



HHS Public Access

Author manuscript

Psychiatry Res. Author manuscript; available in PMC 2023 March 01.

Published in final edited form as:

Psychiatry Res. 2022 March ; 309: 114404. doi:10.1016/j.psychres.2022.114404.

Semantic and phonetic similarity of verbal fluency responses in early-stage psychosis

Nancy B. Lundin^a, Michael N. Jones^b, Evan J. Myers^c, Alan Breier^d, Kyle S. Minor^e

^aDepartment of Psychological and Brain Sciences and Program in Neuroscience, Indiana University, Bloomington, IN, USA; Department of Psychiatry and Behavioral Health, The Ohio State University, Columbus, OH, USA

^bDepartment of Psychological and Brain Sciences and Cognitive Science Program, Indiana University, Bloomington, IN, USA

^cDepartment of Psychology, Indiana University Purdue University Indianapolis, Indianapolis, IN, USA

^dDepartment of Psychiatry, Indiana University School of Medicine, Indianapolis, IN, USA; Eskenazi Midtown Prevention and Recovery Center for Early Psychosis, Indianapolis, IN, USA

^eDepartment of Psychology, Indiana University Purdue University Indianapolis, Indianapolis, IN, USA; Eskenazi Midtown Prevention and Recovery Center for Early Psychosis, Indianapolis, IN, USA

Abstract

Linguistic abnormalities can emerge early in the course of psychotic illness. Computational tools that quantify response similarity in standardized tasks such as the verbal fluency test could efficiently characterize the nature and functional correlates of these deficits. Participants with early-stage psychosis ($n=20$) and demographically matched controls without a psychiatric diagnosis ($n=20$) performed category and letter verbal fluency. Semantic similarity was measured via predicted context co-occurrence in a large text corpus using Word2Vec. Phonetic similarity was measured via edit distance using the VFClust tool. Responses were designated as clusters (related items) or switches (transitions to less related items) using similarity-based thresholds. Results revealed that participants with early-stage psychosis compared to controls had lower fluency scores, lower cluster-related semantic similarity, and fewer switches; mean cluster size and

Corresponding author: Kyle S. Minor, Department of Psychology, Indiana University Purdue University Indianapolis Indianapolis School of Medicine, LD 124, 402 N. Blackford Street, Indianapolis, IN 46202, USA Phone: (317) 274-2933 | ksminor@iupui.edu.
Author contributors

Nancy Lundin: Conceptualization, Data Curation, Formal analysis, Project administration, Software, Visualization, Writing - original draft. **Michael Jones:** Data Curation, Resources, Software, Supervision, Validation, Writing - review & editing. **Evan Myers:** Writing - original draft, Writing - review & editing. **Alan Breier:** Resources, Supervision, Writing - review & editing. **Kyle Minor:** Conceptualization, Data curation, Funding acquisition, Investigation, Project administration, Resources, Supervision, Writing - review & editing.

Conflict of interest

The authors have declared that there are no conflicts of interest in relation to the subject of this study.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

phonetic similarity did not differ by group. Lower fluency semantic similarity was correlated with greater speech disorganization (Communication Disturbances Index), although more strongly in controls, and correlated with poorer social functioning (Global Functioning: Social), primarily in the psychosis group. Findings suggest that search for semantically related words may be impaired soon after psychosis onset. Future work is warranted to investigate the impact of language disturbances on social functioning over the course of psychotic illness.

Keywords

early psychosis; computational linguistics; phonetic similarity; semantic coherence

1. Introduction

The ability to think in a coherent, organized manner is critical for making decisions and communicating one's needs to others. This fluidity of thought is often impaired in individuals experiencing psychosis, evidenced by disrupted speech patterns often referred to as formal thought disorder (Bleuler, 1911; Docherty, 2012). While thought disorder has historically been characterized using manual rating systems (e.g., Andreasen, 1986; Solovay et al., 1986), computational linguistic tools have the capacity to harness fine-grained semantic information from speech with greater efficiency and reliability (Corcoran et al., 2020). Recent studies have demonstrated that automated tools can provide unique information compared to clinical rating scales of disorganization (Minor et al., 2019) and detect subtle speech disturbances which predict conversion to psychosis in individuals at high clinical risk (Bedi et al., 2015; Corcoran et al., 2018). However, these findings do not always replicate (Iter et al., 2018), and there is a great deal of variability in the analytic methods and speech prompts used (Hitzenko et al., 2021). More research is needed to characterize the nature of linguistic alterations early in the course of psychotic illness, particularly using automated tools and standardized laboratory tasks that are easy to administer such as the verbal fluency test (Holmlund et al., 2019).

