This is the accepted manuscript of Alonso-Sánchez MF, Limongi R, Gati J, Palaniyappan L. Language network self-inhibition and semantic similarity in first-episode schizophrenia: A computational-linguistic and effective connectivity approach. Schizophr Res. 2022 May 11:S0920-9964(22)00160-8. doi: 10.1016/j.schres.2022.04.007. Licensed CC-BY-NC-ND. https://creativecommons.org/licenses/by-nc-nd/4.0/

Language network self-inhibition and semantic similarity in first-episode schizophrenia: A computational-linguistic and effective connectivity approach

Authors: María Francisca Alonso-Sánchez[1,2]*, Roberto Limongi [2], Joseph Gati[2-5] and Lena Palaniyappan [2-5].

[1] CIDCL, Fonoaudiología, Facultad de Medicina, Universidad de Valparaíso, Chile.

[2] Robarts Research Institute, Western University, London Ontario, Canada

[3] Centre for Youth Mental Health Service Innovation, Research and Training, Douglas Mental Health University Institute, Montreal, Quebec, Canada

[4] Department of Psychiatry, McGill University, Montreal, Quebec, Canada.

[5] Department of Medical Biophysics, Schulich School of Medicine and Dentistry, Western University, London Ontario, Canada.

*Corresponding author: mariafrancisca.alonso@uv.cl, Robarts Research Institute, 1151 Richmond St., UWO, London, Ontario, Canada, N6A 5B7

Abstract Word Count: 250

Total Word Count: 3906

Abstract

Introduction: A central feature of schizophrenia is the disorganisation and impoverishment of language. Recently, we observed higher semantic similarity in first-episode-schizophrenia (FES) patients. In this study, we investigate if this aberrant similarity relates to the 'causal' connectivity between two key nodes of the word production system: inferior frontal gyrus (IFG) and the semantic-hub at the ventral anterior temporal lobe (vATL). **Methods:** Resting-state fMRI scans were collected from 60 participants (30 untreated FES and 30 healthy controls). The semantic distance was measured with the CoVec semantic tool based on GloVe. A spectral dynamic causal model with Parametrical Empirical Bayes was constructed modelling the intrinsic self-inhibitory and extrinsic-excitatory connections within the brain regions. We estimated the parameters of a fully connected model with the semantic distance as a covariate. **Results:** FES patients chose words with higher semantic similarity when describing the pictures compared to the HC group. Among patients, an increased semantic similarity was related with an increase in intrinsic connections within both the vATL and IFG, suggesting that reduced 'synaptic gain' in these regions likely contribute to aberrant sampling of the semantic space during discourse in schizophrenia. **Conclusions:** Lexical impoverishment relates to increased self-inhibition in both the IFG and vATL. The associated reduction in synaptic gain may relate to reduced precision of locally generated neural activity, forcing the choice of words that are already 'activated' in a lexical network. One approach to improve word sampling may be via promoting synaptic gain via supra-physiological stimulation within the Broca's-vATL network; this proposal needs verification.

Keywords: psychosis; lexical access; fMRI; Spectral Dynamic Causal Modeling; Broca's area; disorganization; formal thought disorder

1. Introduction

When we speak, we choose words relevant to the specific context by drawing on our knowledge of what those words represent. Semantic memory is seen as the storage of words (”bird”) and many related concepts they represent (they fly, chirp, have feathers or even that some of them only swim, but they are still in the same category). Word selection process involves overcoming competition (i.e., resolution of conflicts) but this occurs swiftly without affecting the efficiency of a discourse. A top-down control is said to drive the lexical unit selection within the semantic system while preserving the thematic progression of discourse aligned to the communicational goal (Levelt, 2001). The rapid and precise production of diverse words is critical to fluent language; restricting the use of words to a limited repertoire of concepts may interfere with the efficiency of communication.

Semantic impairment is a prominent linguistic feature of schizophrenia (Brown & Kuperberg, 2015). Previous studies have demonstrated a failure to inhibit the ‘retrieval’ or ‘activation’ i.e., the use of conceptually related words that are not appropriate for the context in patients with schizophrenia (Almeida & Radanovic, 2021). This increased semantic priming is seen especially in those with clinically detectable thought disorder (Pomarol-Clotet et al., 2008), with a higher degree of similarity between words contributing to redundant discourse with reduced information content (Alonso-Sánchez et al., 2022). Semantic deficits have been generally described as part of disorganization symptoms (Elvevåg et al., 2007; Hoffman et al., 2020) and poverty of content (e.g., lower semantic density(Rezaii et al., 2019)). These features often appear before the onset of psychosis (Mota et al., 2017; Spencer et al., 2021) and relate to social and occupational outcomes in early stages of the illness (Palaniyappan et al., 2019).

A wide variety of computational tools (Natural Language Processing, a.k.a. NLP) are now accessible to detect speech features that are not readily perceptible to the human eye/ear. These tools can readily capture surface features such as word repetitions (type-token ratio), and also

unravel the content of the texts and the relationship between the words that compose them (Levy et al., 2015; Mikolov et al., 2013). One of these approaches, termed distributional semantics, clusters words into neighbourhoods based on the frequency of their co-occurrence in everyday language (e.g. 'snack' and 'plate') while distancing widely those that are highly improbable to appear together ('snack' and 'planet'). Specifically, the GloVe semantic similarity tool benefits from a high-dimensional word vectorization framework and is based on a log-bilinear regression model that uses both local and global statistics of a corpus, unlike other methods (Word2Vec uses only local while Latent Semantic Analysis uses only global statistics) (Pennington et al., 2014). The local statistic computation captures the relation between words, as shown in the analogy task (king - man + woman = queen) while the global statistics profit from all the document and not only by the surrounding words. This approach computes the ratio of co-occurrence probabilities between adjacent words within a frame. Thus, we can estimate the cosine similarity within words to unravel the semantic diversity in a discourse sample. In our study, we infer the lexical unit selection from the putative semantic library using the GloVe cosine similarity from a sample of a descriptive discourse (Kanerva, 2009; Pennington et al., 2014).

The anatomical aspect of brain network recruitment in lexical/semantic tasks is still a matter of intense debate, but a growing body of evidence supports the view that the ventral Anterior Temporal Lobe (vATL) is critical for semantic cognition, constituting a putative Semantic Hub (Visser et al., 2010). The hub is part of a larger network that comprises the uncinate fasciculus, frontal-occipital fasciculus and ventral occipitotemporal areas (Chiou et al., 2018; Jackson et al., 2019; Krieger-Redwood & Jefferies, 2014). The Inferior Frontal Gyrus (IFG) is related to the regulation of automatic spreading activation/inhibition of candidates/competitors within the lexical network (Chiou et al., 2018; Python et al., 2018). The top-down control mediated by IFG contributes to both lexical retrieval (the activation of several possible relevant responses) and

selection (the narrowing of choices to the most appropriate one) (Conner et al., 2019; Snyder et al., 2011; Wagner et al., 2001; Whitney et al., 2009).

