Schizophrenia Research 216 (2020) 175–183

Contents lists available at ScienceDirect

Schizophrenia Research

journal homepage: www.elsevier.com/locate/schres

Dynamic Causal Modelling suggests impaired effective connectivity in patients with schizophrenia spectrum disorders during gesture-speech integration

Adrian Wroblewski [a,b,*], Yifei He [a,b,c], Benjamin Straube [a,b]

[a] Translational Neuroimaging Marburg (TNM), Department of Psychiatry and Psychotherapy, University of Marburg, Germany
[b] Center for Mind, Brain and Behavior (CMBB), University of Marburg and Justus Liebig University Giessen, Germany
[c] Faculty of Translation, Language, and Cultural Studies, University of Mainz, Germersheim, Germany

ARTICLE INFO

Article history:
Received 10 October 2019
Received in revised form 26 November 2019
Accepted 15 December 2019
Available online 24 December 2019

Keywords:
fMRI
DCM
Multisensory-integration
Gesture
Speech
Schizophrenia

ABSTRACT

Integrating visual and auditory information during gesture-speech integration (GSI) is important for successful social communication, which is often impaired in schizophrenia. Several studies suggested the posterior superior temporal sulcus (pSTS) to be a relevant multisensory integration site. However, intact STS activation patterns were often reported in patients. Thus, here we used Dynamic Causal Modelling (DCM) to analyze whether information processing in schizophrenia spectrum disorders (SSD) is impaired during GSI on network level.

We investigated GSI in three different samples. First, we replicated a recently published connectivity model for GSI in a healthy subject group ($n = 19$). Second, we investigated differences between patients with SSD and a matched healthy control group ($n = 17$ each). Participants were presented videos of an actor performing intrinsically meaningful gestures accompanied by spoken sentences in German or Russian, or just telling a German sentence without gestures.

Across all groups, fMRI analyses revealed similar activation patterns, and DCM analyses resulted in the same winning model for GSI. This finding directly replicates previous results. However, patients revealed significantly reduced connectivity in the verbal pathway (from left middle temporal gyrus (MTG) to left STS). The clinical significance of this connection is supported by its correlations with the severity of concretism and a subscale of negative symptoms (SANS).

Our model confirms the importance of the pSTS as integration site during audio-visual integration. Patients showed generally intact connectivity during GSI, but revealed impaired information transfer via the verbal pathway. This might be the basis of interpersonal communication problems in patients with SSD.

1. Introduction

The integration of visual and auditory information during gesture-speech integration (GSI) is important for successful social communication (McNeill, 1992). Gestures are commonly used to emphasize or to provide additional meaning, thereby influencing speech perception (Kelly et al., 2010; Wu and Coulson, 2010) and memory processes (Straube et al., 2009, 2011). To achieve this, the listener needs to ultimately integrate information from both visual and auditory modalities. However, for patients with schizophrenia spectrum disorders (SSD), they might be impaired in this crucial integration process, which might contribute directly to dysfunctional social communication.

There is evidence that patients suffering from SSD show deficits in the perception and production of gestures (Bucci et al., 2008; Grüsser et al., 1990; Matthews et al., 2013; Nagels et al., 2019) and the processing of other complex actions (Takahashi et al., 2010). As for language, one characteristic aspect of language disturbances in SSD is concretism, which describes impaired understanding of metaphors, irony or abstract content in general (Bergemann et al., 2008; de Bonis et al., 1997; Iakimova et al., 2010; Kircher et al., 2007), and constitutes a clinical manifestation of a broader language dysfunction called Formal Thought Disorder (FTD). Besides, patients with SSD also show dysfunctional integration of visual and auditory information on a low cognitive level, i.e. lip movements (Beauchamp et al., 2010), as well as in more complex GSI processes (Surguladze et al., 2001; Szycik et al., 2009; Walther et al., 2013a; Walther and Mittal, 2016), especially for abstract (metaphoric) gestures (Cuevas et al., 2019; Straube et al., 2013a, 2013b). Together, these abnormalities might play an important role for deficits in social behavior in SSD (Lavelle et al., 2014). Despite

* Corresponding author at: Translational Neuroimaging Marburg (TNM), Department of Psychiatry and Psychotherapy, University of Marburg, Rudolf-Bultmann-Str. 8, 35039 Marburg, Germany.
E-mail address: Adrian.wroblewski@med.uni-marburg.de (A. Wroblewski).

https://doi.org/10.1016/j.schres.2019.12.005
0920-9964/

substantial report of dysfunctional audio-visual integration in SSD, the underlying brain mechanisms of these deficits are far from clear. Specifically, although the posterior superior temporal sulcus (pSTS), across domains, has been found to be crucial for multimodal integration (Beauchamp et al., 2004a, 2004b; Calvert, 2001; Dick et al., 2009, 2014; He et al., 2015; Holle et al., 2008; Hubbard et al., 2009; Kircher et al., 2009; Straube et al., 2011; Surguladze et al., 2001; Werner and Noppeney, 2010; Willems et al., 2009a, 2009b), interestingly, pSTS functioning has not always been found to be impaired in patients with SSD even during GSI (Straube et al., 2013b). However, multimodal integration is not only based on the activation of certain brain regions, but also relies on the information transfer between them (Nath and Beauchamp, 2011; Noppeney et al., 2008; Werner and Noppeney, 2010). So far, only few studies have investigated functional or effective connectivity during GSI. Straube and colleagues showed in a functional connectivity study that patients with schizophrenia and healthy controls share common connectivity from pSTS to middle temporal gyrus (MTG) and ventral inferior frontal gyrus (IFG) during integration of co-verbal iconic and metaphoric gestures, but, however, patients revealed reduced connectivity during metaphoric gesture processing between STS and bilateral frontal, parietal, and left temporal structures as well as the anterior cingulate cortex and subcortical regions (Straube et al., 2013b). In a recent study from our group, we suggested a three-region connectivity model for GSI, with bidirectional coupling between pSTS, visual gesture (occipital cortex; OC) and auditory speech processing (middle temporal gyrus; MTG) areas. The connectivity pattern revealed that the pSTS had inhibitory influence on the connectivity between OC and MTG, and thereby information transfer is strengthened towards the pSTS as integration site (Straube et al., 2018). This model for the first time corroborated the role of the pSTS during GSI on a network-level, by showing a top-down inhibition on direct verbal-gestural connectivity. While dysfunctional activation and seed-based (PPI) connectivity for processing of gestures in an abstract sentence context had been observed in patients with schizophrenia, the detailed information transfer between verbal, gestural and integration sites during GSI of simple intrinsically meaningful gestures and corresponding speech is unknown (Straube et al., 2013a, 2013b).

The purpose of the current study was to investigate impaired GSI in SSD on network-level by 1) replicating the connectivity model published in Straube et al. (2018) with the pSTS as central integration site in healthy subjects, and 2) using this model to identify differences between patients with SSD and a matched healthy control sample. We hypothesized that patients show impaired gestural or verbal information transfer towards the pSTS, which would be reflected in reduced MTG-STS or OC-STS coupling strength, respectively. We carried out two studies: study 1 was focusing on replicating the GSI connectivity model in a recently published sample of healthy participants (He et al., 2018b). In study 2 we collected data from patients with SSD and a matched healthy control group.