The verbal fluency test is an efficient and widely administered task in clinical and research contexts in which participants are asked to name as many different items belonging to a particular category (e.g., animals, foods) or beginning with a particular letter (e.g., S) as possible within a fixed timeframe (e.g., 1 minute). Participants tend to produce semantically or phonetically related groupings of words in bursts over time, described as "clustering" and "switching" to new clusters (Bousfield and Sedgewick, 1944; Gruenewald and Lockhead, 1980; Troyer et al., 1997). Letter fluency is typically more difficult for participants than category fluency, perhaps because category fluency resembles more common mental activities such as generating items for a grocery list (Shao et al., 2014). However, individuals with schizophrenia tend to exhibit greater impairments in category than letter fluency (Bokat and Goldberg, 2003), and this heightened semantic deficit is already present in first-episode psychosis (Mesholam-Gately et al., 2009). Further research is needed to understand the specific processes that underlie semantic functioning deficits at this early stage of illness.

Studies have examined whether clustering and switching processes in verbal fluency are altered in schizophrenia compared to healthy controls using hand-coded categorization schemes (e.g., Troyer et al., 1997). These schemes provide a framework for grouping responses in particular types of category fluency data (e.g., animals) into common subcategories (e.g., pets, farm animals) and provide a set of rules for scoring phonetic clusters in letter fluency data, such as grouping words that start with the same sound. Many studies using these types of hand-coded methods have found fewer switches and similar cluster sizes in schizophrenia compared to controls (Elvevåg et al., 2002; Moore et al., 2006; Piras et al., 2019), while some have found fewer switches and cluster-related words (Bozikas et al., 2005; Robert et al., 1998), and still others have reported smaller cluster sizes with no difference in the number of switches (van Beilen et al., 2004). Given these mixed findings and the variety of ways that clustering has been measured, automated metrics of semantic and phonetic similarity could be leveraged to characterize linguistic search processes more efficiently in psychosis.

Advances in natural language processing have facilitated opportunities to quantify specific response patterns in verbal fluency performance to study psychopathology (Holmlund et al., 2019). Automated methods can also be adapted to different cultures and languages more feasibly than manual coding methods (Kim et al., 2019). Tools such as VFClust (Ryan et al., 2013) can analyze the phonetic similarity between words, although this tool has rarely been applied in psychosis research. Semantic similarity between words can be measured efficiently and reliably with computational semantic models such as Latent Semantic Analysis (LSA; Landauer and Dumais, 1997) and Word2Vec (Mikolov et al., 2013). Models of this class provide indices of semantic coherence by training a learning algorithm on the co-occurrence of words within contexts across large text corpora. Studies have found reduced semantic coherence of category fluency responses in schizophrenia, particularly in individuals with disorganized speech, compared to non-psychiatric controls (Elvevåg et al., 2007; Pauselli et al., 2018). While the measure of overall semantic coherence is informative, work examining the distinct “local” (i.e., finding related words) and “global” (i.e., searching for a new set of words) phases of verbal fluency search (Hills et al., 2012; Lundin et al., 2020; Troyer et al., 1997) raises the question of whether individuals with psychosis produce less semantically associated responses within clusters than controls, switch to new clusters that are more semantically distant from prior clusters than controls, or both.

Finally, research has yet to determine whether semantic search behavior during verbal fluency indexes speech organization more broadly in individuals with psychotic illness. Does one’s ability to retrieve semantically related words during a brief search task relate to the production of orderly discourse, or are these non-overlapping processes? Moreover, does semantic coherence in verbal fluency responses relate to better social functioning with family members and peers? Evidence suggests that individuals with schizophrenia and first-episode psychosis who exhibit more disorganized speech tend to experience modestly greater difficulties with social functioning (Marggraf et al., 2020; Oeztuerk et al., 2021; Roche et al., 2016), but associations between these variables and verbal fluency semantic search properties have rarely been explored. These relationships are particularly important to understand within the context of early-stage psychosis, as an efficient index

of communication disturbance prior to long-term functional impairment could inform assessment and early intervention.