In patients with schizophrenia, the frontotemporal network was reported as underactive during semantic processing and the authors suggested that these results might be explained by an inhibition impairment (Arcuri et al., 2012). Horn reported a deviant semantic network with increased severity of formal thought disorder in which the difference was due to a decreased contribution of the left pars orbitalis, and the authors highlighted the possible influence of semantic retrieval demands in the context of selection conflict (Horn & Federspiel, 2012). Adopting a similar interpretation, when patients with schizophrenia showed greater activation in the left inferior frontal gyrus (BA45) while performing a semantic task, the authors argued that this greater activation was due to a difficulty in semantic representation access (Chen et al., 2013). Surbeck reported structural aberrations in white matter tracts underlying semantic processing in the ventral language stream of the brain (Surbeck et al., 2020). Taken together, these studies support a role for the inferior frontal gyrus in the disruption of semantic performance in patients with schizophrenia, specifically related to the cognitive control of lexical/semantic selection. Surprisingly, few published studies have examined the aspects of the causal connectivity of brain regions in relation to aberrant lexical selection in schizophrenia.

The term effective connectivity refers to the causal influence that one neural system exerts over other neural systems, while discounting other influences. This directed causal relation is widely different from the structural connectivity that uses white matter tracts (physical architecture), or from functional connectivity that estimates only statistical dependencies (correlation) between brain regions. Dynamic causal modelling (DCM) is a generative model framework for effective connectivity hypothesis testing between neural populations (Zeidman, Jafarian, Corbin, et al., 2019; Zeidman, Jafarian, Seghier, et al., 2019). DCM uses a forward model to estimate the influence of one brain region over another by modelling the neural fluctuation through Bayesian inversion of the brain blood oxygen level response (Friston, 2011; Friston et al., 2003). For

resting-state fMRI instead of estimating the neuronal states time-varying fluctuation, we estimate the parameters of the cross-correlation functions (spectral DCM) by modelling endogenous neural fluctuations.

Notwithstanding the considerable amount of literature that relates semantic memory with the anterior temporal lobe and cognitive control with the inferior frontal gyrus (Binder et al., 2009; Chiou & Lambon Ralph, 2019; Jackson, 2021; Jackson et al., 2016; Rice et al., 2015; Visser et al., 2010), to our knowledge, there have been no attempts to examine the effective connectivity between frontal and temporal regions in patients with schizophrenia and subtle semantic disturbances. Several confounds need to be addressed when undertaking such investigations. First, as long-term exposure to antipsychotic medications can affect connectivity parameters and the inferred causal influences (Schmidt et al., 2016), we recruited a sample of untreated, but first-episode patients. Second, subjective measures of speech are replete with problems of reliability; employing automated measures with corpus-based quantification of statistical co-occurrence of semantic distance is critical for a reliable approach. We employ an NLP measure of lexical similarity and explore the effective connectivity between two core areas of the semantic network: IFG and vATL in untreated first episode of schizophrenia patients (mean 6 days of lifetime daily dose exposure). We hypothesize that lexical inefficiency, indexed by higher semantic similarity between words, will relate to the effective connectivity of IFG and vATL in patients with first episode of schizophrenia.

2. Materials and methods

2.1 Participants

Resting-state fMRI scans were collected from 60 participants at the Prevention and Early Intervention Program for Psychosis in London, Ontario, Canada. The sample included 30 subjects with untreated First Episode of schizophrenia (FES group) and 30 healthy controls (HC group).

68% of this sample has been included in our prior work that focussed on longitudinal speech data (Alonso-Sánchez et al., 2022). Severity of symptoms was confirmed with the Positive and Negative Syndrome Scale (PANSS) (Kay, Fiszbein, et al., 1987). Language disorganization, poverty of speech, and perseveration of ideas were assessed with the Thought and Language Index (TLI) scale. The Impoverishment in Thinking is the sum of poverty of speech, weakening of goal and perseveration of ideas; while disorganization of thinking is the scores sum of looseness, distractibility, peculiar use of words, sentences, and logic. The FEP group was drug-naïve for antipsychotics at the time of assessment (total antipsychotic use less than 14 days). The inclusion criteria for the HC group were not having (a) personal or family mental illness or (b) neurological diseases. All participants provided written informed consent before assessment and ethics approval was granted by the Human Research Ethics Board at Western University, London, Ontario.

2.2 Assessments

All participants were assessed with the Social and Occupational Functioning Assessment Scale (SOFAS) to quantify the functioning level in social and occupational domains (Rybarczyk, 2011) and with the Socioeconomic Status (SES) to measure the parental level of occupation and employment from 1 (Managerial and professional occupations) to 5 (routine occupations)(Boyd, 2008). In addition, we used The Edinburgh Inventory (Oldfield, 1971) to define the handedness of the participants, the scale is scored from -12 (totally left-handed) to +12 (totally right-handed).

2.3 Semantic Measures

To elicit the language sample, the participants were instructed to describe 3 pictures of the Thematic Apperception Test images (Murray, 1943) during one minute per picture and were prompted if they stopped speaking before one-minute time window elapsed. The task was recorded and later transcribed by research assistants. We analyzed the semantic distance between words with the Covington Vector semantic tool (CoVec version 1.0.5912) (Covington, 2016).

CoVec is a natural language processing tool based on data from GloVe Project, with 840 billion words in English with 300-element vectors (Pennington et al., 2014). CoVec measures the likelihood of co-occurrence of words through vector cosine similarity. The text preprocessing removes punctuation, marks 'stopwords' (eg. "a"), and finally, ignores words which are not found in the GloVe dataset (displays a warning of all the missing words). CoVec reports the average of similarity of successive words with an n-word frame segment, using all the positions of the frame. In this analysis we used a full frame including all the text per each picture.