2. Methods

2.1. Participants

In study 1, 20 healthy participants (H-group) with no histories of medical or mental illnesses were included, of which one had to be excluded due to bad data quality, leading to a final sample of $n = 19$ (12 females, mean age 22.65 years, range 19–32 years). For more details about the H-group, please see (He et al., 2018b) where data were analyzed without considering connectivity analyses. In study 2, 17 medicated patients (P-group) with SSD and 18 matched healthy controls (C-group) were included, of which one had to be excluded due to missing activation in the needed brain regions for connectivity analyses, leading to a final sample in the C-group of $n = 17$. Demographic and clinical descriptions of the participants are summarized in Table 1. Patients were recruited at the Department of Psychiatry and Psychotherapy Marburg and diagnosed according to ICD-10 with schizophrenia (F20.0, $n = 13$, and F20.3, $n = 1$) or schizoaffective disorder (F25.0, $n = 2$, and F25.3, $n = 1$). All except one SSD patient received antipsychotic treatment; six were additionally treated with antidepressive or other psychiatric medication. Scale for the Assessment of Positive Symptoms (SAPS) (Andreasen, 1984a) and Scale for the Assessment of Negative Symptoms (SANS) (Andreasen, 1984b) were used to characterize negative and positive symptoms in patients. Exclusion criteria were substance use or alcohol abuse within the past six months, brain injury, and neurological diseases. In all samples we conducted neuropsychological tests to assess working memory function, digital span, trail making (TMT), verbal IQ (MWST-B) (Lehrl, 1999), and metaphoric language processing (concretism, evaluated with the Proverb Interpretation Task) (Barth and Küfferle, 2001) (Table 1). Additionally, we reported scores from subscales of SAPS and SANS, word fluency test, as well as a questionnaire on self-reported gesture production and perception [BAG, Brief Assessment of Gesture (Nagels et al., 2015)] in Supplementary Table 1. All participants were native German speakers without any knowledge of Russian language, and had normal or corrected to normal vision. In both C– and P-groups, 16/17 participants were right-handed. Written informed consent was obtained prior to the participation in the study. The study was approved by the local ethics committee.

Table 1
Demographic, medication, symptom, and neuropsychological measures in study 2 for patients (P-group) and matched healthy controls (C-group).

	P-group	C-group	Difference (p)
Age (years)	33.12 ± 12.35	32.65 ± 10.07	0.904
Gender male/female	13/4	13/4	
Education (years)	11.82 ± 1.77	12.76 ± 1.39	0.095
TMT A (seconds)	31.49 ± 10.73	26.45 ± 10.12	0.168
TMT B (seconds)	68.56 ± 37.8	54.18 ± 19.44	0.175
Digit Span forward	7.94 ± 1.75	8.06 ± 2.51	0.875
Digit Span backward	6.35 ± 1.93	6.59 ± 2.58	0.765
Verbal IQ	28.88 ± 5.25	28.47 ± 3.91	0.797
Concretism*	1.38 ± 0.46	1.13 ± 0.20	0.048
SAPS (global)	15 ± 6.89		
SANS (global)	9 ± 6.02		
CPZ Equivalent	562.52 ± 372.63		

Values are presented as mean ± standard deviation. TMT: trail making test; CPZ: chlorpromazine; * Concretism evaluated with the Proverb Interpretation Task.

2.2. Materials

In all three samples, we applied the same materials and procedures. The experimental paradigm has already been described in detail elsewhere (He et al., 2015, 2018b; Straube et al., 2018), so we will provide only a short overview here. During data acquisition, participants were presented short video sequences showing an actor performing an intrinsically meaningful gesture (IMG) accompanied by one spoken sentence in German or Russian language, or isolated speech only (Fig. 1). Two gesture types were presented, either with social (emblematic) or tool-related content. All videos lasted 5 s and contained only one simple sentence each. The actor performed either 1) an IMG in context of a German sentence (GG), 2) an IMG in context of a Russian sentence (GR), or 3) a German sentence with emblematic or tool-related content without being accompanied by any gesture (SG). In total 312 videos (52 per condition * 3 conditions * 2 sets) as well as 26 filler videos with Russian sentences and meaningless gestures were constructed. A similar approach to investigate GSI had been used previously by our group (Green et al., 2009; He et al., 2015; Kircher et al., 2009; Straube et al., 2011).

Fig. 1. Example of the GSI paradigm for the Gesture-German (GG) condition. The German sentence was translated into English for illustrative purposes.

2.3. Experimental procedure

Two experimental blocks were presented with a duration of 14 min and 182 videos each. The order of the videos was pseudo-randomized and counterbalanced across subjects. Each video sequence was followed by a gray screen with an average jitter of 4000 ms (2154-5846 ms). During the experiment, participants had to fulfill a content judgement task in which they had to decide (via button press with the index or middle finger of the left hand) whether the presented stimulus was rather object- or person-related. Participants were instructed to respond as soon as they decided on an answer. In all samples, simultaneous EEG-fMRI data were collected (He et al., 2018b), but the focus of the current study is effective connectivity based on fMRI data. Therefore, we will describe only this part briefly here.

2.4. MRI data acquisition

All MRI data were acquired using a 3-Tesla scanner (Siemens TIM Trio). Functional images were obtained using a T2*-weighted gradient-echo echo-planar imaging sequence sensitive for the BOLD contrast (TE = 30 ms, TR = 2000 ms, flip angle 90°, matrix size 64 × 64 voxels, slice thickness 4.0 mm, inter-slice gap 0.36 mm, FoV = 230 mm^2, 30 slices, ascending phase encoding direction). Slices were positioned transaxially parallel to the intercommissural (AC-PC) plane. In total, 425 volumes were collected per block. In addition, for each subject a high-resolution structural image was acquired using a three-dimensional T1-weighted magnetization-prepared rapid gradient-echo (3D MPRAGE) sequence in sagittal plane (TE = 2.26 ms, TR = 1900 ms, inversion time 900 ms, flip angle 9°, voxel size 1x1x1 mm^3, FOV = 256 mm^2, 176 slices).

2.5. Behavioral and correlational analyses

Statistical data analyses were performed using IBM SPSS v.20 for Windows. We conducted ANOVAs and paired one-tailed *t*-tests (Bonferroni-corrected) to investigate differences between groups and conditions. Furthermore, we used correlational analyses (Spearman correlation) to explore the relation between connectivity parameters and clinical and neuropsychological measures. Correlations were not corrected for multiple comparisons because of their exploratory nature.

2.6. Whole-brain fMRI data analysis

Functional MRI data were preprocessed and analyzed using the SPM12 software package (Statistical Parametric Mapping, Wellcome Trust Center for Neuroimaging, London; http://www.fil.ion.ucl.ac.uk) based on Matlab R2017a (version 9.2.0; MathWorks). After discarding the first five volumes to minimize T1-saturation effects, all images were high-pass filtered (cut-off period 128 s), co-registered to the anatomical T1 images, segmented, spatially realigned, normalized into the MNI space using the MNI template (resulting voxel size $2 \times 2 \times 2$ mm^3) and smoothed (6 mm isotropic Gaussian filter). The statistical whole-brain analysis was performed in a two-level, mixed-effects procedure. The voxel-wise BOLD activity was modeled by means of a single subject first-level General Linear Model (GLM) comprising the onsets of each event [critical word onset (Green et al., 2009; Kircher et al., 2009)] with a duration of 1 s. The six realignment parameters were additionally included as nuisance covariates to account for residual motion artifacts. The hemodynamic response was modeled by the canonical hemodynamic response function (HRF). Parameter estimate (ß-) images for the HRF were calculated for the two blocks per subject separately. Parameter estimates were then entered into a within-subject flexible factorial analysis including a subject-factor.

