals with dyslexia show multisensory processing deficits with reduced ability to integrate AV information (Hahn et al., 2014; Rüsseler et al., 2015, 2018; Ye et al., 2017). Likewise, schizophrenia has also been reported to lead to impaired AV speech perception (De Gelder et al., 2003; Ross et al., 2007b; Szycik et al., 2009b, 2013) and to prolonged reaction times in AV target detection tasks (Williams et al., 2010). Subjects with different kinds of synesthesia show reduced AV integration of simple AV stimuli (Neufeld et al., 2012) as well as of AV speech (Sinke et al., 2014).

Autism spectrum disorder (ASD) is a neurodevelopmental condition characterized by persistent deficits in social communication and restricted, repetitive patterns of behavior, interests or activities (American Psychiatric Association, 2013). Eye contact and eye gaze is generally atypical in autism, a feature that persists into adulthood (Baron-Cohen et al., 2001a; Klin et al., 2002; Pelphrey et al., 2005; Pierce et al., 2011; Jones and Klin, 2013). In addition to the impact of an impairment of gaze behavior on social contexts, these impairments might also impact AV integration of speech. Indeed, there is growing evidence for AV integration deficits in ASD (Feldman et al., 2018a). Changes in multisensory function have been demonstrated using simple auditory and visual stimuli (Brandwein et al., 2013) and the double-flash illusion (Foss-Feig et al., 2010). Subjects with ASD are less susceptible to the speech related McGurk illusion (Taylor et al., 2010; Bebko et al., 2014; Stevenson et al., 2014) and show lower sensitivity for temporal synchrony of AV speech events (Bebko et al., 2006). But the picture is not fully consistent (Zhang et al., 2018). Studies exist showing no differences in AV speech perception between ASD and typically developed subjects (Williams et al., 2004). The inconsistencies of the findings are likely a result of the heterogeneity of the studied subjects who differed often in age and severity of symptomatology of ASD across studies (Feldman et al., 2018a).

In the DSM-5 (American Psychiatric Association, 2013), the categorization into autistic disorder and Asperger's syndrome (AS) no longer exists. Autism is now defined as ASD with three gradient levels of severity categorized according to the severity of symptoms and need of support. Until the publication of DSM-5 in 2013, the former DSM-IV-TR subsumed at least two different main-subcategories of ASD within pervasive development disorders, i.e., autistic disorder and AS. Thereby diagnosis of autistic disorder was related to a delay in the development of spoken language. According to the DSM-5, individuals with an assured DSM-IV-TR diagnosis of an AS shall now be classified as ASD. Therefore results of earlier studies have to be interpreted keeping in mind the fact that some of them analyzed only subpopulations of the recent ASD group. In this manuscript we refer to ASD according to DSM-5 and AS as a subpopulation of ASD according to DSM-IV-TR.

Whereas children with AS showed a regular rate of McGurk illusions (Bebko et al., 2014), adolescent ASD subjects (Stevenson et al., 2014) showed reduced rates and adult AS

subjects have been shown to have qualitative differences in perception of McGurk stimuli compared to control subjects (Saalasti et al., 2012) and to show further differences in AV speech perception (Saalasti et al., 2011). On the other hand children with AS also show deficits in speech perception (Saalasti et al., 2008) and these deficits seem to be amplified in acoustically noisy environments (Alcantara et al., 2004), i.e., situations in which visual information is most important (Ross et al., 2007a). Thus, autistic children have less benefit from AV information in speech perception of whole words whereas perception of phonemes is almost intact (Stevenson et al., 2017). The authors interpret their finding in the manner of affected sensory processing as origin for reduced multisensory abilities in autism. This may be reflected in atypical patterns of sensory responsiveness related to multisensory speech perception and integration problems (Feldman et al., 2018b). Interestingly, AV deficits in speech perception seem to disappear in early adolescence (Foxe et al., 2015). It is therefore possible that multisensory integration problems are directly related to disturbed or prolonged maturation of the sensory system in AS (Brandwein et al., 2015) or ASD (Beker et al., 2018). This could be dependent on the specific subtype respective severity level when taking into account the genetic heterogeneity of the ASD population (Robertson and Baron-Cohen, 2017) or specific age of analyzed subjects (Beker et al., 2018; Feldman et al., 2018a).

To summarize, the knowledge about AV integration deficits in ASD and especially AS is quite sparse and ambiguous with growing evidence for worse performance within the ASD population compared to typically developed subjects. These deficits are more pronounced in younger subjects and are correlated with autism symptom severity. Recently different mechanisms underlying these deficits have been postulated (Robertson and Baron-Cohen, 2017). These mechanisms can be subdivided in sensory-first and top-down accounts.

Within the sensory-first account AV integration deficits have been attributed to unisensory deficits, e.g., a weaker visual influence during multisensory processing, conceptualized as a rather peripheral problem (Gelder et al., 1991). This reduction affects predominantly tasks involving human faces or socially relevant stimuli (Mongillo et al., 2008; Irwin et al., 2011) and may be also based on the specific deficit related to the ability to process biological motion (Saalasti et al., 2012) or problems with processing temporal aspects of sensory information (Kwakye et al., 2011; Woynaroski et al., 2013).

Top-down accounts explain such AV integration deficits by problems with combining stimulus details or specific features into a coherent percept as it is postulated by the "weak central coherence theory" (Happe and Frith, 2006). Following the theory of weak central coherence, unisensory representations of specific environmental objects or events are not or rather not sufficiently integrated into a multisensory representation of these objects or events (Megnin et al., 2012). In this case alterations in functionality of typical multisensory brain areas could be responsible for AV integration deficits.

To identify brain areas responsible for AV speech integration in AS, we used functional magnetic resonance imaging (fMRI) in adult AS subjects who were exposed to AV congruent

(video and audio signal matched, e.g., visual: hotel, auditory: hotel) and incongruent speech (video and audio signal did not match, e.g., visual: island, auditory: hotel) stimulation. As far as unisensory deficits underlie multisensory integration problems in AS, this should be reflected in activity differences within specific unisensory brain areas. Otherwise differences within typical multisensory brain areas should be detected.

MATERIALS AND METHODS

All procedures had been approved by the local Ethics Committee of the Hannover Medical School and have been performed in accordance with the ethical standards laid down in the Declaration of Helsinki. The participants gave written informed consent and participated for a small monetary compensation.