2.4 fMRI acquisition and analysis

The acquisition and processing are described in detail in our prior work reporting baseline data from this sample (Dey et al., 2021; Limongi et al., 2020). In brief, all data were acquired at the Centre for Functional and Metabolic Mapping (CFMM) at the University of Western Ontario on a Siemens 7T Plus (Erlangen, Germany) using a multi-band EPI acquisition sequence (Feinberg et al., 2010; Moeller et al., 2010) with the following imaging parameters: TE/TR = 20 ms / 1000 ms, flip angle = 30◦, 63 slices with multi-band factor = 3, iPat = 3 and an isotropic resolution of 2 mm. A total of 360 functional volumes were collected for a total acquisition time of 6 minutes, 26 seconds. T1-weighted MP2RAGE anatomical volume (TE/TR = 2.83/6000 ms) at 750 μm isotropic resolution was also acquired for anatomical registration. fMRI data was pre-processed with SPM12 (available free at http://www.fil.ion.ucl.ac.uk/spm/ software/spm12/) on Matlab. The pre-processing included manual reorientation on the anterior and posterior commissure, voxel displacement maps calculation for inhomogeneities in the magnetic field, realignment, co-registration and segmentation. A DARTEL template was built with all T1 images of the subjects, and the fMRI images were normalized to the Montreal Neurological Institute (MNI) space. We fitted a general linear model with six head movement parameters and the white matter and cerebrospinal fluid time series and a cosine basis set with frequencies ranging from 0.0078 to 0.1 Hz as regressors. All images were high pass filtered to discard slow frequency drifts (0.0078

Hz). The volumes of interest were defined a priori as Brodmann area 45 (BA45) in the inferior frontal gyrus (MNI: -44 28 22) due to its role in lexical selection (Heim et al., 2008) and the sematic hub located in the ventral anterior temporal lobe (MNI: -40 -15 -33) (Chiou & Lambon Ralph, 2019) (shown in Fig 1).

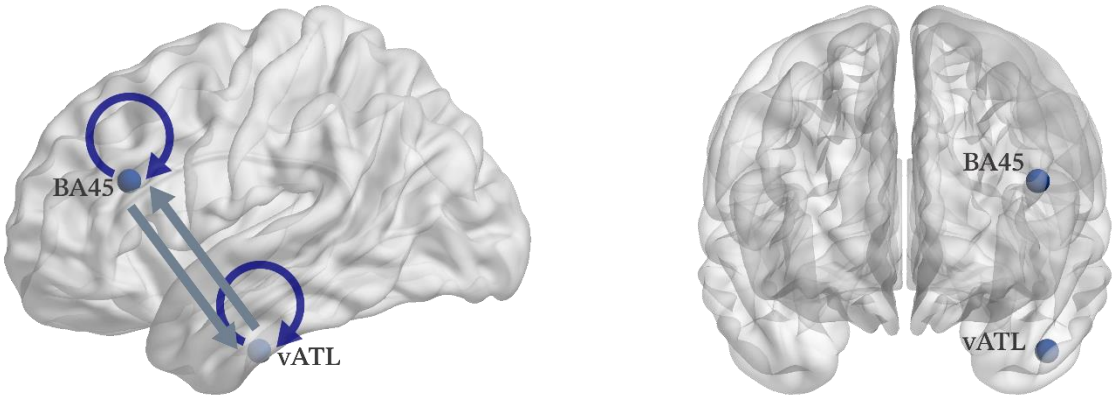

Figure 1. Anatomical map of voxel positioning for dynamic causal modelling. a) Brodmann area 45, b) ventral anterior temporal lobe. Model for DCM. Plot made with BrainNet Viewer (Xia et al., 2013).

2.5 Effective connectivity

To compare the effective connectivity between groups we used a spectral dynamic causal modelling (DCM) approach with Parametrical Empirical Bayes (PEB) (Friston, 2011; Friston et al., 2014) . PEB is a hierarchical Bayesian model that uses both non-linear and linear analyses to measure the group level effects and subject-by-subject variability on parameter estimates. We use bilinear modulatory effects with two states per region, excitatory (inter-region) and inhibitory (within-region), without stochastic effects (Friston et al., 2014).

We evaluate the evidence in support of the hypothesis that both group membership and semantic similarity better explained the effective connectivity within the two-node network than group alone. We also assessed the evidence for an interaction between group and semantic similarity in explaining the 2-node network's effective connectivity. To this end, we compared

the following as regressors predicting connectivity: Model 1: Group; Model 2: Group and Semantic similarity; Model 3: Group, Semantic Similarity, and group × semantic similarity interaction. The semantic regressor was mean-centered and the interaction between variables was summarized as the product of the linguistic measure and group. For the comparison we estimated the free energy (accuracy minus complexity) of the three models considering group and semantic co-occurrence and interaction as covariates. We select the model with higher free energy and analyze the parameter of each volume of interest with the intrinsic and extrinsic connectivity strength and the corresponding posterior probabilities.

2.3 Statistical analysis

Demographic and clinical data were analyzed using descriptive and Bayesian statistics. We use Bayesian t test for continuous variables, and a Bayesian independent multinomial test for categorical variables. Log odd ratios were computed with the 95% credible interval for categorical variables. Bayes factors were considered substantial over 3, strong over 10 and very strong over 30 for all the hypotheses tested according to Jeffreys schema (Jarosz & Wiley, 2014). All the statistical analyses used JASP version 0.14.0.1 (JASP Team, 2020) and the figures plots were made on Python in Jupyter Notebook 6.1.5 (Kluyver et al., 2016).

3. Results

3.1 Demographic

Clinical and demographic data are shown in Table 1. The groups did not differ regarding age, gender or educational level. There was no group difference in the use of English as the first language (83% FES and 90% HC had English as the first language, BF_{10}= 0.281, 95% CI: -0.93, 1.96), in the self-reported immigrant status as first generation (20% FES and 26% HC were first generation immigrants, BF_{10}= 0.317, 95% CI: -1.50, 0.81) and in the ethnicity of the participants (67% FES and 73% HC were white, BF_{10}= 0.334, 95% CI: -1.37, 0.77). All the participants had

English as their transactional language. In the FES group the 50% was antipsychotics naïve, while the other 50% was expose to a mean of 6 days of a lifetime daily dose (SD: 3.2, range: 2-12).

	HC (n= 30) Mean ±SD	FES (n= 30) Mean ±SD	BF_{10}	Effect size δ 95% CI
Age	21.4 ± 3.5	21.4± 2.2	0.262	-0.46, 0.46
Gender	73% male	73% male	0.277	-1.11, 1.13
Educational level (<12/≥12 years)	30% / 70%	46% / 54%	0.713	-1.73, 0.35
Edinburgh Handedness Inventory	10.4 ± 3.1	9.8 ±4.2	0.308	-0.32, 0.60
PANSS-8 Positive	-	12.3 ±3.5	-	-
PANSS-8 Negative	-	7.6 ±4.3	-	-
PANSS-8 total	-	26.4 ±7.4	-	-
SOFAS	81.9 ±5.0	39.0 ±14.7	>10000	
Parental SES (<3/ ≥3)	44% / 56%	30% / 70%	1.00	-0.24, 1.95
Total TLI	0.29 ±0.3	1.46 ±1.3	736.2	-1.64, -0.53
TLI Impoverishment in thinking	0.13 ±0.2	0.47 ±0.6	6.8	-1.16, -0.14
TLI Disorganization in thinking	0.15 ±0.2	0.99 ±1.1	97.6	-1.45, -0.37
Antipsychotic exposure days at time of scan	-	3.07 ±10.9	-	-

Duration of untreated psychosis	-	9.5 ±13.2	-	-

Table 1. Mean and Standard deviations are shown for continuous variables, with percentages for categorical variables. BF_{10}: Bayes Factor. PANSS: Positive and Negative syndrome scale. SOFAS: Social and Occupational Functioning Assessment Scale. TLI – Thought and Language Index.