As in previous studies (Straube et al., 2018), we defined three experimental contrasts of interest. Firstly, two contrasts were created by subtracting the Gesture-Russian (GG > GR) for the "verbal effect" and the Speech-German from the Gesture-German condition (GG > SG) for the "gestural effect", respectively (Monte-Carlo simulation: $p = .001$, $k = 65$). Secondly, to investigate integration processes we then calculated the conjunction (conjunction null) (Friston et al., 2005; Nichols et al., 2005) between both unimodal conditions (GG > GR ∩ GG > SG; $p = .001$).

2.7. Dynamic Causal Modelling

To investigate GSI on network level we used Dynamic Causal Modelling (DCM; version 12 as implemented in SPM12) to analyze effective connectivity (Frässle et al., 2016; Tak et al., 2018). DCM relies on two steps for the connectivity analyses (Zeidman et al., 2019). Firstly, specific forward models are defined, which are based on concrete hypotheses about coupling and information transfer between a certain number of regions. In other words the method is used to investigate how one neuronal system exerts influence over another and how experimental conditions modulate this influence (Friston et al., 2003). In the first step of classical DCM for fMRI one needs to define certain parameters: 1) The impact of experimental conditions on specific regions (driving input; C-matrix), 2) the intrinsic, context-independent coupling between two regions (A-matrix), and 3) the influence of an experimental condition on the coupling strength between two regions (modulatory input, B-matrix). In non-linear DCM a fourth parameter is modeled that describes the modulatory influence one region exerts on the coupling between two other regions (Friston et al., 2000; Stephan et al., 2008). Secondly, after model specification random-effects Bayesian Model Selection (BMS) and Bayesian Model Averaging (BMA) are used to identify the model or group of models that fit the data best and to average parameters over subjects (Stephan et al., 2007). Resulting coupling parameters were entered into one-sample *t*-tests to be characterized by classical statistics (see Supplementary Table 4). A Kolmogorov-Smirnov test was used to test for normality on the main connections of interest (MTG-STS and OC-STS), indicating normal distribution in all relevant parameters. For the comparison between groups we used one-way ANOVAs and *t*-tests (unpaired, two-samples, one-tailed, Bonferroni-corrected).

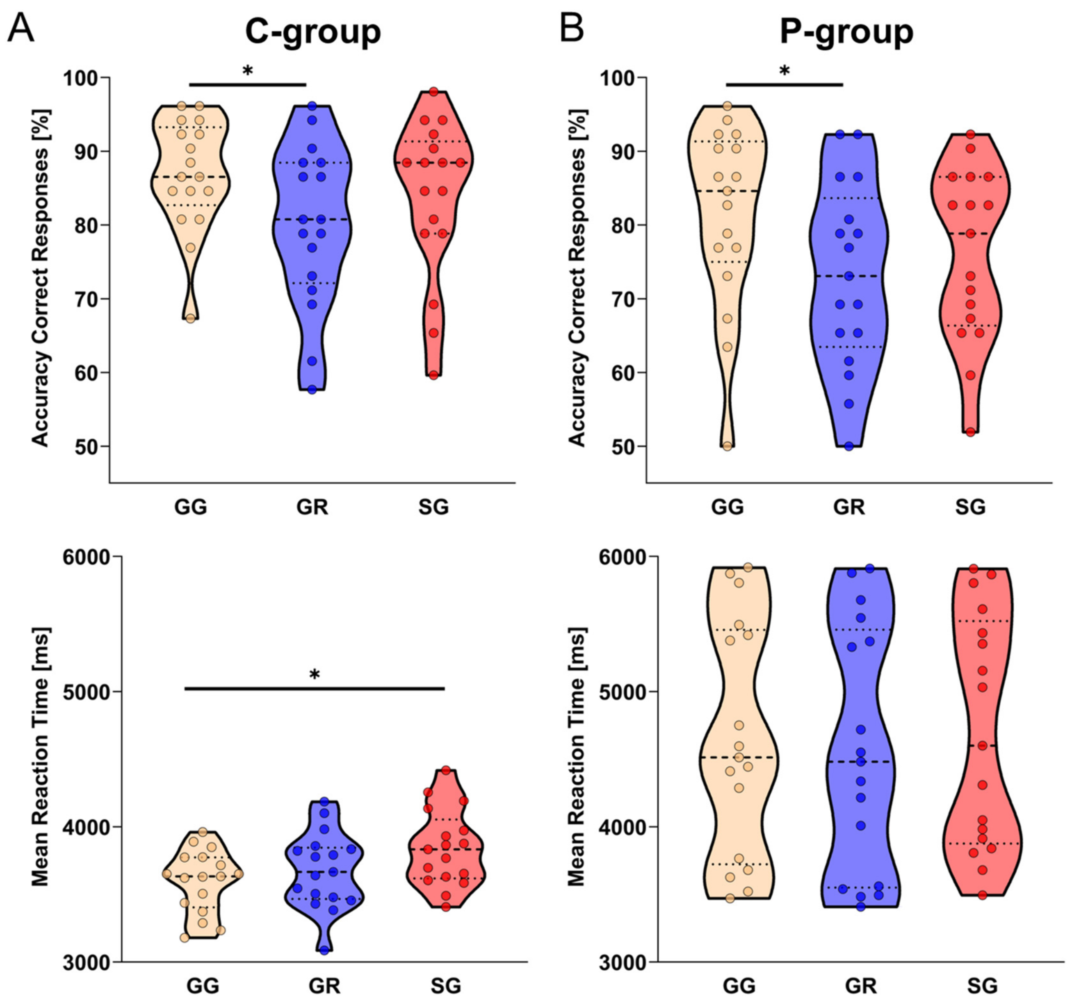

Fig. 2. Violin plots showing content judgement task performance in study 2 for the condition Gesture-German (GG), Gesture-Russian (GR) and Speech-German (SG). A: Results of the matched healthy controls (C-group). B: Results of the patients (P-group). Top: Distribution of accuracy rates of correct responses across subjects. Bottom: Distribution of reaction times across subjects. *Significant at $p < .05$ Bonferroni corrected.

2.8. Time series extraction and model space

Time series were extracted following a previous approach from our group (Straube et al., 2018), as well as the recommendations of Zeidman et al. (2019). In short we followed a five-step approach: 1) A GLM for every subject was defined to identify brain regions showing a main effect for the contrasts we described in Section 2.6. 2) On group level, we selected three brain region clusters (left STS, left MTG and left OC) that were used as masks in the single-subject analyses in the next step. Each of the three samples (H-, C– and P-group) was analyzed separately. 3) Using the previously created masks, we identified the peak coordinates for each individual subject and region (Supplementary Table 3). Furthermore, activation in the region of interest (ROI) had to exceed a liberal statistical threshold of $p < .01$ uncorrected. Note that the threshold was lowered to $p < .05$ uncorrected in case of no resulting activity (Zeidman et al., 2019) (one subject in C-group and four subjects in P-group). 4) Time series were then extracted as the first eigenvariates from all voxels inside a 4 mm sphere around the peak voxel, adjusted for the Effects-of-Interest contrast. Contrasts used for time series extraction were the same as described in 2.6. (left MTG: GG > GR, left OC: GG > SG, left pSTS: GG > GR ∩ GG > SG). 5) For model estimation new first-level design matrices were specified, containing an “integration regressor” (GG condition), a “verbal regressor”, encompassing all meaningful auditory speech information (GG and SG condition), and a “gestural regressor”, encompassing all meaningful visual gesture information (GG and GR condition).