The present study used automated tools Word2Vec and VFClust to characterize the semantic and phonetic similarity of words produced during verbal fluency tests in early psychosis. The first study aim was to test whether individuals with early-stage psychosis differ from healthy controls in automated metrics of category and letter verbal fluency performance. We hypothesized that individuals with early-stage psychosis compared to controls would produce fewer responses, with greater impairments in category than letter fluency (Mesholam-Gately et al., 2009). We also predicted that individuals with early-stage psychosis would produce fewer switches between clusters yet have similar cluster sizes (Elvevåg et al., 2002; Lundin et al., 2020; Moore et al., 2006; Piras et al., 2019). Regarding similarity, we predicted that early-stage psychosis participants would exhibit lower semantic similarity than controls (Elvevåg et al., 2007; Pauselli et al., 2018), particularly while searching within clusters, given recent findings of prolonged within-cluster search times in schizophrenia (Lundin et al., 2020). The second study aim was to test whether higher semantic similarity in category fluency responses was associated with greater speech organization and better social functioning in individuals with and without early-stage psychosis.

2. Methods

2.1 Participants

Early-stage psychosis participants ($n=20$) were recruited from a Midwestern outpatient early psychosis treatment center. They were adult outpatients with a primary diagnosis of schizophrenia, schizoaffective disorder, schizophreniform disorder, or psychotic disorder not otherwise specified within five years of illness onset, as assessed by the Structured Clinical Interview for DSM-IV-TR Disorders-Patient Edition (SCID-I/P; First et al., 2002). Other inclusion criteria were as follows: age 18–40, English fluency, no current substance dependence, no history of neurological illness or injury resulting in loss of consciousness >5 minutes, and verbal IQ>70. All early-stage psychosis participants were prescribed an oral antipsychotic and in nonacute phases of illness, with no change in medication or illness phase during the past month.

Healthy controls ($n=20$) were recruited from the community (e.g., flyers, local ads) and were matched with early-stage psychosis participants on sex and race. Inclusion criteria were the same as those for early-stage psychosis with the exception that current psychiatric diagnosis or reported history of psychotic symptoms was exclusionary. All participants underwent clinical assessment of psychotic symptoms with the Positive and Negative Syndrome Scale (PANSS; Kay et al., 1987). Study procedures were approved by local Institutional Review Boards, and all participants provided informed consent.

2.2 Verbal fluency tests

Verbal fluency tests were administered from the Brief Assessment of Cognition in Schizophrenia (BACS) neuropsychological battery (Keefe et al., 2004). In the category

fluency test, participants were instructed to verbally list as many different animals as they could in one minute. In the letter fluency tests, participants were asked to list as many different items as they could that began with the letters F and S, separately for one minute each. They were instructed not to list proper nouns or the same word with different endings (e.g., “look” and “looking”).

2.3 Semantic similarity of category fluency responses

Category verbal fluency data were analyzed using the skip-gram Word2Vec model with negative sampling (Mikolov et al., 2013), which was trained on a Google News corpus containing three billion word tokens. Using Word2Vec, word tokens were treated as high-dimensional vector representations, and cosine-based similarity values were computed between these vectors based on learned predictions of the frequency of direct and indirect co-occurrence across contexts in the corpus. Thus, a higher cosine value between word tokens indicates greater predicted context co-occurrence and is used as a proxy for semantic similarity (e.g., “cat” and “dog” typically have higher cosine values than “cat” and “zebrafish”). A semantic similarity matrix of 676 animal names generated from Word2Vec used previously (Lundin et al., 2020) was applied to the participant responses, with cosine-based similarity values ranging from 0 to 1. In the few instances in which animal responses were absent from the similarity matrix, replacement animals were used (e.g., “gold raven” was replaced with “raven”). Repeated responses were included in semantic analysis but not counted as correct responses (Troyer et al., 1997).

Semantic clusters and switches were designated using the similarity drop model, in which a relative drop in pairwise semantic similarity values in the participants’ response stream is designated as a switch (Hills et al., 2012). In particular, if the similarity between words A and B is represented by $S(A, B)$ and the participant produces consecutive words A, B, C, D , a switch is designated after response B if $S(A, B) > S(B, C)$ and $S(B, C) < S(C, D)$. Mean cluster size was computed for each participant as the average number of responses in each cluster including the first cluster response. Switches were counted as the number of transitions to new clusters, not including the first response produced. Fluency score was calculated as the total number of unique animal responses produced. Pairwise similarity values were averaged separately for cluster and switch responses. Of note, the experimenter handwriting for the first two responses of one participant could not be deciphered. To prioritize data inclusion, similarity analysis for this participant started at the third response, and the first two responses were included in the fluency score total as they were marked as correct at the time of the study.