Participants

Participants were divided into two groups: 16 adult subjects diagnosed with AS meeting the DSM-IV-TR criteria (American Psychiatric Association, 2000) and 16 age-matched healthy controls (HC). Nine of the AS subjects met DSM-IV-TR criteria for one or more additional psychiatric diagnoses (agoraphobia, $n = 2$; social phobia, $n = 2$; somatization disorder, $n = 1$; depression, $n = 2$; attention deficit hyperactivity disorder, $n = 1$; dysthymia, $n = 2$; and obsessive-compulsive disorder, $n = 1$). All HC subjects were free of previous and recent neurological or psychiatric diseases. Two additional AS subjects were excluded from the study due to excessive movements (more than 3 mm or 3° rotation) within the scanner ($n = 1$) and incomplete fulfilment of diagnostic criteria ($n = 1$). Two additional HC subjects were excluded because of excessive movement during MRI data acquisition. All AS subjects were recruited from the long-standing patient pool of the Department of Psychiatry of Hannover Medical School. All of them were consulted for a first-time diagnosis. As the diagnostic assessment was primarily made in the context of the psychiatric health care system and participants were recruited from that existing pool, only diagnostic data according to DSM-IV-TR-criteria was available. As shown in **Table 1** both groups were matched for age, gender, and estimated verbal IQ as assessed by the MWT-B – "Mehrfachwahl-Wortschatz-Intelligenztest" (Lehrl et al., 1995).

All participants were native speakers of German. The mean autism-spectrum quotient (AQ) for the AS subjects group was

TABLE 1 | Characteristics of the experimental groups.

	AS (n = 16)	HC (n = 16)	Significance
Age in years (M, SD)	39.50, 11.17	33.75, 8.22	$t = 1.76$; n.s. ($p = 0.09$, $d = 0.59$)
No of males/females	13/3	15/1	$\chi^2_{0.05;1} = 1.14$; n.s. ($p = 0.29$, $V = 0.19$)
Verbal IQ[a]	30.28, 4.37	31.72, 4.10	$t = -0.84$; n.s. ($p = 0.34$, $d = 0.34$)
AQ[b]	40.37, 5.27		

[a]IQ, raw score assessed by the MWT-B – "Mehrfachwahl-Wortschatz-Intelligenztest". [b]Autism-spectrum Quotient.

40.37 ± 5.27 (Baron-Cohen et al., 2001b) and thus in the higher range. Two AS subjects declared ambidexterity, whereas two HC subjects were exclusively left-handed.

DSM-IV-TR criteria for AS in adulthood (American Psychiatric Association, 2000) were thoroughly explored by a self-developed semi-structured interview ("Diagnostic interview: AS in adulthood"). After a general section focusing on medical anamnesis (somatic, psychiatric, and social histories, including childhood development), the interview continued with a special section involving AS that included the following items with respect to childhood and adulthood: social interaction and communication (e.g., friendships with/relationship to/interest in peers, and being a loner and suffering from loneliness); special interests (e.g., spending leisure time and interest in specific objects/topics); stereotypic behavior (e.g., rituals and reaction toward disturbances of rituals); and other characteristics (e.g., clumsiness and sensitivity toward noises/smells/tactile stimuli). The interview contained items and descriptions of all relevant criteria for the diagnosis of AS as defined in DSM-IV-TR. The result of the interview was confirmed for every AS-subject by verifying the threshold value of the AQ. Additionally, we observed eye contact, facial expression, prosody, and "mirroring" of affections and clumsiness during the interview. The interview was conducted by the same experienced psychiatrist and had a duration of approximately 90 min. At the time of diagnostic investigation, the investigator was blind to the research questions. Diagnosis was completed with information from personal interviews, gained by telephone or in written form, and observers in child- and/or adulthood (e.g., partners, friends, and parents or siblings). In some cases, school reports were incorporated. All DSM-IV-TR criteria had to be clearly fulfilled to confirm diagnosis. Moreover, retrospective data on the development of speech were assessed. An additional examination for axis-I co-morbidity was undertaken by using the German version of the Structured Clinical Interview for DSM-IV Axis I Disorders (SCID-I) (Wittchen et al., 1997). We had no exclusion-criteria based on medication.

Stimuli and Design

The chosen paradigm to elicit brain activity for AV speech stimulation has been successfully used in healthy subjects and different neuropsychiatric populations (Szycik et al., 2009a,b; Rüsseler et al., 2017). Stimuli, taken from the German part of the CELEX-Database (Baayen et al., 1993), comprised 70 disyllabic nouns with a Mannheim frequency 1,000,000 (MannMln) of at least one (see, **Supplementary Table S1** for the stimulus list in our study). MannMln frequency serves as frequency measure indicating the occurrence of a word within the 6,000,000 words of the Mannheim word corpus. The stimuli were spoken by a female native speaker of German with linguistic experience and recorded by means of a digital camera and a microphone. The videos (400^2 pixels resolution, 6° visual angle) showed the frontal view of the whole face of the speaker and were divided into periods of 2 s duration, accompanied by audio streams in mono-mode. The stimuli were randomly divided into two sets of 35 items each. The first set contained video segments with congruent AV information: Lip movements were fitting to the

word spoken by the speaker. The second set consisted out of video sequences with incongruent AV information: Lip movements did not fit to the spoken words; e.g., video: Engel/angel and audio: Hase/rabbit. Auditory and visual information in the incongruent stimuli started simultaneously to the onset of vocalization. The participants were instructed to carefully watch and listen to the stimuli without being informed about the AV incongruence of some stimuli. The subjects were asked to keep attention on both modalities. To ensure the attention of the subjects to the stimuli, we used a simple semantic categorization task and analyzed the detection rate (answer rate for each stimulus). Subjects were asked to respond for each stimulus by pressing the left or right response device for stimuli describing living objects (five target stimuli) vs. objects of other categories (remaining 30 stimuli). The loudness of the presented stimuli was individually adjusted to the almost audible threshold for auditory comprehension. Firstly, the interaural loudness difference (due to individually applied ear plugs) for each subject was corrected by presenting a simple test tone and changing the sound pressure level (SPL) bilaterally until subjects signaled that the audible signal was equally loud in both ears. In the second step, we presented some test stimuli during the real scanner noise and increased SPL until subjects gave us a signal by the response device that they heard the stimuli well.

Presentation software (Neurobehavioral Systems, Inc., Albany, CA, United States) was used to deliver the stimuli. A "slow event related" design was used for the stimulus presentation. Each stimulation event was followed by a fixed resting period of 16 s duration. During this time a dark screen with a fixation cross at the position of the speaker's mouth was presented. The duration of the whole functional stimulation part of the experiment was therefore 21 min.