3.2 Semantic similarity

The groups did not differ in the number of words spoken during the descriptive discourse (HC: 70.5 ±12.8, FES: 72.7 ±33.9, BF= 0.266, effect size δ 95% CI = -0.50, 0.42), but as hypothesized, the FES group had higher semantic similarity than the HC group with very strong evidence favoring the between-group difference hypothesis (HC: 0.330 ±0.02, FES: 0.364 ±0.03, BF= 1623, effect size δ 95% CI = -1.71, -0.59) (Figure 2).

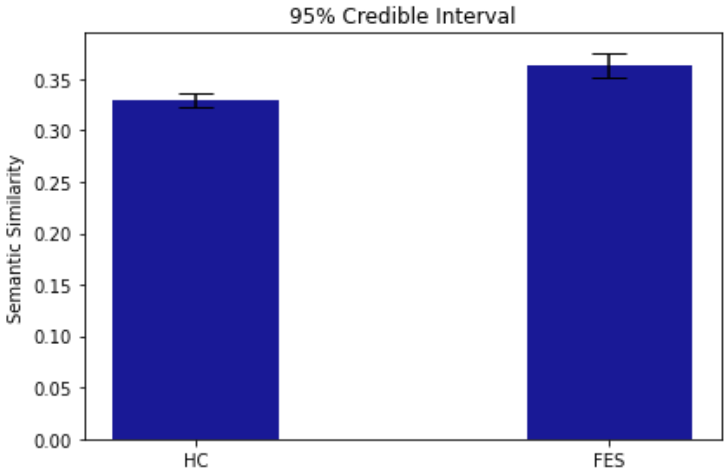

Figure 2. 95% highest credible interval of semantic similarity per group. HC: Healthy controls; FES: first-episode schizophrenia.

We also examined whether the higher semantic similarity in FEP comes from generating content that is non-specific to the picture (i.e., specificity) and/or from content that is 'stuck' in describing limited features of the picture at hand (i.e. insufficient diversity). These results, presented in the Supplementary Materials, indicate that when describing the pictures offered,

patients generate content that is likely to be as specific and as diverse as healthy subjects, but still choose words that are semantically close to each other.

3.3 Model selection

When comparing the models regarding the regressors included on them, the free energy of the second model (Group and semantic similarity) exceeded the other models' performance (Figure 3). The third model (Group, semantic similarity, and interaction) was an overfit for the data.

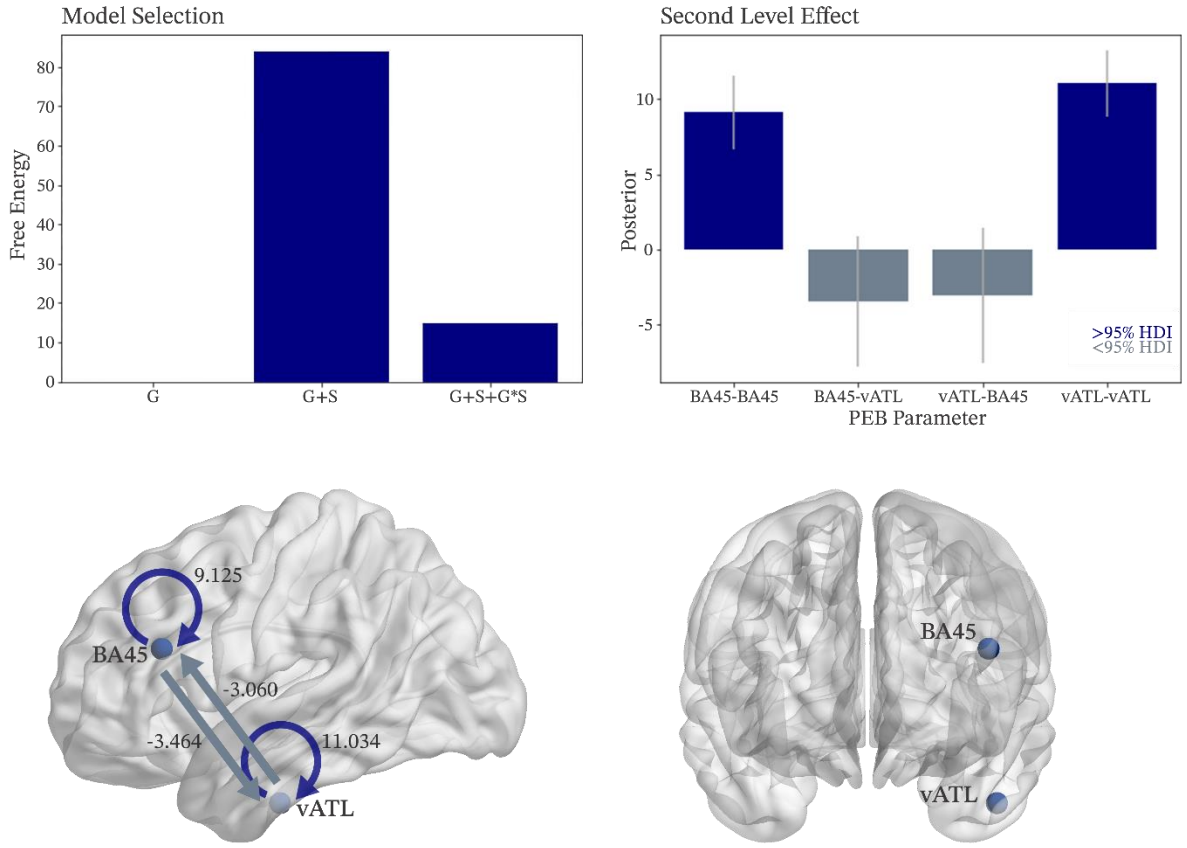

Figure 3: Top left: Model selection. G: Model group; G+S: Model group and semantic similarity; G+S+G*S: Model group, semantic similarity and Interaction between group and semantic similarity. Top right: Second level effect of Semantic similarity on intrinsic inhibitory and extrinsic excitatory connections. Posterior parameter estimates with bars represent the estimated effect size and the lines represent 90% credible interval. Bottom: Brain network with the connectivity strength.