The time series were then entered into the predefined models. For a detailed description of the model space please refer to Straube et al. (2018). In short, the model space consisted of 17 simple models with bidirectional intrinsic connections between all three nodes (A-matrix). As driving inputs we used the verbal regressor as input into the MTG and the gestural regressor was used as input into the OC (C-matrix). The integration regressor was used as modulatory input (B-matrix) and reflected several hypotheses about the integration processes during GSI. Furthermore, models M12 to M17 contained non-linear modulations of the STS onto unidirectional or bidirectional connectivity between MTG and OC.

3. Results

The results of the behavioral and fMRI analyses of the H-group have already been reported in a previous study (He et al., 2018b), so in Sections 3.1 and 3.2 we will focus only on the results from study 2.

3.1. Behavioral results

During the experiment participants had to fulfill a content judgement task in which they had to decide (via button press) whether the

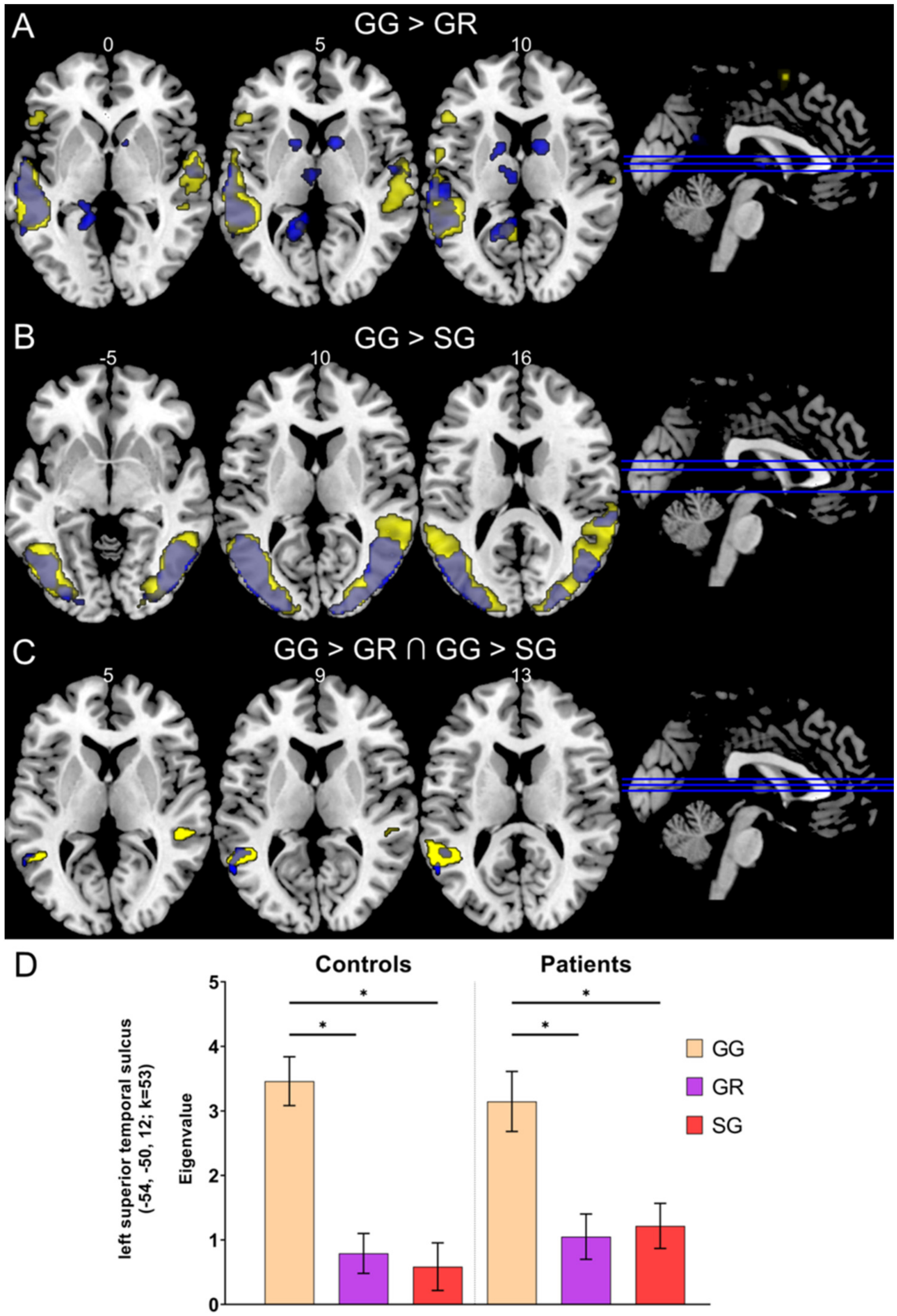

Fig. 3. The fMRI data analysis revealed similar activation patterns for patients (blue) and matched healthy controls (yellow) in study 2. A: BOLD activation during the processing of German speech reflected in the Gesture-German and Gesture-Russian (GG > GR) contrast. B: BOLD activation during the processing of gestures reflected in the Gesture-German and Speech-German (GG > SG) contrast. C: BOLD activation for the gesture-speech integration reflected in the conjunction of contrasts from (A) and (B) (GG > GR∩GG > SG). D: Activation pattern in the left superior temporal sulcus for controls and patients, resulting from the conjunction analysis (controls ∩ patients) of the contrast from (C).

presented stimulus was rather object- or person-related as soon as they knew the answer (Fig. 2). The main effect "Group" revealed that patients performed less accurate ($F_{(1,16)} = 3.328$, $p = .087$) on trend level and responded significantly slower ($F_{(1,16)} = 14.021$, $p = .002$) across conditions. Pairwise comparisons (Bonferroni-corrected) between conditions (GG > SG and GG > GR) revealed a significant difference in accuracy for GG > GR in controls ($t_{(16)} = 3.027$, $p = .008$) and patients ($t_{(16)} = 3.855$, $p = .001$). Furthermore, we found an interaction on trend level concerning reaction times (group X condition; $F_{(2,32)} = 2.492$, $p = .099$, $\eta^2 = .135$). The post-hoc analysis revealed that controls responded slower for GG < SG ($t_{(16)} = 4.105$, $p = .001$).

3.2. fMRI results

BOLD activation in study 2 revealed comparable activation patterns for C– and P-group in all contrasts, with overlapping activation in the left pSTS for the integration contrast across groups (conjunction analysis, MNIx,y,z: [−54, −50, 12], $t = 4.01$, k = 53, $p < .001$ uncorr.; see