Presenting the stimuli and communicating between the examination and control rooms was possible due to an fMRI compatible audio system integrated into earmuffs for reduction of residual background scanner noise. Visual stimuli were presented on a MRI compatible screen positioned in the front of the scanner. Subjects were able to see the screen through a mirror positioned on the top of the head coil. To ensure good visibility a detailed test picture with similar size and resolution as the video sequences was presented prior to the experiment and all participants were asked to report the content of this picture.

Image Acquisition

Magnetic-resonance images were acquired on a 3-T Siemens Skyra Scanner (Siemens, Erlangen, Germany) equipped with a standard head coil. A total of 640 T2*-weighted volumes of the whole brain [EPI-sequence; time to repeat (TR) 2000 ms, echo time (TE) 30 ms, flip angle 80°, and field of View (FOV) 192 mm, matrix 64^2, 30 slices, slice thickness 3.5 mm, interslice gap 0.35 mm] near to standard bicommissural (AC-PC) orientation were collected. After the functional measurement a 3D high resolution T1-weighted volume for anatomical information (MPRAGE-sequence, 192 slices, FOV = 256 mm, isovoxel 1 mm, TR 2.5 s, TE 4.37 ms, flip angle 7°) was recorded. The subject's head was fixed during the entire measurement to avoid head movements.
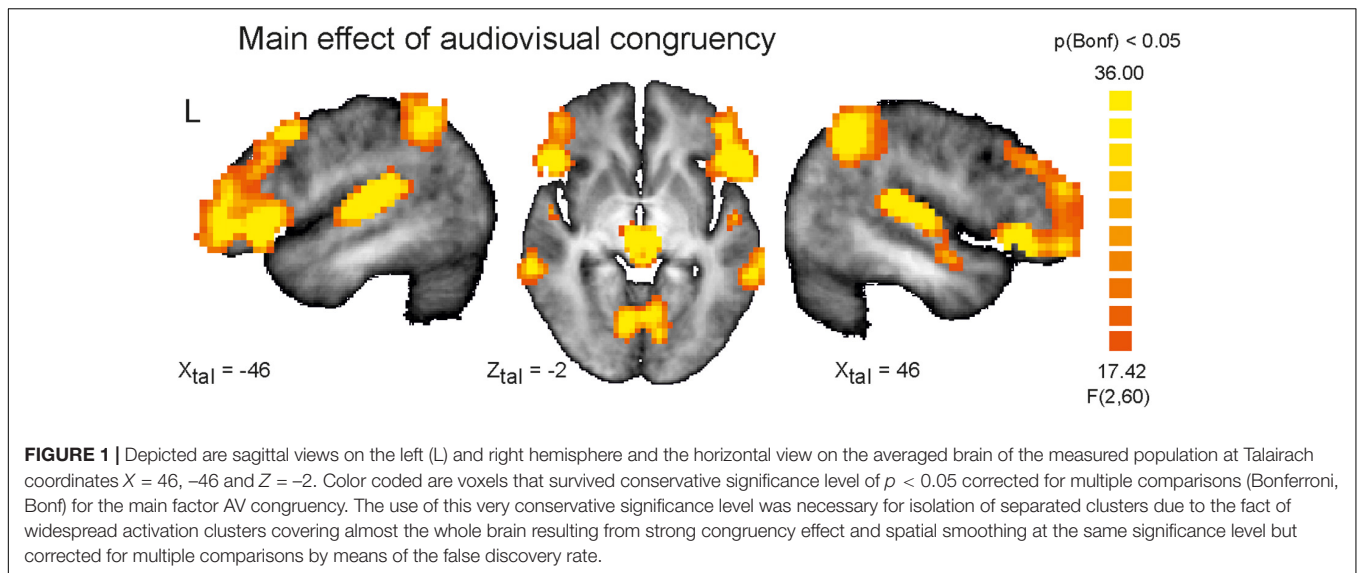
fMRI Data Analysis

Analysis and visualization of the data were performed using Brain Voyager QX (Brain Innovation BV, Maastricht, Netherlands) software (Goebel et al., 2006). First, a correction for the temporal offset between the slices acquired in one scan was applied. For this purpose the data was interpolated. After this slice scan time correction a 3D motion correction was performed by realignment of the entire measured volume set to the first volume by means of trilinear interpolation. Thereafter, linear trends were removed and a high pass filter was applied resulting in filtering out signals occurring less than two cycles in the whole time course. Structural and functional data were spatially transformed into the Talairach standard space (Talairach and Tornoux, 1988) using a 12-parameter affine transformation. Functional EPI volumes were spatially smoothed with an 8 mm full-width half-maximum isotropic Gaussian kernel to accommodate residual anatomical differences across participants.

For the statistical model a design matrix including all conditions of interest was specified using a hemodynamic response function. This function was created by convolving the rectangle function with the model of Boynton et al. (1996) using $\Delta = 2.5$, $\tau = 1.25$ and $n = 3$. Thereafter, a multi-subject random effects (RFX) analysis of variance model (ANOVA) with stimulation (AV-congruent vs. AV-incongruent) as the first main within-subject factor and group (AS vs. HC) as the second main between-subject factor was used for identification of significant differences in hemodynamic responses. Main effects of both factors and their interaction were considered. The false discovery rate threshold of q(FDR) < 0.05 (Genovese et al., 2002) was chosen for identification of the activated voxels. The centers of mass of suprathreshold regions were localized using Talairach coordinates (Talairach and Tornoux, 1988) and the Talairach Daemon tool (Lancaster et al., 2000).

RESULTS

Since each participant responded to almost all stimuli and the missing rate was nearly zero (ceiling effect) for each of the participants, all participants have attentively completed the experiments. Widespread differences of activation involving large parts of the brain were observed for the main within-subject factor congruency (AV congruent vs. incongruent) at the chosen significance level of q(FDR) < 0.05. Even after changing the significance threshold to $p < 0.0001$ (Bonferroni corrected for multiple comparisons) multiple significant clusters covering brain structures commonly involved in auditory and visual processing and attention functions were observed indicating the high impact of AV incongruency (**Figure 1** and **Table 2**).

The main result of the study is a significant difference in activation for the main between-subjects factor (AS vs. HC). **Figure 2** demonstrates the only one suprathreshold [q(FDR) < 0.05] cluster identified for this contrast. This left hemispheric cluster with an extent of 27 functional voxels occupies the transverse Temporal Gyrus and has its center of mass at Talairach coordinates (x, y, z) −42, −25, 12, i.e., Brodmann area 41. The analysis of mean betas extracted

FIGURE 1 | Depicted are sagittal views on the left (L) and right hemisphere and the horizontal view on the averaged brain of the measured population at Talairach coordinates $X = 46$, -46 and $Z = -2$. Color coded are voxels that survived conservative significance level of $p < 0.05$ corrected for multiple comparisons (Bonferroni, Bonf) for the main factor AV congruency. The use of this very conservative significance level was necessary for isolation of separated clusters due to the fact of widespread activation clusters covering almost the whole brain resulting from strong congruency effect and spatial smoothing at the same significance level but corrected for multiple comparisons by means of the false discovery rate.