3.4 Effective connectivity

The parameter estimates from the winning model indicated that the intrinsic inhibitory connections of BA45 and vATL increased as a function of semantic similarity in the FES group (FES coded as 1; HC coded as -1). The parameters (Figure 3) indicates that the intrinsic inhibitory connection of BA45 and vATL increased linearly with semantic similarity. The mean estimated connection strength within BA45 [intrinsic] is 9.125 (posterior probability, PP = 1), from BA45 to vATL -3.464 (PP: 0.90), from vATL to BA45 -3.060 (PP: 0.87) and, within vATL [intrinsic] is 11.034 (PP: 1.0).

4. Discussion

In this spectral DCM study leveraging an NLP measure of semantic similarity, we demonstrate 3 major findings. First, patients with FES chose words with higher degree of semantic similarity (i.e., lower semantic distance from each other), despite speaking comparable number of words when describing pictures. Second, at resting state, we recovered a fully connected network between the temporal semantic hub (vATL) and inferior frontal Broca's area (BA45); the connectivity of this network was better explained by considering inter-individual differences in semantic similarity along with diagnosis, compared to diagnostic status alone. Third, we show that in FES, higher intrinsic inhibitory connectivity within vATL and BA45 at rest relates to an aberrant increase in semantic similarity of word choices when participants describe a picture. These findings provide a neural model linking speech production and semantic storage regions in the lexical impairment seen in first episode schizophrenia.

Higher inhibitory connections reflect a loss of synaptic gain within cortical microcircuits. In an excitation/inhibition framework, an increase in inhibitory tone may be seen as a reflection of a primary pyramidal dysfunction in schizophrenia (as proposed by Adams et al., 2022) or an increase interneuron-related GABAergic activity (Pinotsis & Friston, 2012). Based on the findings from N-methyl-d-aspartate receptor antagonist (ketamine), autoantibody models,

postmortem studies and knockout mouse models, Nakazawa, Jeevakumar and Nakao (Jeevakumar & Kroener, 2016; Nakazawa et al., 2017) proposed that NMDA hypofunction initially occurs in cortical GABAergic neurons during early postnatal development, reducing intrinsic excitability and GABA release, leading to a disinhibition of pyramidal neurons and subsequent homeostatic down regulation of pyramidal excitability.

Reduced synaptic gain (reflected in increased inhibitory tone), may relate to reduced precision of locally generated neural activity, as well as a reduction in the generation of spontaneous neural oscillations at rest. One physiological consequence of higher intrinsic inhibitory tone is a reduction in the neuronal disinhibition or 'release' required for excitatory activity. In terms of neuronal recruitment in service of a cognitive task, this may make a brain region more dependent on extrinsically driven stimulatory processes. If this increased regional inhibitory tone continues during speech production, for our 2-node language network model, this may mean that aberrant word sampling cannot be overcome in the presence of a physiologically normal level of top-down stimulation (from IFG) arriving at the vATL.

The strength of the connection between the vATL and IFG was not affected by diagnostic status or semantic similarity in our sample. It is possible that such an effect on extrinsic connections may be present when speaking, but not at the resting state when fMRI scans were obtained. In fact, dynamic causal influence between left IFG and temporal cortex has been linked to the 'priming' activity in a previous study (Matsumoto & Kakigi, 2013). Experimental verification of this possibility requires a speech production task in the scanner wherein demands placed on the control system from IFG to vATL (thus, extrinsic signals to and from the semantic hub), can be studied directly (Jackson, 2021).

Our results complement previous observations of aberrant Broca's area connectivity in first-episode schizophrenia (Du et al., 2021; Li et al., 2017). Interestingly, while previous studies have focused on the IFG, few studies have examined the association between the anterior temporal

lobe and schizophrenia features of speech. As NLP measures become more commonly available, they have the potential to replace subjective rating scales as instruments of psychopathological quantification. In this context, eliciting the neural basis of key NLP measures in schizophrenia is of critical importance. While a plethora of derived measures are available, we chose a lexical measure based on world choices during a discourse, given the wealth of information available on semantic abnormalities in schizophrenia (Pomarol-Clotet et al., 2008), This approach provides a precedence for examining the neural basis of syntactic abnormalities (e.g. connective use (Mackinley et al., 2021)) and linguistic style (e.g. analytic thinking (Silva et al., 2021)) that are shown to be relevant in schizophrenia. However, such studies do not address the inhibitory or excitatory connectivity of the vATL so it is challenging to compare our results with previous reports.

Our study has several strengths as well as some limitations. To our knowledge, this is the first study exploring the effective connectivity of IFG and vATL in patients with first episode of schizophrenia in relation to lexical efficiency. We recruited an extremely early-stage sample who had a mean of 3 days of lifetime antipsychotic exposure and were still in the untreated state when speech and fMRI data were obtained. We also used unambiguous picture stimuli which are shown to be more reliable in NLP studies of schizophrenia (Morgan et al., 2021). Given the head motion challenges when people are speaking in the scanner, we used resting state fMRI data to make inferences about speech outside the scanner. This has conceptual limitations and may explain why Broca-vATL connectivity was not related to the semantic similarity measure as expected. The relationship between task and resting-state effective connectivity (as inferred using DCM) is less well understood. Jung and colleagues demonstrated a strong correlation in the effective connectivity among regions engaged in a n-back task between the rest and the task states and concluded that the task-dependent connectivity endows resting- state connectivity with a context sensitivity that predicts task performance (Jung et al., 2018). Applying these results to our findings, we can expect that in the context of online speech production, the BA45-

vATL connectivity pattern may exhibit 'context sensitive' changes in line with their resting-state connectivity patterns. More specifically, we can expect the intrinsic inhibitory tone of BA45 and vATL to be modulated by speech production when in-scanner speech is generated. We caution the readers that this is a speculative interpretation; a conclusive answer requires empirical data to test within-subject rest-task concordance. Finally, we had access to limited longitudinal imaging data from this sample, precluding an examination of changes in brain connectivity over time to complement our recent report of worsening semantic similarity in FES (Alonso-Sánchez et al., 2022).

5. Conclusion

In conclusion, we report an aberrant inhibitory connection strength in the IFG and the vATL in relation to inefficiency of semantic selection in schizophrenia. Our results suggest that NLP-derived semantic anomalies in schizophrenia have a likely neurophysiological basis, and lexical selection may be linked to reduced regional synaptic gain within the left IFG and the vATL.