2.4 Phonetic similarity of letter fluency responses

Letter verbal fluency data were analyzed using VFClust version 0.1.1 (Ryan et al., 2013) with Python 2.7. VFClust assesses phonetic similarity by first using a modified version of the CMU Pronouncing Dictionary (version *cmudict.0.7a*) to convert verbal fluency responses into phonetic representations. Next, the “edit distance” method assigns similarity values equal to 1 minus the Levenshtein distance (Levenshtein, 1966) between two consecutive responses, normalized to the length of the longer response. Similarity values range from 0 to 1, with higher values indicating greater phonetic similarity, or fewer

conversions needed to convert the response from one to the other (e.g., “seen” and “screen” are more phonetically similar than “seen” and “shirt”). Repeated responses were included in phonetic similarity analysis but not counted as correct responses.

Phonetic clusters (“chains” in VFClust) were designated as non-overlapping groupings of consecutive words whose pairwise similarity exceeds an empirically validated threshold (F:0.333; S: 0.286). Mean chain size was computed as the average of the number of responses within a chain including the first chain response and including singletons (chains of length 1). Switch count was defined as the sum of the number of transitions to new chains and singletons, not including the first response produced. Fluency score was counted as the total number of unique responses, excluding exact repetitions, stem repetitions (e.g., “traveled” not counted if “travel” was said), and phonemic errors (e.g., “center” generated for S fluency). Pairwise phonetic similarity values for chain and switch responses were averaged separately for analysis.

2.5 Speech organization

Speech samples were obtained of participants discussing neutrally valenced memories that focused on either their daily routine or place of residence for two minutes while speaking into a head-mounted microphone (see Minor et al., 2016 for further detail). Disorganized speech was measured in the transcribed speech samples using the Communication Disturbances Index (CDI; Docherty et al., 1996), a validated, behaviorally based instrument that identifies specific instances of disorganization from natural speech via a summary score (number of disorganized segments / total words X 100) (Cohen et al., 2014; Merrill et al., 2017). Because the score is generated as a ratio of instances per 100 words, it accounts for differences in the amount of speech generated by participants. Instances of disorganized speech are counted if the intended meaning of a word or phrase is unclear and if the lack of clarity impairs understanding of the larger communication. CDI scores can be sensitive to lower levels of disorganization than clinician observations alone, allowing for more direct comparisons with healthy controls and demonstrating sensitivity for distinguishing healthy controls and unaffected biological relatives of individuals with schizophrenia (Docherty et al., 1998). The last author and a trained graduate student blind to group membership rated the speech samples and discussed discrepancies in weekly consensus meetings (interrater reliability on 30 randomly selected narratives rated prior to consensus meeting: $r=0.84$).

2.6 Social functioning

Social functioning was measured in all participants using the Global Functioning (GF): Social scale (Cornblatt et al., 2007), a clinician-administered interview. The GF: Social assesses age-appropriate social contact with friends, family members, and intimate relationships. Functioning is rated on a 10-point scale, with higher scores indicating superior functioning and lower scores indicating social dysfunction. The GF: Social has previously shown high interrater reliability and good construct validity in clinical high risk (Cornblatt et al., 2007) and early-stage psychosis (Piskulic et al., 2011) samples.

¹Of note, the VFClust code was modified to remove the steps of (1) rejecting words not included in the English Open Word List and (2) combining instances of stem repetitions into one response.

2.7 Data analysis

Data analyses and plotting were conducted using SPSS (IBM SPSS Statistics, Version 27.0) and R statistical software package version 4.0.2 (R Core Team, 2020) in the RStudio environment version 1.3.1073 (RStudio Team, 2020). All tests were considered significant at a threshold of $p < .05$. First, independent samples t -tests were conducted between early-stage psychosis and control groups on total unique responses produced in each fluency task, switch count, mean cluster or chain size, cluster-related similarity, and switch-related similarity. For variables with non-normal distributions, group differences were verified with non-parametric Mann-Whitney U Tests. Next, non-parametric Spearman correlations were conducted collapsed across and separately within diagnostic groups to test the strength of the associations between verbal fluency semantic similarity and (1) CDI score and (2) GF: Social scores.