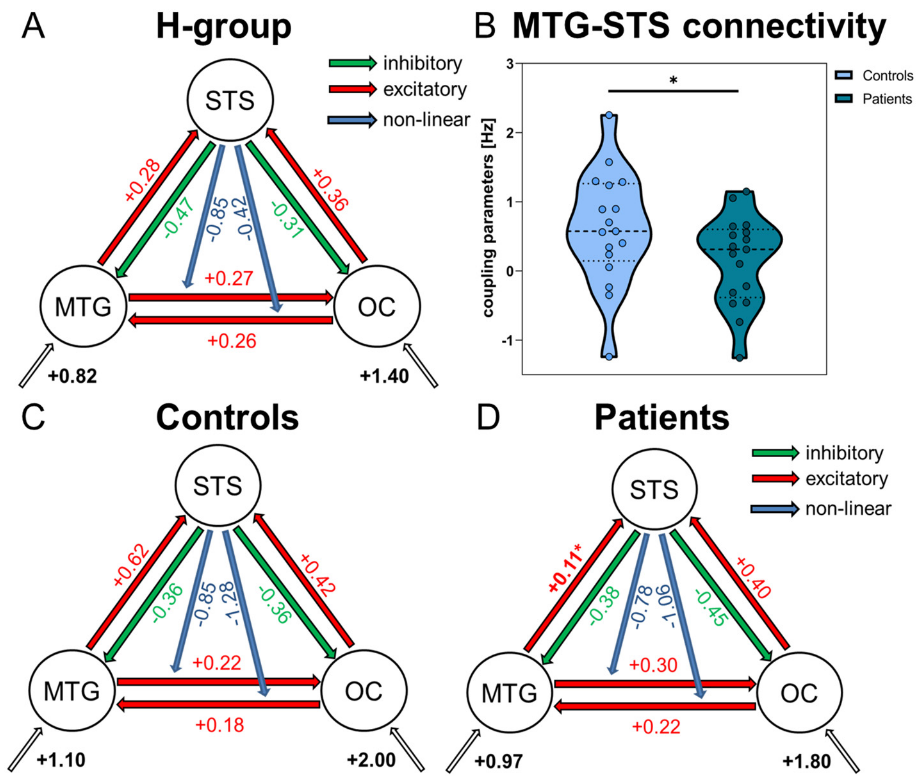

Fig. 4. DCM results. A: Averaged coupling parameters (A- and B-matrix) with excitatory (red), inhibitory (green) and non-linear (blue) connections for the H-group (B). B: Violin plots showing distribution of coupling parameters of the connection from middle temporal gyrus (MTG) to superior temporal sulcus (STS) for controls and patients. The comparison between the groups revealed a significantly weaker coupling strength between MTG and STS in patients. C-D: Averaged coupling parameters (A- and B-matrix) with excitatory (red), inhibitory (green) and non-linear (blue) connections for the controls (C) and patients (D). *Significant at $p < .05$ Bonferroni-corrected.

Fig. 3D). The main effect "group" revealed only activation in the fusiform gyrus (MNIx,y,z: [36, −52, −12], $F = 23.70$, k = 92, $p < .001$ corr.). The interaction between conditions (GG > GR and GG > SG, respectively) and group did not reach significance. In the contrast for verbal processing (GG > GR), we found activation in bilateral middle temporal gyrus (MTG) and posterior superior temporal sulcus (pSTS) in the C-group; in the P-group activation was found only in left MTG/pSTS (Fig. 3A). In the contrast for gestural processing (GG > SG) main activation was located in bilateral parts of the occipital cortex (OC), extending to posterior MTG/STS in both groups (Fig. 3B). The bimodal integration effect (GG > GR ∩ GG > SG) revealed activation mainly in the left STS in both groups (Fig. 3C). For a detailed overview over all activated clusters, please see Supplementary Table 2. These findings are a direct replication of previous fMRI studies using the same or a similar paradigm to investigate GSI (He et al., 2015, 2018b; Straube et al., 2018), showing a similar pattern for patients and healthy participants.

3.3. DCM results

3.3.1. Study 1

To determine the most likely of the 17 DCMs we implemented random-effects BMS. The analysis for the H-group revealed M17 (Fig. 4A) as the winning model (exceedance probability: 75%). Fig. 4B shows the resulting excitatory and inhibitory connections, as well as the non-linear modulation of the STS upon MTG-OC coupling after BMA. This pattern is a direct replication of previously published results (Straube et al., 2018).

3.3.2. Study 2

In both the C– and P-group BMS again revealed M17 as the winning model (exceedance probabilities: C-group: 77%; P-group: 85%). In Fig. 4C (C-group) and 4D (P-group) resulting connectivity parameters after BMA are visualized. Based on our hypothesis that patients show abnormal information transfer towards the STS, we compared the modulatory inputs, as well as the resulting parameters for MTG-STS and OC-STS connections between C– and P-group. Patients showed significantly weaker coupling from MTG to STS (B-matrix: $t_{(16)} = 2.095$, $p = .044$, A + B-matrix: $t_{(16)} = 2.063$, $p = .047$). No significant difference for OC-STS coupling was observed (B-matrix: $t_{(16)} = 0.1384$, $p = .891$, A + B-matrix: $t_{(16)} = 0.074$, $p = .942$). Detailed coupling parameters for all samples are provided in Supplementary Table 4.

3.4. Correlational analyses

To investigate the relation between MTG-STS connectivity strength with clinical and neuropsychological (word fluency and gesture perception) measures, we performed exploratory correlational analyses. Patients revealed negative correlations between MTG-STS connectivity and the severity of concretism ($r = -0.570$, $p = .017$), as well as between intrinsic MTG-STS coupling and the SANS attention subscale ($r = -0.498$, $p = .042$). Regarding BAG scores (Nagels et al., 2015), we found a negative correlation with between gesture perception score and the modulatory input to MTG-STS connection ($r = -0.490$, $p = .046$) in patients. Regarding word fluency we found positive correlations for WFT animals ($r = 0.626$, $p = .007$) and WFT with alternating content ($r = 0.519$, $p = .033$) in controls. Note that the analyses were not corrected for multiple comparisons because of their exploratory nature.

4. Discussion

Integration of auditory speech and visual gesture related information has profound implications for successful social communication, which is increasingly investigated in SSD. Despite substantial report of deficits in audio-visual and gesture-speech integration (GSI) in SSD (Cuevas et al., 2019; Straube et al., 2013a, 2013b; Surguladze et al., 2001; Szycik et al., 2009; Walther et al., 2013b, 2016; Walther and Mittal, 2016), underlying brain mechanisms are not yet understood. In the current study, we provided further evidence that 1) the pSTS plays a central role during GSI, by replicating a recently published connectivity model (Straube et al., 2018) in three different samples, 2) patients with SSD show largely similar underlying brain mechanisms for GSI, but 3) their MTG-STS connectivity seems to be impaired. This impairment might be the basis of dysfunctional integration of co-verbal intrinsically meaningful gestures and thus a possible explanation for interpersonal communication problems in SSD.