Acknowledgements

We appreciate all the participants and their families for the time and effort to contribute to this study. We thank Michael Covington (Covington Innovations) for providing us his CoVec NLP tool. We thank all research team members of the NIMI lab and all the staff members of the PEPP London team, particularly Drs. Kara Dempster (currently at Dalhousie University), Julie Richard, Priya Subramanian and Hooman Ganjavi and (late) Dr. Raj Harricharan for their assistance in patient recruitment and supporting clinical care. We also appreciate the assistance offered by Peter Jeon and Michael Mackinley for data acquisition, Dr. Ali Khan for the fMRI protocol, Hannah Ke, Jenny Chan and Sabrina Ford for data curation.

Competing interests

LP reports personal fees from Otsuka Canada, SPMM Course Limited, UK, Canadian Psychiatric Association; book royalties from Oxford University Press; investigator-initiated educational grants from Janssen Canada, Sunovion and Otsuka Canada outside the submitted work. All other authors report no relevant conflicts.

Funding

This study was funded by The Canadian Institutes of Health Research (CIHR) Foundation Grant (375104/2017) and the Canada First Research Excellence Fund (CFREF) to BrainsCAN. This work was also supported by the National Agency for Research and Development (ANID), Scholarship Program, Becas Chile 2019, Postdoctoral Fellow 74200048 (MFA); We also acknowledge support from the Bucke Family Fund, The Chrysalis Foundation and The Arcangelo Rea Family Foundation (London, Ontario). LP acknowledges personal chair support from the Tanna Schulich Endowment (Schulich School of Medicine, Western University) and Monique H. Bourgeois Endowment (Douglas Research Centre, McGill University).

References

Adams, R. A., Pinotsis, D., Tsirlis, K., Unruh, L., Mahajan, A., Horas, A. M., Convertino, L., Summerfelt, A., Sampath, H., Du, X. M., Kochunov, P., Ji, J. L., Repovs, G., Murray, J. D., Friston, K. J., Hong, L. E., & Anticevic, A. (2022). Computational Modeling of Electroencephalography and Functional Magnetic Resonance Imaging Paradigms Indicates a Consistent Loss of Pyramidal Cell Synaptic Gain in Schizophrenia. *Biological Psychiatry*, *91*(2), 202–215. https://doi.org/10.1016/j.biopsych.2021.07.024

Almeida, V. N., & Radanovic, M. (2021). Semantic priming and neurobiology in schizophrenia: A theoretical review. *Neuropsychologia*, 108058. https://doi.org/10.1016/j.neuropsychologia.2021.108058

Alonso-Sánchez, M. F., Ford, S. D., MacKinley, M., Silva, A. M., Limongi, R., & Palaniyappan, L. (2022). Progressive changes in descriptive discourse in First Episode Schizophrenia: a longitudinal computational semantics study. *Schizophrenia*, *8*(1), 1–9. https://doi.org/10.1038/s41537-022-00246-8

Arcuri, S. M., Broome, M. R., Giampietro, V., Jr, E. A., Kircher, T. T. J., Williams, S. C. R., Andrew, C. M., Brammer, M., Morris, R. G., & Mcguire, P. K. (2012). *Faulty Suppression of Irrelevant Material in Patients with Thought Disorder Linked to Attenuated Frontotemporal Activation*. *2012*. https://doi.org/10.1155/2012/176290

Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex*, *19*(12), 2767–2796. https://doi.org/10.1093/cercor/bhp055

Boyd, M. (2008). A socioeconomic scale for Canada: Measuring occupational status from the census. *Canadian Review of Sociology*, *45*(1), 51–91. https://doi.org/10.1111/j.1755-618X.2008.00003.x

Brown, M., & Kuperberg, G. R. (2015). *Framework of Language Processing : Linking Language Perception , Interpretation , and Production Abnormalities in Schizophrenia*. *9*(November), 1–23. https://doi.org/10.3389/fnhum.2015.00643

Chen, P. J., Fan, L. Y., Hwang, T. J., Hwu, H. G., Liu, C. M., & Chou, T. L. (2013). The deficits on a cortical-subcortical loop of meaning processing in schizophrenia. *NeuroReport*, *24*(3), 147–151. https://doi.org/10.1097/WNR.0b013e32835df562

Chiou, R., Humphreys, G. F., Jung, J. Y., & Lambon Ralph, M. A. (2018). Controlled semantic cognition relies upon dynamic and flexible interactions between the executive 'semantic control' and hub-and-spoke 'semantic representation' systems. *Cortex*, *103*, 100–116. https://doi.org/10.1016/j.cortex.2018.02.018

Chiou, R., & Lambon Ralph, M. A. (2019). Unveiling the dynamic interplay between the hub- and spoke-components of the brain's semantic system and its impact on human behaviour. *NeuroImage*, *199*(April), 114–126. https://doi.org/10.1016/j.neuroimage.2019.05.059

Conner, C. R., Kadipasaoglu, C. M., Shouval, H. Z., Hickok, G., & Tandon, N. (2019). Network dynamics of Broca's area during word selection. *PLoS ONE*, *14*(12), 1–30. https://doi.org/10.1371/journal.pone.0225756

Covington, M. A. (2016). *Covington Vector Semantics Tools*. Covington Innovations.

Dey, A., Dempster, K., MacKinley, M., Jeon, P., Das, T., Khan, A., Gati, J., & Palaniyappan, L. (2021). Conceptual disorganization and redistribution of resting-state cortical hubs in untreated first-episode psychosis: A 7T study. *Npj Schizophrenia*, *7*(1), 1–9. https://doi.org/10.1038/s41537-020-00130-3

Du, J., Palaniyappan, L., Liu, Z., Cheng, W., Gong, W., Zhu, M., Wang, J., Zhang, J., & Feng, J. (2021). The genetic determinants of language network dysconnectivity in drug-naïve early stage schizophrenia. *Npj Schizophrenia*, *7*(1). https://doi.org/10.1038/s41537-021-00141-8

Elvevåg, B., Foltz, P. W., Weinberger, D. R., & Goldberg, T. E. (2007). Quantifying incoherence in speech: An automated methodology and novel application to schizophrenia. *Schizophrenia Research*, *93*(1–3), 304–316. https://doi.org/10.1016/j.schres.2007.03.001

Feinberg, D. A., Moeller, S., Smith, S. M., Auerbach, E., Ramanna, S., Glasser, M. F., Miller, K. L., Ugurbil, K., & Yacoub, E. (2010). Multiplexed echo planar imaging for sub-second whole brain fmri and fast diffusion imaging. *PLoS ONE*, *5*(12). https://doi.org/10.1371/journal.pone.0015710

Friston, K. J. (2011). Functional and Effective Connectivity: A Review. *Brain Connectivity*, *1*(1), 13–36. https://doi.org/10.1089/brain.2011.0008

Friston, K. J., Harrison, L., & Penny, W. (2003). Dynamic causal modelling. *NeuroImage*, *19*(4), 1273–1302. https://doi.org/10.1016/S1053-8119(03)00202-7