3. Results

3.1 Automated verbal fluency measures

Early-stage psychosis and healthy control participants were well-matched on age, sex, and race/ethnicity, but differed by education level (Table 1; Minor et al., 2016). Visualizations of selected participants' verbal fluency response trajectories over the course of category and letter tasks are shown in Figures 1–2. In preliminary analyses between F and S letter fluency variables, total unique responses ($r=0.72$, $p<.001$) and switch count ($r=0.48$, $p=.002$) were significantly correlated between tasks, yet mean chain size was not significantly correlated using Spearman's rho ($\rho=0.21$, $p=.189$). Therefore, F and S fluency variables were examined separately.

Participants with early-stage psychosis produced fewer correct responses than healthy controls in all verbal fluency tasks (Table 2). While both groups produced more correct responses in category than letter fluency, differences between controls and early-stage psychosis groups were larger for category than letter fluency tasks (Figure 3). The number of repeated responses across participants ranged from 0 to 3 for category and letter tasks and did not significantly differ between diagnostic groups.

The early-stage psychosis group compared to the control group switched less frequently during both category and letter fluency tasks; however, groups had similar average semantic and phonetic cluster or chain sizes and did not differ in the percentage of responses designated as switches (Table 2). Early-stage psychosis participants had reduced mean cluster-related semantic similarity than control participants (Figure 4). Diagnostic groups did not significantly differ from one another in mean cluster-related phonetic similarity or mean switch-related semantic or phonetic similarity. A follow-up linear regression confirmed that the early-stage psychosis group still exhibited lower within-cluster similarity than controls when fluency score was included as a predictor ($\beta=-0.04$, $t(37)=-2.44$, $p=.02$). This finding also remained significant when repeated responses were excluded from semantic analysis ($t(38)=-2.94$, $p=.006$).

3.2 Verbal fluency semantic similarity, disorganized speech, and social functioning

In the full sample, greater mean verbal fluency cluster- and switch-related semantic similarity values were significantly correlated with lower levels of speech disorganization as measured by CDI score (Table 3; Figure 5). Correlations with CDI score were in the same negative direction for both groups but were stronger in the control group. Greater verbal fluency cluster-related semantic similarity was also correlated with higher GF: Social score in the full sample, but this relationship was driven by the early-stage psychosis group.

4. Discussion

This study characterized responses from an early psychosis sample on verbal fluency tests using automated semantic (Word2Vec; Mikolov et al., 2013) and phonetic (VFClust; Ryan et al., 2013) similarity analysis. Findings demonstrated that individuals with early-stage psychosis produced clusters of related animal responses that were less semantically associated than those of non-psychiatric controls. Additionally, higher semantic similarity of verbal fluency responses in the full participant sample was correlated with lower speech disorganization, although more strongly within the control group, as well as correlated with higher social functioning, primarily in the psychosis group. Findings suggest that alterations in semantic memory processes may begin early in the course of psychotic illness (Bedi et al., 2015; Corcoran et al., 2018), and that the capacity to produce semantically similar words relates in part to the production of organized speech as well as the quantity and quality of interpersonal relationships. These results contribute to the broader literature illustrating the value of applying computational methods to linguistic data to index fine-grained components of generating coherent speech in individuals with psychosis (Hitczenko et al., 2021).

Theories of disorganized semantic memory representations and/or retrieval processes in psychotic disorders (Kuperberg, 2010) have led to the hypothesis that verbal fluency semantic clusters are smaller compared to normative populations; however, evidence supporting this hypothesis has not been frequently reported. As in the present work, studies have often found individuals with psychotic disorders produce fewer switches between clusters but exhibit clusters of similar average size to controls (Elvevåg et al., 2002; Lundin et al., 2020; Moore et al., 2006; Piras et al., 2019; but see van Beilen et al., 2004). The nuanced measure of semantic similarity was more informative than cluster size, as there was specifically lower similarity between responses while searching within clusters but not while switching in the early-stage psychosis group compared to controls. This broadly aligns with our past study finding prolonged within-cluster search time in schizophrenia (Lundin et al., 2020), again pointing to alterations in the search for semantically associated concepts. However, this past study showed intact local search cue salience from a cognitive foraging model, suggesting comparable within-cluster similarity between the groups, in contrast with the current results. Given that timing information was not available for the present data, future studies should investigate whether lower within-cluster similarity and/or longer within-cluster search times replicate in early and later-stage psychosis.