The left pSTS plays an important role in unspecific crossmodal integration of sensory inputs (Beauchamp et al., 2004b; Calvert, 2001), and is more specifically involved in the processing of speech and gesture information (Dick et al., 2009; Green et al., 2009; Holle et al., 2008, 2010; Kircher et al., 2009; Straube et al., 2011, 2018). Our data support this assumption, as main activation in the integration condition [reflected by the conjunction of the bimodal condition (GG) compared to both unimodal conditions (GR and SG)] was found in the left pSTS, as well as in both conditions of speech and gesture alone (Fig. 3). Despite behavioral evidence for dysfunctional multisensory integration in SSD (Surguladze et al., 2001; Szycik et al., 2009), several studies report no general schizophrenia-associated dysfunctions in brain regions relevant for the processing of speech and gesture (e.g. Straube et al., 2013b showed comparable effects in patients and healthy control subjects for integration of iconic coverbal gestures). This is in line with our findings in which we could show a similar activation pattern in patients and healthy subjects with differences only being present in the fusiform gyrus. Yet, in other studies reduced activation in the left IFG, STS or inferior parietal lobe (Thakkar et al., 2014; Viher et al., 2018) was reported, which we could not find. This could be due to experimental differences, as we looked directly at the difference between multimodal vs. unimodal conditions, while the cited experiments focused exclusively at gesture or hand action performance. Another possible explanation for missing differences between patients and controls on fMRI level could be our small sample size ($n = 17$). However, multimodal integration processes do not only rely on activation in discrete brain regions, but particularly on the information transfer between those regions (Nath and Beauchamp, 2011; Noppeney et al., 2008; Werner and Noppeney, 2010). Following the disconnectivity hypotheses in schizophrenia, some studies already investigated connectivity during GSI. So far, heterogeneous results have been reported. There is evidence that basic connectivity in patients is intact (Straube et al., 2013). Our data support this assumption, as the same winning model results from the connectivity analyses in patients and healthy subjects, suggesting similar processes of information transfer during GSI. Various resting-state and task-based studies, however, reported heterogeneous results concerning mainly Default Mode Network (DMN) connectivity. It can be assumed that SSD patients show increased functional (Camchong et al., 2011; Mannell et al., 2009; Salvador et al., 2010; Whitfield-Gabrieli et al., 2009) and structural (Camchong et al., 2011) connectivity in rather anterior regions (i.e. medial frontal cortex) of the DMN. In contrast, functional (Bluhm et al., 2007; Mannell et al., 2009) and effective connectivity (Zhou et al., 2018) of rather posterior regions (i.e. posterior cingulate) seems to be decreased in patients. Furthermore, Wu and colleagues reported decreased functional connectivity in areas of the language network in patients with schizophrenia (Wu et al., 2017), but again based on resting-state results. To date, only few studies have investigated connectivity during GSI in schizophrenia or SSD. It has been reported that patients showed decreased connectivity between frontal and temporal regions during the processing of metaphoric gestures (Jeong et al., 2009; Straube et al., 2013b). Recent reports suggested that temporal regions (i.e. pSTS) are more relevant in rather unspecific low-level integration of simple audio-visual input, e.g. simple sounds and visual stimuli, whereas frontal regions (i.e. IFG) are more relevant for rather complex semantic processing functions, e.g. evaluation of the presented metaphoric gesture (He et al., 2018a; Straube et al., 2018). This assumption is further supported by the correlation of concretism, i.e. the understanding of rather complex figurative speech, and activation of the IFG in schizophrenia (Straube et al., 2013b). Following the hypothesis that in SSD basic information transfer towards the integration site (pSTS) is disturbed, we found impaired connectivity from the speech-processing region (left MTG) towards the left STS, which means that patients showed decreased transfer of meaningful verbal information during GSI. The clinical significance of this connection is supported by its negative correlations with SANS subscales and concretism, suggesting that those patients who show stronger MTG-STS coupling have fewer problems with social attention and understanding metaphors. Several studies have already shown concretism, a characteristic feature of Formal Thought Disorders, to be one of the leading symptoms of SSD and being correlated to neuronal brain activation (Bergemann et al., 2008; Iakimova et al., 2010; Kircher et al., 2007). However, we are the first relating these symptoms to the connectivity pattern during GSI.

This study has to be interpreted in the light of some limitations. First, patients who participated were mostly medicated, which might have an influence on connectivity. Second, the model space we used is restricted to basic processes during GSI. Extending the model space to further relevant regions, e.g. the IFG (Dick et al., 2012; He et al., 2018b; Straube et al., 2013b) should be part of future approaches, for example on processing of metaphoric gestures and related abnormalities in SSD (Straube et al., 2013b). Including rather basic speech perception regions, e.g. Heschl's gyrus, planum temporale or medial geniculate body (Brodersen et al., 2011; Schofield et al., 2012), as well as investigating differences in interhemispheric effective connectivity between patients with SSD and healthy participants (Chang et al., 2015, 2019; Li et al., 2019; Steinmann et al., 2014) would be further possible extensions of our DCM model space.

Taken together our results demonstrate the relevance of the pSTS as integration site during GSI. We were able to replicate a model, which reflects basic processes of co-verbal speech integration in three independent samples (two samples including healthy participants and one including patients with SSD) and show, that these processes are generally intact in SSD. However, we provide evidence that patients might suffer from neural impairments within the verbal pathway during GSI, which is related to concretism and SANS subscores. This contributes to a better understanding of the neural basis for interpersonal communication problems in patients with SSD and a possible treatment target in future studies.

Contributors

Authors Y.H. and B.S. designed the study and wrote the protocol. Authors A.W. and Y.H. undertook the statistical analysis, and author A.W. wrote the first draft of the manuscript. All authors contributed to and have approved the final manuscript.

Role of the funding sources

This research project is supported by a grant from the “Von-Behring-Röntgen-Stiftung” (project no. 59-0002 and 64-0001) and by the “Deutsche Forschungsgemeinschaft” (project no. DFG: STR1146/11-2 & Ki588/6-2 and CRC/TRR 135/2 project A3). A.W. is supported by the Marburg University Research Academy. Y.H. is supported by the “Von-Behring-Röntgen-Stiftung” (project no. 64-0001). B.S. is supported by the DFG (project no. STR 1146/15-1). The funding sources were not further involved.

Declaration of competing interest

All authors declare that they have no conflicts of interest.

Acknowledgements

We thank the Core-Facility NeuroImaging for the support with the MRI data acquisition.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.schres.2019.12.005.

References

Andreasen, N.C., 1984a. The Scale for the Assessment of Positive Symptoms (SAPS). University of Iowa Press, Iowa.

Andreasen, N.C., 1984b. The Scale for the Assessment of Positive Symptoms (SANS). University of Iowa Press, Iowa.

Barth, A., Küfferle, B., 2001. Development of a proverb test for assessment of concrete thinking problems in schizophrenic patients. Nervenarzt 72, 853–858.

Beauchamp, M.S., Argall, B.D., Bodurka, J., Duyn, J.H., Martin, A., 2004a. Unraveling multisensory integration: patchy organization within human STS multisensory cortex. Nat. Neurosci. 7, 1190–1192.

Beauchamp, M.S., Lee, K.E., Argall, B.D., Martin, A., 2004b. Integration of auditory and visual information about objects in superior temporal sulcus. Neuron 41, 809–823.

Beauchamp, M.S., Nath, A.R., Pasalar, S., 2010. fMRI-guided transcranial magnetic stimulation reveals that the superior temporal sulcus is a cortical locus of the McGurk effect. J. Neurosci. 30, 2414–2417.

Bergemann, N., Parzer, P., Jaggy, S., Auler, B., Mundt, C., Maier-Braunleder, S., 2008. Estrogen and comprehension of metaphoric speech in women suffering from schizophrenia: results of a double-blind, placebo-controlled trial. Schizophr. Bull. 34, 1172–1181.

Bluhm, R.L., Miller, J., Lanius, R.A., Osuch, E.A., Boksman, K., Neufeld, R., Theberge, J., Schaefer, B., Williamson, P., 2007. Spontaneous low-frequency fluctuations in the BOLD signal in schizophrenic patients: anomalies in the default network. Schizophr. Bull. 33, 1004–1012.

de Bonis, M., Epelbaum, C., Deffez, V., Féline, A., 1997. The comprehension of metaphors in schizophrenia. Psychopathology 30, 149–154.

Brodersen, K.H., Schofield, T.M., Leff, A.P., Ong, C.S., Lomakina, E.I., Buhmann, J.M., Stephan, K.E., 2011. Generative embedding for model-based classification of fMRI data. PLoS Comput. Biol. 7, e1002079.

Bucci, S., Startup, M., Wynn, P., Baker, A., Lewin, T.J., 2008. Referential delusions of communication and interpretations of gestures. Psychiatry Res. 158, 27–34.