Friston, K. J., Kahan, J., Biswal, B., & Razi, A. (2014). NeuroImage A DCM for resting state fMRI. *NeuroImage*, *94*, 396–407. https://doi.org/10.1016/j.neuroimage.2013.12.009

Heim, S., Eickhoff, S. B., & Amunts, K. (2008). Specialisation in Broca's region for semantic, phonological, and syntactic fluency? *NeuroImage*, *40*(3), 1362–1368. https://doi.org/10.1016/j.neuroimage.2008.01.009

Hoffman, P., Cogdell-Brooke, L., & Thompson, H. E. (2020). Going off the rails: Impaired coherence in the speech of patients with semantic control deficits. *Neuropsychologia*, *146*(June), 107516. https://doi.org/10.1016/j.neuropsychologia.2020.107516

Horn, H., & Federspiel, A. (2012). *Semantic Network Disconnection in Formal Thought Disorder*. 14–23. https://doi.org/10.1159/000337133

Jackson, R. L. (2021). The neural correlates of semantic control revisited. *NeuroImage*, *224*(October 2020), 117444. https://doi.org/10.1016/j.neuroimage.2020.117444

Jackson, R. L., Cloutman, L. L., & Lambon, M. A. (2019). ScienceDirect Exploring distinct default mode and semantic networks using a systematic ICA approach. *CORTEX*, *113*, 279–297. https://doi.org/10.1016/j.cortex.2018.12.019

Jackson, R. L., Hoffman, P., Pobric, X. G., & Ralph, X. M. A. L. (2016). *The Semantic Network at Work and Rest : Differential Connectivity of Anterior Temporal Lobe Subregions*. *36*(5), 1490–1501. https://doi.org/10.1523/JNEUROSCI.2999-15.2016

Jarosz, A., & Wiley, J. (2014). What Are the Odds? A Practical Guide to Computing and Reporting Bayes Factors. *Journal of Problem Solving*, *7*, 2–9. https://doi.org/10.7771/1932-6246.1167

JASP Team. (2020). *JASP* (Version 0.14.1). University of Amsterdam. https://jasp-stats.org/

Jeevakumar, V., & Kroener, S. (2016). Ketamine Administration during the Second Postnatal Week Alters Synaptic Properties of Fast-Spiking Interneurons in the Medial Prefrontal Cortex of Adult Mice. *Cerebral Cortex*, *26*(3), 1117–1129. https://doi.org/10.1093/cercor/bhu293

Jung, K., Friston, K. J., Pae, C., Choi, H. H., Tak, S., Choi, Y. K., Park, B., Park, C. A., Cheong, C., & Park, H. J. (2018). Effective connectivity during working memory and resting states: A DCM study. *NeuroImage*, *169*(December 2017), 485–495. https://doi.org/10.1016/j.neuroimage.2017.12.067

Kanerva, P. (2009). Hyperdimensional computing: An introduction to computing in distributed representation with high-dimensional random vectors. *Cognitive Computation*, *1*(2), 139–159. https://doi.org/10.1007/s12559-009-9009-8

Kay, S. R., Fiszbein, A., & Opler, L. A. (1987). The Positive and Negative Syndrome Scale (PANSS) for Schizophrenia. *Schizophrenia Bulletin*, *13*(2), 261–276. https://doi.org/10.1093/schbul/13.2.261.

Kluyver, T., Ragan-Kelley, B., Pérez, F., Granger, B., Bussonnier, M., Frederic, J., Kelley, K., Hamrick, J., Grout, J., Corlay, S., Ivanov, P., Avila, D., Abdalla, S., Willing, C., & Team, J. development. (2016). Jupyter Notebooks – a publishing format for reproducible computational workflows. In F. Loizides

& B. Scmidt (Eds.), *In Positioning and Power in Academic Publishing: Players, Agents and Agendas.* (pp. 87–90). IOS Press. https://doi.org/doi:10.3233/978-1-61499-649-1-87

Krieger-Redwood, K., & Jefferies, E. (2014). TMS interferes with lexical-semantic retrieval in left inferior frontal gyrus and posterior middle temporal gyrus: Evidence from cyclical picture naming. *Neuropsychologia*, *64*, 24–32. https://doi.org/10.1016/j.neuropsychologia.2014.09.014

Levelt, W. J. M. (2001). Spoken word production: A theory of lexical access. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(23), 13464–13471. https://doi.org/10.1073/pnas.231459498

Levy, O., Goldberg, Y., & Dagan, I. (2015). Improving Distributional Similarity with Lessons Learned from Word Embeddings. *Transactions of the Association for Computational Linguistics*, *3*, 211–225. https://doi.org/10.1162/tacl_a_00134

Li, Y., Li, P., Yang, Q. X., Eslinger, P. J., Sica, C. T., & Karunanayaka, P. (2017). Lexical-semantic search under different covert verbal fluency tasks: An fMRI study. *Frontiers in Behavioral Neuroscience*, *11*(August), 1–15. https://doi.org/10.3389/fnbeh.2017.00131

Limongi, R., Jeon, P., Mackinley, M., Das, T., Dempster, K., Théberge, J., Bartha, R., Wong, D., & Palaniyappan, L. (2020). Glutamate and Dysconnection in the Salience Network: Neurochemical, Effective Connectivity, and Computational Evidence in Schizophrenia. *Biological Psychiatry*, *88*(3), 273–281. https://doi.org/10.1016/j.biopsych.2020.01.021

Mackinley, M., Chan, J., Ke, H., Dempster, K., & Palaniyappan, L. (2021). Linguistic determinants of formal thought disorder in first episode psychosis. *Early Intervention in Psychiatry*, *15*(2), 344–351. https://doi.org/10.1111/eip.12948

Matsumoto, A., & Kakigi, R. (2013). Subliminal Semantic Priming Changes the Dynamic Causal Influence between the Left Frontal and Temporal Cortex. *Journal of Cognitive Neuroscience*, *26*, 165–174. https://doi.org/doi:10.1162/jocn_a_00472

Mikolov, T., Sutskever, I., Chen, K., Corrado, G., & Dean, J. (2013). Distributed Representations ofWords and Phrases and their Compositionality. *Advances in Neural Information Processing Systems*, 1389–1399. https://doi.org/10.18653/v1/d16-1146

Moeller, S., Yacoub, E., Olman, C. A., Auerbach, E., Strupp, J., Harel, N., & Uğurbil, K. (2010). Multiband Multislice GE-EPI at 7 Tesla, With 16-Fold Acceleration Using Partial Parallel Imaging With Application to High Spatial and Temporal Whole-Brain FMRI. *Magn Reson Med*, *63*(5), 1144–1153. https://doi.org/doi:10.1002/mrm.22361

Morgan, S. E., Diederen, K., Vértes, P. E., Ip, S. H. Y., Wang, B., Thompson, B., Demjaha, A., & Micheli, A. de. (2021). Assessing psychosis risk using quantitative markers of disorganised speech. *Npj Schizophrenia*, *December 2020*. https://doi.org/10.1101/2021.01.04.20248717

Mota, N. B., Copelli, M., & Ribeiro, S. (2017). Thought disorder measured as random speech structure classifies negative symptoms and schizophrenia diagnosis 6 months in advance. *Npj Schizophrenia*, *3*(1), 1–10. https://doi.org/10.1038/s41537-017-0019-3

Murray, H. (1943). *Thematic Apperception Test*. Harvard University Press.