This study replicated prior findings of greater impairment in category fluency compared to letter fluency in early-stage psychosis (Mesholam-Gately et al., 2009). Category and letter fluency recruit overlapping cognitive processes such as processing speed and executive

functioning (Shao et al., 2014; Unsworth et al., 2011). They also recruit distinct processes, as category fluency performance benefits more from the generation of semantic associations, whereas letter fluency performance benefits more from usage of phonetic cues (Shao et al., 2014; Troyer et al., 1997). This study used VFClust to analyze phonetic similarity as a novel application in an early psychosis population and demonstrated comparable phonetic similarity of responses between diagnostic groups. Usage of automated phonetic analysis is less time- intensive and error-prone than manual scoring methods and in this case provided information as to which linguistic processes are intact in early psychosis. Overall, findings support a partially selective impairment in finding semantically related concepts in psychosis—rather than a more generalized deficit in memory search.

This study also revealed associations between generation of semantically related verbal fluency responses, organization of free speech, and social functioning. Past work has shown that more disordered speech relates to poorer social functioning in psychosis (Marggraf et al., 2020; Oeztuerk et al., 2021; Roche et al., 2016), but the relationship between social functioning and semantic relatedness of verbal fluency responses is rarely examined. The present findings indicated that poorer clinician-rated social functioning modestly correlated with lower cluster- related semantic similarity of fluency responses, and that this relationship was driven by the early-stage psychosis group. Further work in larger samples is needed to investigate how various neurocognitive and communication-related functions may explain this relationship. Interestingly, lower verbal fluency response semantic similarity significantly correlated with greater disorganized speech, but this relationship was stronger in the control than psychosis group, despite the lower variability of CDI scores in controls. This finding suggests that in normative functioning, the retrieval of semantic associations in memory relates to the production of organized speech. The weaker relationship in the clinical sample was unexpected. This may suggest that disrupted verbal fluency performance and disorganized speech index unique facets of psychosis-related linguistic disturbances. Alternatively, future studies are encouraged to test whether fluency response patterns relate more strongly to free speech organization at deeper levels of semantic cohesion using automated tools such as Coh-Metrix (McNamara et al., 2014).

The present findings should be interpreted within the context of the study's limitations. One limitation relates to the automated methods used to designate clusters and switches. The similarity drop model (Hills et al., 2012) applied to continuous corpus-based semantic similarity values likely performs better than hand-coded categorization schemes at capturing a variety of dimensions on which animals are similar as well as individual differences in response production (e.g., one participant listing various dog breeds vs. another listing pets including “dog”). This algorithm is limited though in that it may designate a switch for a relatively small drop in similarity where a cluster designation may fit better. This being said, follow-up analyses coding the present clusters and switches using the widely used Troyer norms (Troyer et al., 1997) showed a correspondence of 70% on average with the designations from the similarity drop model, indicating that the current method performed similarly to existing studies. A related limitation is that the distinct switching designation methods used for category and letter fluency data make it difficult to compare search behavior between the tasks. Future verbal fluency switching models could incorporate weighted parameters of both semantic and phonetic similarity and comparable data-driven

thresholds to achieve a more comprehensive understanding of the linguistic features of memory retrieval. A final limitation is the small sample size of each diagnostic group; future studies are needed to verify the present results in larger participant samples.

Overall, this study demonstrated altered search for semantically related words in early-stage psychosis using automated linguistic analysis and provided preliminary evidence for associations between verbal fluency search, speech organization, and social functioning. Future studies are warranted to directly compare these variables between individuals in a clinical high- risk state, early-stage psychosis, and long-term psychosis to better understand the potential impact of linguistic disturbances on social functioning over the course of illness. Future work is also encouraged to further automate the analysis of verbal fluency responses using computerized speech-to-text transcription and response onset time designation tools (Holmlund et al., 2019) to more efficiently examine patterns of semantic and phonetic search in individuals across the psychotic spectrum.