TABLE 2 | Significant brain clusters identified for the main factor AV congruency.

Regions	$\approx$BA	Hemisphere	Talairach center of mass			Cluster size (mm³)
			x	y	z	
Incongruent > congruent						
Inferior frontal gyrus	47	R	49	23	−3	621
Transverse temporal gyrus	41	R	48	−25	11	2592
Brainstem, Midbrain		R/L	−1	−20	0	2079
Medial frontal gyrus	8	L	−1	27	40	729
Transverse temporal gyrus	41	L	−44	−27	12	3996
Inferior frontal gyrus	47	L	−53	21	1	3051

p(Bonf) < 0.0001; BA Brodmann area; L left; R right.

for this cluster and each group and condition revealed a stronger activation in this area for HC subjects (AV-congruent: 2.47 ± 0.64; AV-incongruent: 2.44 ± 0.73) in comparison to AS subjects (AV-congruent: 1.35 ± 0.75; AV-incongruent: 1.30 ± 0.81), $F(1) = 21.05$, $p = 0.00$, $n^2{}_p = 0.41$. The analysis of interaction of both main factors revealed no significant clusters at the chosen threshold level.
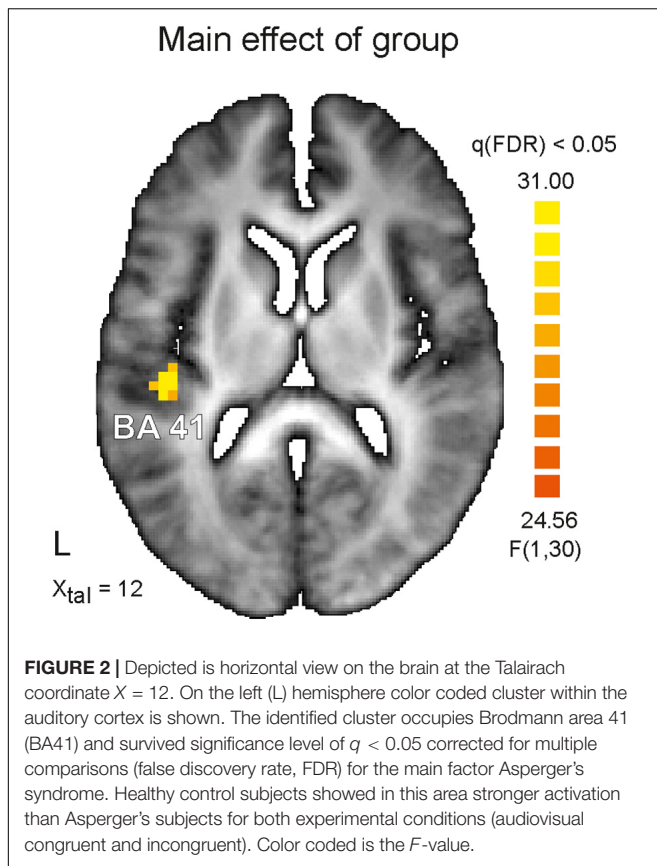
DISCUSSION

In this study we analyzed brain activity for AV congruent and incongruent speech stimuli in a group of adult AS subjects in comparison to typically developed healthy adult control subjects. We tested two competing hypotheses regarding multisensory processing deficits in ASD: Activity differences in brain areas typically associated with multisensory processing would indicate a primary deficit in multisensory processing occurring on later stages of the processing cascade. By contrast, group differences in unisensory brain areas would indicate rather low-level deficits which impact later multisensory processing.

The main finding of the study is a group difference in the left temporal cortex. As the mapping of the brain activity

to specific brain locations within auditory cortex by mean of Talairach space leads to some systematic errors (Gaschler-Markefski et al., 2006) with a strong tendency for caudal displacement of BA 41 this activity cluster can by assigned to BA 41 and/or BA 42. Nevertheless both areas are known to support unisensory auditory processing stages at cortical level. While AS subjects show less activation in this cluster than healthy controls, this difference is independent of the AV congruency of speech stimuli. This finding therefore suggests a difference in unisensory auditory processing at a lower level of the processing cascade in AS.

As in several other experiments employing similar stimuli (Szycik et al., 2009a,b; Rüsseler et al., 2018), a very strong congruency effect was observed in the present study which involved superior temporal and frontal areas. In particular, the superior temporal sulcus (STS) typically displays increased activation for AV stimuli compared to unimodal presentation and also shows congruency effects (Calvert et al., 2000; Sekiyama et al., 2003; Wright et al., 2003; Brefczynski-Lewis et al., 2009; Szycik et al., 2009a, 2012; Lee and Noppeney, 2011; Nath and Beauchamp, 2011, 2012; Stevenson et al., 2011). Also, the inferior frontal cortex has been repeatedly shown to be involved in AV integration which has been interpreted as providing a motor plan

FIGURE 2 | Depicted is horizontal view on the brain at the Talairach coordinate $X = 12$. On the left (L) hemisphere color coded cluster within the auditory cortex is shown. The identified cluster occupies Brodmann area 41 (BA41) and survived significance level of $q < 0.05$ corrected for multiple comparisons (false discovery rate, FDR) for the main factor Asperger's syndrome. Healthy control subjects showed in this area stronger activation than Asperger's subjects for both experimental conditions (audiovisual congruent and incongruent). Color coded is the F-value.

for the production of the phoneme that the listener is about to hear (Skipper et al., 2007).

The fact that no group by congruency interaction effect is present in the current data set would suggest that AS participants have alterations in unisensory processing stages. Whereas the current results suggest low-level auditory processing differences in AS, it is puzzling that these changes do not propagate to subsequent multimodal processing stages either leading to compensatory overactivation or to decreased activation due to insufficient input. Moreover, the current experiment was designed against a background of multiple studies that have shown behavioral AV integration deficits in AS, but did not entail a behavioral readout assessing multisensory integration. Therefore, it is impossible to assess whether the decreased activity in left temporal cortex in AS has a behavioral correlate, or represents in contrast to typically developed subjects just more efficient neuronal processing of the auditory input at this stage. To address these points, further experiments are necessary linking activity differences in auditory and multisensory areas to behavioral effects (e.g., Rüsseler et al., 2015).