Nakazawa, K., Jeevakumar, V., & Nakao, K. (2017). Spatial and temporal boundaries of NMDA receptor hypofunction leading to schizophrenia. *Npj Schizophrenia*, *3*(1), 0–1. https://doi.org/10.1038/s41537-016-0003-3

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*(1), 97–113. https://doi.org/10.1016/0028-3932(71)90067-4

Palaniyappan, L., Mota, N. B., Oowise, S., Balain, V., Copelli, M., Ribeiro, S., & Liddle, P. F. (2019). Speech structure links the neural and socio-behavioural correlates of psychotic disorders. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *88*(April 2018), 112–120. https://doi.org/10.1016/j.pnpbp.2018.07.007

Pennington, J., Socher, R., & Manning, C. D. (2014). GloVe: Global Vectors for Word Representation. *Conference: Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 1532–1543. https://doi.org/10.3115/v1/D14-1162

Pomarol-Clotet, E., Oh, T. M. S. S., Laws, K. R., & McKenna, P. J. (2008). Semantic priming in schizophrenia: Systematic review and meta-analysis. *British Journal of Psychiatry*, *192*(2), 92–97. https://doi.org/10.1192/bjp.bp.106.032102

Python, G., Glize, B., & Laganaro, M. (2018). The involvement of left inferior frontal and middle temporal cortices in word production unveiled by greater facilitation effects following brain damage. *Neuropsychologia*, *121*(October), 122–134. https://doi.org/10.1016/j.neuropsychologia.2018.10.026

Rezaii, N., Walker, E., & Wolff, P. (2019). OPEN A machine learning approach to predicting psychosis using semantic density and latent content analysis. *Npj Schizophrenia*, *December 2018*. https://doi.org/10.1038/s41537-019-0077-9

Rice, G. E., Hoffman, P., & Lambon Ralph, M. A. (2015). Graded specialization within and between the anterior temporal lobes. *Annals of the New York Academy of Sciences*, *1359*(1), 84–97. https://doi.org/10.1111/nyas.12951

Rybarczyk, B. (2011). Encyclopedia of Clinical Neuropsychology. In J. Kreutzer, J. DeLuca, & B. Caplan (Eds.), *Encyclopedia of Clinical Neuropsychology*. Springer. https://doi.org/10.1007/978-0-387-79948-3

Schmidt, A., Palaniyappan, L., Smieskova, R., Simon, A., Riecher-Rössler, A., Lang, U. E., Fusar-Poli, P., McGuire, P., & Borgwardt, S. J. (2016). Dysfunctional insular connectivity during reward prediction in patients with first-episode psychosis. *Journal of Psychiatry and Neuroscience*, *41*(6), 367–376. https://doi.org/10.1503/jpn.150234

Silva, A., Limongi, R., MacKinley, M., & Palaniyappan, L. (2021). Small Words That Matter: Linguistic Style and Conceptual Disorganization in Untreated First-Episode Schizophrenia. *Schizophrenia Bulletin Open*, *2*(1), 1–10. https://doi.org/10.1093/schizbullopen/sgab010

Snyder, H. R., Banich, M. T., & Munakata, Y. (2011). Choosing our words: Retrieval and selection processes recruit shared neural substrates in left ventrolateral prefrontal cortex. *Journal of Cognitive NeuroscienceNeurosci.*, *23*(11), 3470–3482. https://doi.org/10.1162/jocn_a_00023

Spencer, T. J., Thompson, B., Oliver, D., Diederen, K., Demjaha, A., Weinstein, S., Morgan, S. E., Day, F., Valmaggia, L., Rutigliano, G., de Micheli, A., Mota, N. B., Fusar-Poli, P., & McGuire, P. (2021). Lower speech connectedness linked to incidence of psychosis in people at clinical high risk. *Schizophrenia Research*, *228*, 493–501. https://doi.org/10.1016/j.schres.2020.09.002

Surbeck, W., Hänggi, J., Scholtes, F., Viher, P. v., Schmidt, A., Stegmayer, K., Studerus, E., Lang, U. E., Riecher-Rössler, A., Strik, W., Seifritz, E., Borgwardt, S., Quednow, B. B., & Walther, S. (2020). Anatomical integrity within the inferior fronto-occipital fasciculus and semantic processing deficits in schizophrenia spectrum disorders. *Schizophrenia Research*, *218*, 267–275. https://doi.org/10.1016/j.schres.2019.12.025

Visser, M., Jefferies, E., & Lambon Ralph, M. A. (2010). Semantic processing in the anterior temporal lobes: A meta-analysis of the functional neuroimaging literature. *Journal of Cognitive Neuroscience*, *22*(6), 1083–1094. https://doi.org/10.1162/jocn.2009.21309

Wagner, A., Paré-Blagoev, J., Clark, J., & Poldrack, R. (2001). Recovering Meaning: Left Prefrontal Cortex Guides Controlled Semantic Retrieval. *Neuron*, *31*, 329–338. https://doi.org/10.1016/S0026-0495(96)90074-8

Whitney, C., Grossman, M., & Kircher, T. T. J. (2009). The influence of multiple primes on bottom-up and top-down regulation during meaning retrieval: Evidence for 2 distinct neural networks. *Cerebral Cortex*, *19*(11), 2548–2560. https://doi.org/10.1093/cercor/bhp007

Xia, M., Wang, J., & He, Y. (2013). BrainNet Viewer: A Network Visualization Tool for Human Brain Connectomics. *PLoS ONE*, *8*(7). https://doi.org/10.1371/journal.pone.0068910

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*(May), 174–190. https://doi.org/10.1016/j.neuroimage.2019.06.031

Zeidman, P., Jafarian, A., Seghier, M. L., Litvak, V., Cagnan, H., Price, C. J., & Friston, K. J. (2019). A guide to group effective connectivity analysis, part 2: Second level analysis with PEB. *NeuroImage*, *200*(May), 12–25. https://doi.org/10.1016/j.neuroimage.2019.06.032