Acknowledgements

We would like to thank the participants of this study, Beshawn Davis, Matthew Marggraf, and research assistants in the Cognition, Language, and Affect in Serious Psychopathology (CLASP) Laboratory for assistance in data collection and speech sample scoring, and Peter Todd for valuable discussions that informed this work.

Funding

This research was supported by the National Institutes of Health, National Center for Advancing Translational Sciences, Clinical and Translational Sciences Award (KL2TR001106 and UL1TR001108 to A. Shekhar, with sub-award to K.S.M.), the National Institute of Mental Health (T32 MH103213 to N.B.L.), and the National Science Foundation Graduate Research Fellowship Program (NSF GRFP 1342962 to N.B.L.).

References

- Andreasen NC, 1986. Scale for the assessment of thought, language, and communication (TLC). *Schizophr Bull* 12, 473–482. 10.1093/schbul/12.3.473 [PubMed: 3764363]
- Bedi G, Carrillo F, Cecchi GA, Slezak DF, Sigman M, Mota NB, Ribeiro S, Javitt DC, Copelli M, Corcoran CM, 2015. Automated analysis of free speech predicts psychosis onset in high-risk youths. *NPJ Schizophr* 1, 15030. 10.1038/npjzsch.2015.30 [PubMed: 27336038]
- Bleuler E, 1911. *Dementia praecox or the group of schizophrenias* (Zinkin J, Translation, 1950). International Universities Press, New York, NY.
- Bokat CE, Goldberg TE, 2003. Letter and category fluency in schizophrenic patients: a meta-analysis. *Schizophr Res* 64, 73–78. 10.1016/s0920-9964(02)00282-7 [PubMed: 14511803]
- Bousfield WA, Sedgewick CHW, 1944. An Analysis of Sequences of Restricted Associative Responses. *The Journal of General Psychology* 30, 149–165. 10.1080/00221309.1944.10544467
- Bozikas VP, Kosmidis MH, Karavatos A, 2005. Disproportionate impairment in semantic verbal fluency in schizophrenia: differential deficit in clustering. *Schizophr Res* 74, 51–59. 10.1016/j.schres.2004.05.001 [PubMed: 15694754]
- Cohen AS, Auster T, Callaway D, MacAulay RK, Minor KS, 2014. Neurocognitive underpinnings of language disorder: Contrasting schizophrenia and mood disorders. *Journal of Experimental Psychopathology* 5, 492–502. 10.5127/jep.034213
- Corcoran CM, Carrillo F, Fernández-Slezak D, Bedi G, Klim C, Javitt DC, Bearden CE, Cecchi GA, 2018. Prediction of psychosis across protocols and risk cohorts using automated language analysis. *World Psychiatry* 17, 67–75. 10.1002/wps.20491 [PubMed: 29352548]
- Corcoran CM, Mittal VA, Bearden CE, E Gur R, Hitzenko K, Bilgrami Z, Savic A, Cecchi GA, Wolff P, 2020. Language as a biomarker for psychosis: A natural language processing approach. *Schizophr Res* 226, 158–166. 10.1016/j.schres.2020.04.032 [PubMed: 32499162]