Previous studies of auditory perception in adults with AS have shown abnormalities in unimodal auditory processing even on low-level stages of the auditory stream (Marco et al., 2011). Thus, similarly to our results reduced activity in the left auditory cortex to non-speech auditory stimuli has been seen in adult patients with ASD (Hames et al., 2016). Additionally, weaker ERP signals of subjects with ASD measured by EEG in the temporal

region and representing low-level cortical stages of the auditory system correlated to symptom severity in the autism spectrum (Brandwein et al., 2015).

On low-level stages of auditory processing the basic features of the incoming auditory information have to be extracted. These features may play a crucial role for successful integration of auditory and visual sensory input (e.g., Baumgart et al., 1999; de Heer et al., 2017). Indeed, previous studies have shown higher speech perception (about 3.5 dB) thresholds in AS subjects (Alcantara et al., 2004) which lead to a substantial decrease in speech recognition. Children with AS have been found to have considerable deficits in auditory stream segregation with reduced mismatch negativity amplitude when more than one sound stream is present (Lepisto et al., 2009) pointing to deficits in auditory processing. Similarly, neural responses to basic features of speech sounds within auditory cortex were found in other mismatch negativity experiment with AS children (Kujala et al., 2010).

To summarize, multisensory problems or problems in solving AV speech situations that subjects with AS often show may be related to deficits in ability to combine stimulus details or specific features into a coherent percept, as postulated by the "weak central coherence theory" (Happe and Frith, 2006). For AV speech stimuli this deficit seems to be related to alterations in unisensory auditory processing areas than in multisensory integration areas.

LIMITATIONS

One problem with the current design is that the current experiment did not entail a behavioral task that assessed multisensory integration and additional unisensory stimulation. We made this design decision, as the current paradigm has been proven quite useful in assessing audiovisual integration during speech perception in healthy participants (Szycik et al., 2009a), patients with schizophrenia (Szycik et al., 2009b) or participants with dyslexia (Rüsseler et al., 2018). For this study we used a presentation mode with a lower stimulation frequency which resulted in a small amount of stimuli and a long overall experimental period. A low stimulation rate was designed to give the BOLD response sufficient time to relax after each audiovisual stimulus. However, as we only observed alterations in unimodal auditory areas, a behavioral read-out and unisensory condition would be very helpful for interpretation.

As multiple studies have demonstrated eye-movement abnormalities in AS (Baron-Cohen et al., 2001a; Klin et al., 2002; Pelphrey et al., 2002; Dalton et al., 2005; Neumann et al., 2006; Norbury et al., 2009), the question arises whether altered eye-movements in AS might contribute to AV speech integration. As most previous studies, we did not determine gaze-behavior and eye-fixations in the current experiment. Thus, we cannot speak to this question. One previous study, however, did not find eye-movement changes in AS in an experiment assessing AV integration in speech processing (Saalasti et al., 2012).

Another important limitation of this study concerns the diagnosis of AS. Since our participants were collected from

a clinical population and the diagnostic procedures for this study started prior to widespread German adaptation of DSM-5, we used DSM-IV-TR diagnostic criteria for AS for all study participants. Therefore the results of this study refer strictly speaking only to a subpopulation of subjects with the diagnosis of ASD according to DSM-5.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the manuscript/**Supplementary Files**.

ETHICS STATEMENT

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional (Hannover Medical School) research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Informed consent was obtained from all individual participants included in the study.

AUTHOR CONTRIBUTIONS

F-AT participated in coordination of the study and the statistical analysis, performed the measurement, interpreted the data, and drafted the manuscript. LH and MR

participated in coordination of the study and performed the measurement. MZ participated in coordination of the study and interpreted the data. CS participated in coordination of the study and the statistical analysis, performed the measurement, and helped to draft the manuscript. DW and MW participated in the statistical analysis, interpreted the data, and helped to draft the manuscript. TM participated in study design and the statistical analysis, interpreted the data, and drafted the manuscript. GS conceived of the study, participated in design and coordination of the study and the statistical analysis, interpreted the data, and drafted the manuscript. All authors read and approved the final manuscript.

FUNDING

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2019.02286/full#supplementary-material

TABLE S1 | Stimulus list.

REFERENCES

Alcantara, J. I., Weisblatt, E. J., Moore, B. C., and Bolton, P. F. (2004). Speech-in-noise perception in high-functioning individuals with autism or Asperger's syndrome. *J. Child Psychol. Psychiatry* 45, 1107–1114. doi: 10.1111/j.1469-7610.2004.t01-1-00303.x

American Psychiatric Association (2000). *Diagnostic, and Statistical Manual of Mental Disorders DSM-IV*, 4th rev. Edn. Washington, DC: American Psychiatric Publication.

American Psychiatric Association (2013). *Diagnostic and Statistical Manual of Mental Disorders. DSM-5*, 5th Edn. Washington, DC: American Psychiatric Publication.

Baayen, R. H., Piepenbrock, R., and Van Rijn, H. (1993). *The CELEX Lexical Data Base on CDROM*. Philadephia, PA: Linguistic Data Consortium. https://catalog.ldc.upenn.edu/docs/LDC96L14/celex.readme.html (accessed September 28, 2019).

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., and Plumb, I. (2001a). The "Reading the Mind in the Eyes" Test revised version: a study with normal adults, and adults with Asperger syndrome or high-functioning autism. *J. Child Psychol. Psychiatry* 42, 241–251. doi: 10.1017/s0021963001006643

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., and Clubley, E. (2001b). The autism-spectrum quotient (AQ): evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *J. Autism Dev. Disord.* 31, 5–17.