Calvert, G.A., 2001. Crossmodal processing in the human brain: insights from functional neuroimaging studies. Cereb. Cortex 11, 1110–1123.

Camchong, J., MacDonald, A.W., Bell, C., Mueller, B.A., Lim, K.O., 2011. Altered functional and anatomical connectivity in schizophrenia. Schizophr. Bull. 37, 640–650.

Chang, X., Xi, Y.-B., Cui, L.-B., Wang, H.-N., Sun, J.-B., Zhu, Y.-Q., Huang, P., Collin, G., Liu, K., Xi, M., Qi, S., Tan, Q.-R., Miao, D.-M., Yin, H., 2015. Distinct inter-hemispheric dysconnectivity in schizophrenia patients with and without auditory verbal hallucinations. Sci. Rep. 5, 11218.

Chang, X., Collin, G., Mandl, R.C.W., Cahn, W., Kahn, R.S., 2019. Interhemispheric connectivity and hemispheric specialization in schizophrenia patients and their unaffected siblings. NeuroImage Clin. 21, 101656.

Cuevas, P., Steines, M., He, Y., Nagels, A., Culham, J., Straube, B., 2019. The facilitative effect of gestures on the neural processing of semantic complexity in a continuous narrative. Neuroimage 195, 38–47.

Dick, A.S., Goldin-Meadow, S., Hasson, U., Skipper, J.I., Small, L., Small, S.L.C.-P., 2009. Co-speech gestures influence neural activity in brain regions associated with processing semantic information. Hum. Brain Mapp. 30, 3509–3526.

Dick, A.S., Goldin-Meadow, S., Solodkin, A., Small, S.L., 2012. Gesture in the developing brain. Dev. Sci. 15, 165–180.

Dick, A.S., Mok, E.H., Beharelle, A.R., Goldin-Meadow, S., Small, S.L., 2014. Frontal and temporal contributions to understanding the iconic co-speech gestures that accompany speech. Hum. Brain Mapp. 35, 900–917.

Frässle, S., Paulus, F.M., Krach, S., Jansen, A., 2016. Test-retest reliability of effective connectivity in the face perception network. Hum. Brain Mapp. 37, 730–744.

Friston, K.J., Mechelli, A., Turner, R., Price, C.J., 2000. Nonlinear responses in fMRI: the balloon model, Volterra kernels, and other hemodynamics. Neuroimage 12, 466–477.

Friston, K.J., Harrison, L., Penny, W., 2003. Dynamic causal modelling. Neuroimage 19, 1273–1302.

Friston, K.J., Penny, W.D., Glaser, D.E., 2005. Conjunction revisited. Neuroimage 25, 661–667.

Green, A., Straube, B., Weis, S., Jansen, A., Willmes, K., Konrad, K., Kircher, T., 2009. Neural integration of iconic and unrelated coverbal gestures: a functional MRI study. Hum. Brain Mapp. 30, 3309–3324.

Grüsser, O.J., Kirchhoff, N., Naumann, A., 1990. Brain mechanisms for recognition of faces, facial expression, and gestures: neuropsychological and electroencephalographic studies in normals, brain-lesioned patients, and schizophrenics. Res. Publ. Assoc. Res. Nerv. Ment. Dis. 67, 165–193.

He, Y., Gebhardt, H., Steines, M., Sammer, G., Kircher, T., Nagels, A., Straube, B., 2015. The EEG and fMRI signatures of neural integration: an investigation of meaningful gestures and corresponding speech. Neuropsychologia 72, 27–42.

He, Y., Nagels, A., Schlesewsky, M., Straube, B., 2018a. The role of gamma oscillations during integration of metaphoric gestures and abstract speech. Front. Psychol. 9, 1–11.

He, Y., Steines, M., Sommer, J., Gebhardt, H., Nagels, A., Sammer, G., Kircher, T., Straube, B., 2018b. Spatial – temporal dynamics of gesture – speech integration : a simultaneous EEG-fMRI study. Brain Struct. Funct. 223, 3073–3089.

Holle, H., Gunter, T.C., Rüschemeyer, S.-A., Hennenlotter, A., Iacoboni, M., 2008. Neural correlates of the processing of co-speech gestures. Neuroimage 39, 2010–2024.

Holle, H., Obleser, J., Rueschemeyer, S.-A., Gunter, T.C., 2010. Integration of iconic gestures and speech in left superior temporal areas boosts speech comprehension under adverse listening conditions. Neuroimage 49, 875–884.

Hubbard, A.L., Wilson, S.M., Callan, D.E., Dapretto, M., 2009. Giving speech a hand: gesture modulates activity in auditory cortex during speech perception. Hum. Brain Mapp. 30, 1028–1037.

Iakimova, G., Passerieux, C., Denhière, G., Laurent, J.-P., Vistoli, D., Vilain, J., Hardy-Baylé, M.-C., 2010. The influence of idiomatic salience during the comprehension of ambiguous idioms by patients with schizophrenia. Psychiatry Res. 177, 46–54.

Jeong, B., Wible, C.G., Hashimoto, R.-I., Kubicki, M., 2009. Functional and anatomical connectivity abnormalities in left inferior frontal gyrus in schizophrenia. Hum. Brain Mapp. 30, 4138–4151.

Kelly, S.D., Ozyürek, A., Maris, E., 2010. Two sides of the same coin: speech and gesture mutually interact to enhance comprehension. Psychol. Sci. 21, 260–267.

Kircher, T.T.J., Leube, D.T., Erb, M., Grodd, W., Rapp, A.M., 2007. Neural correlates of metaphor processing in schizophrenia. Neuroimage 34, 281–289.

Kircher, T., Straube, B., Leube, D., Weis, S., Sachs, O., Willmes, K., Konrad, K., Green, A., 2009. Neural interaction of speech and gesture: differential activations of metaphoric co-verbal gestures. Neuropsychologia 47, 169–179.

Lavelle, M., Healey, P.G.T., McCabe, R., 2014. Nonverbal behaviour during face-to-face social interaction in schizophrenia: a review. J. Nerv. Ment. Dis. 202, 47–54.

Lehrl, S., 1999. Mehrfachwahl-Wortschatz-Intelligenztest. MWT-B.

Li, S., Hu, N., Zhang, W., Tao, B., Dai, J., Gong, Y., Tan, Y., Cai, D., Lui, S., 2019. Dysconnectivity of multiple brain networks in schizophrenia: a meta-analysis of resting-state functional connectivity. Front. Psych. 10, 482.

Mannell, M.V., Franco, A.R., Calhoun, V.D., Cañive, J.M., Thoma, R.J., Mayer, A.R., 2009. Resting state and task-induced deactivation: a methodological comparison in patients with schizophrenia and healthy controls. Hum. Brain Mapp. 31 (NA-NA).

Matthews, N., Gold, B.J., Sekuler, R., Park, S., 2013. Gesture imitation in schizophrenia. Schizophr. Bull. 39, 94–101.

McNeill, D., 1992. Hand and Mind: What Gestures Reveal About Thought. Univ. of Chicago Press, Chicago.

Nagels, A., Kircher, T., Steines, M., Grosvald, M., Straube, B., 2015. A brief self-rating scale for the assessment of individual differences in gesture perception and production. Learn. Individ. Differ. 39, 73–80.

Nagels, A., Kircher, T., Grosvald, M., Steines, M., Straube, B., 2019. Evidence for gesture-speech mismatch detection impairments in schizophrenia. Psychiatry Res. 273, 15–21.