Baumgart, F., Gaschler-Markefski, B., Woldorff, M. G., Heinze, H. J., and Scheich, H. (1999). A movement-sensitive area in auditory cortex. *Nature* 400, 724–726. doi: 10.1038/23385

Bebko, J. M., Schroeder, J. H., and Weiss, J. A. (2014). The McGurk effect in children with autism and asperger syndrome. *Autism Res.* 7, 50–59. doi: 10.1002/aur.1343

Bebko, J. M., Weiss, J. A., Demark, J. L., and Gomez, P. (2006). Discrimination of temporal synchrony in intermodal events by children with autism and children with developmental disabilities without autism. *J. Child Psychol. Psychiatry* 47, 88–98. doi: 10.1111/j.1469-7610.2005.01443.x

Beker, S., Foxe, J. J., and Molholm, S. (2018). Ripe for solution: delayed development of multisensory processing in autism and its remediation. *Neurosci. Biobehav. Rev.* 84, 182–192. doi: 10.1016/j.neubiorev.2017.11.008

Boynton, G. M., Engel, S. A., Glover, G. H., and Heeger, D. J. (1996). Linear systems analysis of functional magnetic resonance imaging in human V1. *J. Neurosci.* 16, 4207–4221. doi: 10.1523/jneurosci.16-13-04207.1996

Brandwein, A. B., Foxe, J. J., Butler, J. S., Frey, H. P., Bates, J. C., Shulman, L. H., et al. (2015). Neurophysiological indices of atypical auditory processing and multisensory integration are associated with symptom severity in autism. *J. Autism Dev. Disord.* 45, 230–244. doi: 10.1007/s10803-014-2212-9

Brandwein, A. B., Foxe, J. J., Butler, J. S., Russo, N. N., Altschuler, T. S., Gomes, H., et al. (2013). The development of multisensory integration in high-functioning autism: high-density electrical mapping and psychophysical measures reveal impairments in the processing of audiovisual inputs. *Cereb. Cortex* 23, 1329–1341. doi: 10.1093/cercor/bhs109

Brefczynski-Lewis, J., Lowitszch, S., Parsons, M., Lemieux, S., and Puce, A. (2009). Audiovisual non-verbal dynamic faces elicit converging fMRI and ERP responses. *Brain Topogr.* 21, 193–206. doi: 10.1007/s10548-009-0093-6

Calvert, G. A., Campbell, R., and Brammer, M. J. (2000). Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. *Curr. Biol.* 10, 649–657. doi: 10.1016/s0960-9822(00)00513-3

Dalton, K. M., Nacewicz, B. M., Johnstone, T., Schaefer, H. S., Gernsbacher, M. A., Goldsmith, H. H., et al. (2005). Gaze fixation and the neural circuitry of face processing in autism. *Nat. Neurosci.* 8, 519–526. doi: 10.1038/nn1421

De Gelder, B., Vroomen, J., Annen, L., Masthof, E., and Hodiamont, P. (2003). Audio-visual integration in schizophrenia. *Schizophr Res.* 59, 211–218. doi: 10.1016/S0920-9964(01)00344-9

de Heer, W. A., Huth, A. G., Griffiths, T. L., Gallant, J. L., and Theunissen, F. E. (2017). The hierarchical cortical organization of human speech processing. *J. Neurosci.* 37, 6539–6557. doi: 10.1523/JNEUROSCI.3267-16.2017

Diederich, A., and Colonius, H. (2015). The time window of multisensory integration: relating reaction times and judgments of temporal order. *Psychol. Rev.* 122, 232–241. doi: 10.1037/a0038696

Driver, J. (1996). Enhancement of selective listening by illusory mislocation of speech sounds due to lip-reading. *Nature* 381, 66–68. doi: 10.1038/381066a0

Feldman, J. I., Dunham, K., Cassidy, M., Wallace, M. T., Liu, Y., and Woynaroski, T. G. (2018a). Audiovisual multisensory integration in individuals with autism spectrum disorder: a systematic review and meta-analysis. *Neurosci. Biobehav. Rev.* 95, 220–234. doi: 10.1016/j.neubiorev.2018.09.020

Feldman, J. I., Kuang, W., Conrad, J. G., Tu, A., Santapuram, P., Simon, D. M., et al. (2018b). Brief report: differences in multisensory integration covary with sensory responsiveness in children with and without autism spectrum disorder. *J. Autism Dev. Disord.* 49, 397–403. doi: 10.1007/s10803-018-3667-x

Foss-Feig, J. H., Kwakye, L. D., Cascio, C. J., Burnette, C. P., Kadivar, H., Stone, W. L., et al. (2010). An extended multisensory temporal binding window in autism spectrum disorders. *Exp. Brain Res.* 203, 381–389. doi: 10.1007/s00221-010-2240-4

Foxe, J. J., Molholm, S., Del Bene, V. A., Frey, H. P., Russo, N. N., Blanco, D., et al. (2015). Severe multisensory speech integration deficits in high-functioning school-aged children with Autism Spectrum Disorder (ASD) and their resolution during early adolescence. *Cereb. Cortex* 25, 298–312. doi: 10.1093/cercor/bht213

Gaschler-Markefski, B., Brechmann, A., Szycik, G. R., Kaulisch, T., Baumgart, F., and Scheich, H. (2006). "Definition of human auditory cortex territories based on anatomical landmarks and fMRI activation," in *Dynamics of Speech Production and Perception*, eds P. L. Divenyi, S. Greenberg, and G. Meyer, (Amsterdam: IOS Press), 355–367.

Gelder, B. D., Vroomen, J., and van der Heide, L. (1991). Face recognition and lip-reading in autism. *Eur. J. Cogn. Psychol.* 3, 69–86. doi: 10.1080/09541449108406220

Genovese, C. R., Lazar, N. A., and Nichols, T. (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *Neuroimage* 15, 870–878. doi: 10.1006/nimg.2001.1037

Goebel, R., Esposito, F., and Formisano, E. (2006). Analysis of functional image analysis contest (FIAC) data with BrainVoyager QX: from single-subject to cortically aligned group general linear model analysis and self-organizing group independent component analysis. *Hum. Brain Mapp.* 27, 392–401. doi: 10.1002/hbm.20249

Gondan, M., Niederhaus, B., Rosler, F., and Roder, B. (2005). Multisensory processing in the redundant-target effect: a behavioral and event-related potential study. *Percept. Psychophys.* 67, 713–726. doi: 10.3758/bf03193527

Hahn, N., Foxe, J. J., and Molholm, S. (2014). Impairments of multisensory integration and cross-sensory learning as pathways to dyslexia. *Neurosci. Biobehav. Rev.* 47, 384–392. doi: 10.1016/j.neubiorev.2014.09.007

Hairston, W. D., Wallace, M. T., Vaughan, J. W., Stein, B. E., Norris, J. L., and Schirillo, J. A. (2003). Visual localization ability influences cross-modal bias. *J. Cogn. Neurosci.* 15, 20–29. doi: 10.1162/089892903321107792

Hames, E. C., Murphy, B., Rajmohan, R., Anderson, R. C., Baker, M., Zupancic, S., et al. (2016). Visual, auditory, and cross modal sensory processing in adults with autism: an EEG power and BOLD fMRI investigation. *Front. Hum. Neurosci.* 10:167. doi: 10.3389/fnhum.2016.00167

Happe, F., and Frith, U. (2006). The weak coherence account: detail-focused cognitive style in autism spectrum disorders. *J. Autism Dev. Disord.* 36, 5–25. doi: 10.1007/s10803-005-0039-0

Irwin, J. R., Tornatore, L. A., Brancazio, L., and Whalen, D. H. (2011). Can children with autism spectrum disorders "hear" a speaking face? *Child Dev.* 82, 1397–1403. doi: 10.1111/j.1467-8624.2011.01619.x

Jones, W., and Klin, A. (2013). Attention to eyes is present but in decline in 2-6-month-old infants later diagnosed with autism. *Nature* 504, 427–431. doi: 10.1038/nature12715

Klin, A., Jones, W., Schultz, R., Volkmar, F., and Cohen, D. (2002). Visual fixation patterns during viewing of naturalistic social situations as predictors of social competence in individuals with autism. *Arch. Gen. Psychiatry* 59, 809–816.

Kujala, T., Kuuluvainen, S., Saalasti, S., Jansson-Verkasalo, E., von Wendt, L., and Lepisto, T. (2010). Speech-feature discrimination in children with Asperger syndrome as determined with the multi-feature mismatch negativity paradigm. *Clin. Neurophysiol.* 121, 1410–1419. doi: 10.1016/j.clinph.2010.03.017

Kwakye, L. D., Foss-Feig, J. H., Cascio, C. J., Stone, W. L., and Wallace, M. T. (2011). Altered auditory and multisensory temporal processing in autism spectrum disorders. *Front. Integr. Neurosci.* 4:129. doi: 10.3389/fnint.2010.00129

Lancaster, J. L., Woldorff, M. G., Parsons, L. M., Liotti, M., Freitas, C. S., Rainey, L., et al. (2000). Automated Talairach atlas labels for functional brain mapping. *Hum. Brain. Mapp.* 10, 120–131. doi: 10.1002/1097-0193(200007)10:3<120::aid-hbm30>3.0.co;2-8

Lee, H., and Noppeney, U. (2011). Physical and perceptual factors shape the neural mechanisms that integrate audiovisual signals in speech comprehension. *J. Neurosci.* 31, 11338–11350. doi: 10.1523/JNEUROSCI.6510-10.2011

Lehrl, S., Triebig, G., and Fischer, B. (1995). Multiple choice vocabulary test MWT as a valid and short test to estimate premorbid intelligence. *Acta Neurol. Scand.* 91, 335–345. doi: 10.1111/j.1600-0404.1995.tb07018.x

Lepisto, T., Kuitunen, A., Sussman, E., Saalasti, S., Jansson-Verkasalo, E., Nieminen-von Wendt, T., et al. (2009). Auditory stream segregation in children with Asperger syndrome. *Biol. Psychol.* 82, 301–307. doi: 10.1016/j.biopsycho.2009.09.004

Marco, E. J., Hinkley, L. B., Hill, S. S., and Nagarajan, S. S. (2011). Sensory processing in autism: a review of neurophysiologic findings. *Pediatr. Res.* 69(5 Pt 2), 48R–54R. doi: 10.1203/PDR.0b013e3182130c54

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

Megnin, O., Flitton, A., Jones, C. R., de Haan, M., Baldeweg, T., and Charman, T. (2012). Audiovisual speech integration in autism spectrum disorders: ERP evidence for atypicalities in lexical-semantic processing. *Autism Res.* 5, 39–48. doi: 10.1002/aur.231

Mongillo, E. A., Irwin, J. R., Whalen, D. H., Klaiman, C., Carter, A. S., and Schultz, R. T. (2008). Audiovisual processing in children with and without autism spectrum disorders. *J. Autism Dev. Disord.* 38, 1349–1358. doi: 10.1007/s10803-007-0521-y

Nath, A. R., and Beauchamp, M. S. (2011). Dynamic changes in superior temporal sulcus connectivity during perception of noisy audiovisual speech. *J. Neurosci.* 31, 1704–1714. doi: 10.1523/JNEUROSCI.4853-10.2011

Nath, A. R., and Beauchamp, M. S. (2012). A neural basis for interindividual differences in the McGurk effect, a multisensory speech illusion. *Neuroimage* 59, 781–787. doi: 10.1016/j.neuroimage.2011.07.024

Neufeld, J., Sinke, C., Zedler, M., Emrich, H. M., and Szycik, G. R. (2012). Reduced audio-visual integration in synaesthetes indicated by the double-flash illusion. *Brain Res.* 1473, 78–86. doi: 10.1016/j.brainres.2012.07.011

Neumann, D., Spezio, M. L., Piven, J., and Adolphs, R. (2006). Looking you in the mouth: abnormal gaze in autism resulting from impaired top-down modulation of visual attention. *Soc. Cogn. Affect. Neurosci.* 1, 194–202. doi: 10.1093/scan/nsl030

Norbury, C. F., Brock, J., Cragg, L., Einav, S., Griffiths, H., and Nation, K. (2009). Eye-movement patterns are associated with communicative competence in autistic spectrum disorders. *J. Child Psychol. Psychiatry* 50, 834–842. doi: 10.1111/j.1469-7610.2009.02073.x

Pelphrey, K. A., Morris, J. P., and McCarthy, G. (2005). Neural basis of eye gaze processing deficits in autism. *Brain* 128(Pt 5), 1038–1048. doi: 10.1093/brain/awh404

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 Dev. Disord.* 32, 249–261.

Pierce, K., Conant, D., Hazin, R., Stoner, R., and Desmond, J. (2011). Preference for geometric patterns early in life as a risk factor for autism. *Arch. Gen. Psychiatry* 68, 101–109. doi: 10.1001/archgenpsychiatry.2010.113

Robertson, C. E., and Baron-Cohen, S. (2017). Sensory perception in autism. *Nat. Rev. Neurosci.* 18, 671–684. doi: 10.1038/nrn.2017.112

Ross, L. A., Saint-Amour, D., Leavitt, V. M., Javitt, D. C., and Foxe, J. J. (2007a). Do you see what I am saying? Exploring visual enhancement of speech comprehension in noisy environment. *Cereb. Cortex* 17, 1147–1153. doi: 10.1093/cercor/bhl024

Ross, L. A., Saint-Amour, D., Leavitt, V. M., Molholm, S., Javitt, D. C., and Foxe, J. J. (2007b). Impaired multisensory processing in schizophrenia: deficits in the visual enhancement of speech comprehension under noisy environmental conditions. *Schizophr Res.* 97, 173–183. doi: 10.1016/j.schres.2007.08.008