Nath, A.R., Beauchamp, M.S., 2011. Dynamic changes in superior temporal sulcus connectivity during perception of noisy audiovisual speech. J. Neurosci. 31, 1704–1714.

Nichols, T., Brett, M., Andersson, J., Wager, T., Poline, J.-B., 2005. Valid conjunction inference with the minimum statistic. Neuroimage 25, 653–660.

Noppeney, U., Josephs, O., Hocking, J., Price, C.J., Friston, K.J., 2008. The effect of prior visual information on recognition of speech and sounds. Cereb. Cortex 18, 598–609.

Salvador, R., Sarró, S., Gomar, J.J., Ortiz-Gil, J., Vila, F., Capdevila, A., Bullmore, E., McKenna, P.J., Pomarol-Clotet, E., 2010. Overall brain connectivity maps show cortico-subcortical abnormalities in schizophrenia. Hum. Brain Mapp. 31, 2003–2014.

Schofield, T.M., Penny, W.D., Stephan, K.E., Crinion, J.T., Thompson, A.J., Price, C.J., Leff, A.P., 2012. Changes in auditory feedback connections determine the severity of speech processing deficits after stroke. J. Neurosci. 32, 4260–4270.

Steinmann, S., Leicht, G., Mulert, C., 2014. Interhemispheric auditory connectivity: structure and function related to auditory verbal hallucinations. Front. Hum. Neurosci. 8, 55.

Stephan, K.E., Weiskopf, N., Drysdale, P.M., Robinson, P.A., Friston, K.J., 2007. Comparing hemodynamic models with DCM. Neuroimage 38, 387–401.

Stephan, K.E., Kasper, L., Harrison, L.M., Daunizeau, J., den Ouden, H.E.M., Breakspear, M., Friston, K.J., 2008. Nonlinear dynamic causal models for fMRI. Neuroimage 42, 649–662.

Straube, B., Green, A., Weis, S., Chatterjee, A., Kircher, T., 2009. Memory effects of speech and gesture binding: cortical and hippocampal activation in relation to subsequent memory performance. J. Cogn. Neurosci. 21, 821–836.

Straube, B., Green, A., Bromberger, B., Kircher, T., 2011. The differentiation of iconic and metaphoric gestures: common and unique integration processes. Hum. Brain Mapp. 32, 520–533.

Straube, B., Green, A., Sass, K., Kircher, T., 2013. Superior temporal sulcus disconnectivity during processing of metaphoric gestures in schizophrenia. Schizophr. Bull. 40, 936–944.

Straube, B., Green, A., Sass, K., Kirner-Veselinovic, A., Kircher, T., 2013b. Neural integration of speech and gesture in schizophrenia: evidence for differential processing of metaphoric gestures. Hum. Brain Mapp. 34, 1696–1712.

Straube, B., Wroblewski, A., Jansen, A., He, Y., 2018. The connectivity signature of co-speech gesture integration: the superior temporal sulcus modulates connectivity between areas related to visual gesture and auditory speech processing. Neuroimage 181, 539–549.

Surguladze, S.A., Calvert, G.A., Brammer, M.J., Campbell, R., Bullmore, E.T., Giampietro, V., David, A.S., 2001. Audio-visual speech perception in schizophrenia: an fMRI study. Psychiatry Res. 106, 1–14.

Szycik, G.R., Münte, T.F., Dillo, W., Mohammadi, B., Samii, A., Emrich, H.M., Dietrich, D.E., 2009. Audiovisual integration of speech is disturbed in schizophrenia: an fMRI study. Schizophr. Res. 110, 111–118.

Tak, S., Noh, J., Cheong, C., Zeidman, P., Razi, A., Penny, W.D., Friston, K.J., 2018. A validation of dynamic causal modelling for 7T fMRI. J. Neurosci. Methods 305, 36–45.
Takahashi, H., Kato, M., Sassa, T., Shibuya, T., Koeda, M., Yahata, N., Matsuura, M., Asai, K., Suhara, T., Okubo, Y., 2010. Functional deficits in the extrastriate body area during observation of sports-related actions in schizophrenia. Schizophr. Bull. 36, 642–647.
Thakkar, K.N., Peterman, J.S., Park, S., 2014. Altered brain activation during action imitation and observation in schizophrenia: a translational approach to investigating social dysfunction in schizophrenia. Am. J. Psychiatry 171, 539–548.
Viher, P.V., Stegmayer, K., Kubicki, M., Karmacharya, S., Lyall, A.E., Federspiel, A., Vanbellingen, T., Bohlhalter, S., Wiest, R., Strik, W., Walther, S., 2018. The cortical signature of impaired gesturing: findings from schizophrenia. NeuroImage Clin. 17, 213–221.
Walther, S., Mittal, V.A., 2016. Why we should take a closer look at gestures. Schizophr. Bull. 42, 259–261.
Walther, S., Vanbellingen, T., Müri, R., Strik, W., Bohlhalter, S., 2013a. Impaired gesture performance in schizophrenia: particular vulnerability of meaningless pantomimes. Neuropsychologia 51, 2674–2678.
Walther, S., Vanbellingen, T., Müri, R., Strik, W., Bohlhalter, S., 2013b. Impaired pantomime in schizophrenia: association with frontal lobe function. Cortex 49, 520–527.
Walther, S., Eisenhardt, S., Bohlhalter, S., Vanbellingen, T., Müri, R., Strik, W., Stegmayer, K., 2016. Gesture performance in schizophrenia predicts functional outcome after 6 months. Schizophr. Bull. 42, 1326–1333.
Werner, S., Noppeney, U., 2010. Distinct functional contributions of primary sensory and association areas to audiovisual integration in object categorization. J. Neurosci. 30, 2662–2675.
Whitfield-Gabrieli, S., Thermenos, H.W., Milanovic, S., Tsuang, M.T., Faraone, S.V., McCarley, R.W., Shenton, M.E., Green, A.I., Nieto-Castanon, A., LaViolette, P., Wojcik, J., Gabrieli, J.D.E., Seidman, L.J., 2009. Hyperactivity and hyperconnectivity of the default network in schizophrenia and in first-degree relatives of persons with schizophrenia. Proc. Natl. Acad. Sci. 106, 1279–1284.
Willems, R.M., Ozyürek, A., Hagoort, P., 2009a. Differential roles for left inferior frontal and superior temporal cortex in multimodal integration of action and language. Neuroimage 47, 1992–2004.
Willems, R.M., Toni, I., Hagoort, P., Casasanto, D., 2009b. Body-specific motor imagery of hand actions: neural evidence from right- and left-handers. Neuroscience 3, 1–9.
Wu, Y.C., Coulson, S., 2010. Gestures modulate speech processing early in utterances. Neuroreport 21, 522–526.
Wu, X., Zeng, L.-L., Shen, H., Yuan, L., Qin, J., Zhang, P., Hu, D., 2017. Functional network connectivity alterations in schizophrenia and depression. Psychiatry Res. Neuroimaging 263, 113–120.
Zeidman, P., Jafarian, A., Corbin, N., Seghier, M.L., Razi, A., Price, C.J., Friston, K.J., 2019. A guide to group effective connectivity analysis, part 1: first level analysis with DCM for fMRI. Neuroimage 200, 174–190.
Zhou, Y., Zeidman, P., Wu, S., Razi, A., Chen, C., Yang, L., Zou, J., Wang, G., Wang, H., Friston, K.J., 2018. Altered intrinsic and extrinsic connectivity in schizophrenia. NeuroImage Clin. 17, 704–716.