Rüsseler, J., Gerth, I., Heldmann, M., and Münte, T. F. (2015). Audiovisual perception of natural speech is impaired in adult dyslexics: an ERP study. *Neuroscience* 287, 55–65. doi: 10.1016/j.neuroscience.2014.12.023

Rüsseler, J., Ye, Z., Gerth, I., Szycik, G. R., and Münte, T. F. (2017). Audio-visual speech perception in adult readers with dyslexia: an fMRI study. *Brain Imaging Behav.* 12, 357–368. doi: 10.1007/s11682-017-9694-y

Rüsseler, J., Ye, Z., Gerth, I., Szycik, G. R., and Münte, T. F. (2018). Audio-visual speech perception in adult readers with dyslexia: an fMRI study. *Brain Imaging Behav.* 12, 357–368. doi: 10.1007/s11682-017-9694-y

Saalasti, S., Katsyri, J., Tiippana, K., Laine-Hernandez, M., von Wendt, L., and Sams, M. (2012). Audiovisual speech perception and eye gaze behavior of adults with asperger syndrome. *J. Autism Dev. Disord.* 42, 1606–1615. doi: 10.1007/s10803-011-1400-0

Saalasti, S., Lepisto, T., Toppila, E., Kujala, T., Laakso, M., Nieminen-von Wendt, T., et al. (2008). Language abilities of children with Asperger syndrome. *J. Autism Dev. Disord.* 38, 1574–1580. doi: 10.1007/s10803-008-0540-3

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

Schwartz, J. L., Berthommier, F., and Savariaux, C. (2004). Seeing to hear better: evidence for early audio-visual interactions in speech identification. *Cognition* 93, B69–B78. doi: 10.1016/j.cognition.2004.01.006

Sekiyama, K., Kanno, I., Miura, S., and Sugita, Y. (2003). Auditory-visual speech perception examined by fMRI and PET. *Neurosci. Res.* 47, 277–287. doi: 10.1016/s0168-0102(03)00214-1

Sinke, C., Neufeld, J., Zedler, M., Emrich, H. M., Bleich, S., Münte, T. F., et al. (2014). Reduced audiovisual integration in synesthesia - evidence from bimodal speech perception. *J. Neuropsychol.* 8, 94–106. doi: 10.1111/jnp.12006

Skipper, J. I., van Wassenhove, V., Nusbaum, H. C., and Small, S. L. (2007). Hearing lips and seeing voices: how cortical areas supporting speech production mediate audiovisual speech perception. *Cereb. Cortex* 17, 2387–2399. doi: 10.1093/cercor/bhl147

Stevenson, R. A., Baum, S. H., Segers, M., Ferber, S., Barense, M. D., and Wallace, M. T. (2017). Multisensory speech perception in autism spectrum disorder: From phoneme to whole-word perception. *Autism Res.* 10, 1280–1290. doi: 10.1002/aur.1776

Stevenson, R. A., Siemann, J. K., Woynaroski, T. G., Schneider, B. C., Eberly, H. E., Camarata, S. M., et al. (2014). Brief report: Arrested development of audiovisual speech perception in autism spectrum disorders. *J. Autism Dev. Disord.* 44, 1470–1477. doi: 10.1007/s10803-013-1992-7

Stevenson, R. A., VanDerKlok, R. M., Pisoni, D. B., and James, T. W. (2011). Discrete neural substrates underlie complementary audiovisual speech integration processes. *Neuroimage* 55, 1339–1345. doi: 10.1016/j.neuroimage.2010.12.063

Sumby, W. H., and Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *J. Acoust. Soc. Am.* 26, 212–215. doi: 10.1121/1.1907309

Szycik, G. R., Jansma, H., and Münte, T. F. (2009a). Audiovisual integration during speech comprehension: an fMRI study comparing ROI-based and whole brain analyses. *Hum. Brain Mapp.* 30, 1990–1999. doi: 10.1002/hbm.20640

Szycik, G. R., Münte, T. F., Dillo, W., Mohammadi, B., Samii, A., Emrich, H. M., et al. (2009b). Audiovisual integration of speech is disturbed in schizophrenia: an fMRI study. *Schizophr. Res.* 110, 111–118. doi: 10.1016/j.schres.2009.03.003

Szycik, G. R., Stadler, J., Tempelmann, C., and Münte, T. F. (2012). Examining the McGurk illusion using high-field 7 Tesla functional MRI. *Front. Hum. Neurosci.* 6:95. doi: 10.3389/fnhum.2012.00095

Szycik, G. R., Ye, Z., Mohammadi, B., Dillo, W., te Wildt, B. T., Samii, A., et al. (2013). Maladaptive connectivity of Broca's area in schizophrenia during audiovisual speech perception: an fMRI study. *Neuroscience* 253, 274–282. doi: 10.1016/j.neuroscience.2013.08.041

Talairach, J., and Tornoux, P. (1988). *Co-Planar Stereotactic Atlas of the Human Brain*. Stuttgart: Thieme.

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

Williams, J. H., Massaro, D. W., Peel, N. J., Bosseler, A., and Suddendorf, T. (2004). Visual-auditory integration during speech imitation in autism. *Res. Dev. Disabil.* 25, 559–575. doi: 10.1016/j.ridd.2004.01.008

Williams, L. E., Light, G. A., Braff, D. L., and Ramachandran, V. S. (2010). Reduced multisensory integration in patients with schizophrenia on a target detection task. *Neuropsychologia* 48, 3128–3136. doi: 10.1016/j.neuropsychologia.2010.06.028

Wittchen, H. U., Zaudig, M., and Fydrich, T. (1997). *Achse I: Psychische Störungen, in SKID Strukturiertes Klinisches Interview für DSM-IV*, 1 Edn. Göttingen: Hogrefe.

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

Wright, T. M., Pelphrey, K. A., Allison, T., McKeown, M. J., and McCarthy, G. (2003). Polysensory interactions along lateral temporal regions evoked by audiovisual speech. *Cereb. Cortex* 13, 1034–1043. doi: 10.1093/cercor/13.10.1034

Ye, Z., Rüsseler, J., Gerth, I., and Münte, T. F. (2017). Audiovisual speech integration in the superior temporal region is dysfunctional in dyslexia. *Neuroscience* 356, 1–10. doi: 10.1016/j.neuroscience.2017.05.017

Zhang, J., Meng, Y., He, J., Xiang, Y., Wu, C., Wang, S., et al. (2018). McGurk effect by individuals with autism spectrum disorder and typically developing controls: a systematic review and meta-analysis. *J. Autism Dev. Disord.* 49, 34–43. doi: 10.1007/s10803-018-3680-0