- Cornblatt BA, Auther AM, Niendam T, Smith CW, Zinberg J, Bearden CE, Cannon TD, 2007. Preliminary findings for two new measures of social and role functioning in the prodromal phase of schizophrenia. *Schizophr Bull* 33, 688–702. 10.1093/schbul/sbm029 [PubMed: 17440198]
- Docherty NM, 2012. On identifying the processes underlying schizophrenic speech disorder. *Schizophr Bull* 38, 1327–1335. 10.1093/schbul/sbr048 [PubMed: 21562141]
- Docherty NM, DeRosa M, Andreasen NC, 1996. Communication disturbances in schizophrenia and mania. *Arch Gen Psychiatry* 53, 358–364. 10.1001/archpsyc.1996.01830040094014 [PubMed: 8634014]
- Docherty NM, Rhinewine JP, Labhart RP, Gordinier SW, 1998. Communication disturbances and family psychiatric history in parents of schizophrenic patients. *J Nerv Ment Dis* 186, 761–768. 10.1097/00005053-199812000-00004 [PubMed: 9865814]
- Elvevåg B, Fisher JE, Gurd JM, Goldberg TE, 2002. Semantic clustering in verbal fluency: schizophrenic patients versus control participants. *Psychol Med* 32, 909–917. 10.1017/S0033291702005597 [PubMed: 12171385]
- Elvevåg B, Foltz PW, Weinberger DR, Goldberg TE, 2007. Quantifying incoherence in speech: an automated methodology and novel application to schizophrenia. *Schizophr Res* 93, 304–316. 10.1016/j.schres.2007.03.001 [PubMed: 17433866]
- First MB, Spitzer RL, Gibbon M, Williams JBW, 2002. Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Patient Edition (SCID-I/P). New York: Biometrics Research, New York State Psychiatric Institute.
- Gruenewald PJ, Lockhead GR, 1980. The free recall of category examples. *Journal of Experimental Psychology: Human Learning & Memory* 6, 225–240. 10.1037/0278-7393.6.3.225
- Hills TT, Jones MN, Todd PM, 2012. Optimal foraging in semantic memory. *Psychol Rev* 119, 431–440. 10.1037/a0027373 [PubMed: 22329683]
- Hitzenko K, Mittal VA, Goldrick M, 2021. Understanding Language Abnormalities and Associated Clinical Markers in Psychosis: The Promise of Computational Methods. *Schizophr Bull* 47, 344–362. 10.1093/schbul/sbaa141 [PubMed: 33205155]
- Holmlund TB, Cheng J, Foltz PW, Cohen AS, Elvevåg B, 2019. Updating verbal fluency analysis for the 21st century: Applications for psychiatry. *Psychiatry Res* 273, 767–769. 10.1016/j.psychres.2019.02.014 [PubMed: 31207864]
- Iter D, Yoon JH, Jurafsky D, 2018. Automatic detection of incoherent speech for diagnosing schizophrenia, in: *Proceedings of the Fifth Workshop on Computational Linguistics and Clinical Psychology: From Keyboard to Clinic*. Presented at the Association for Computational Linguistics, New Orleans, LA, pp. 136–146.
- Kay SR, Fiszbein A, Opler LA, 1987. The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophr Bull* 13, 261–276. 10.1093/schbul/13.2.261 [PubMed: 3616518]
- Keefe RSE, Goldberg TE, Harvey PD, Gold JM, Poe MP, Coughenour L, 2004. The Brief Assessment of Cognition in Schizophrenia: reliability, sensitivity, and comparison with a standard neurocognitive battery. *Schizophr Res* 68, 283–297. 10.1016/j.schres.2003.09.011 [PubMed: 15099610]
- Kim N, Kim J-H, Wolters MK, MacPherson SE, Park JC, 2019. Automatic Scoring of Semantic Fluency. *Front Psychol* 10, 1020. 10.3389/fpsyg.2019.01020 [PubMed: 31156496]
- Kuperberg GR, 2010. Language in schizophrenia Part 1: an Introduction. *Lang Linguist Compass* 4, 576–589. 10.1111/j.1749-818X.2010.00216.x [PubMed: 20936080]
- Landauer TK, Dumais ST, 1997. A solution to Plato’s problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychol Rev* 104, 211–240. 10.1037/0033-295X.104.2.211
- Levenshtein VI, 1966. Binary codes capable of correcting deletions, inserts, and reversals. *Soviet Physics, Doklady* 10, 707–710.
- Lundin NB, Todd PM, Jones MN, Avery JE, O’Donnell BF, Hetrick WP, 2020. Semantic Search in Psychosis: Modeling Local Exploitation and Global Exploration. *Schizophr Bull Open* 1, sgaa011. 10.1093/schizbulopen/sgaa011

- Marggraf MP, Lysaker PH, Salyers MP, Minor KS, 2020. The link between formal thought disorder and social functioning in schizophrenia: A meta-analysis. *Eur Psychiatry* 63, e34. 10.1192/j.eurpsy.2020.30 [PubMed: 32200776]
- McNamara DS, Graesser AC, McCarthy PM, Cai Z, 2014. Automated evaluation of text and discourse with Coh-Metrix. Cambridge University Press, New York, NY.
- Merrill AM, Karcher NR, Cicero DC, Becker TM, Docherty AR, Kerns JG, 2017. Evidence that communication impairment in schizophrenia is associated with generalized poor task performance. *Psychiatry Res* 249, 172–179. 10.1016/j.psychres.2016.12.051 [PubMed: 28104564]
- Mesholam-Gately RI, Giuliano AJ, Goff KP, Faraone SV, Seidman LJ, 2009. Neurocognition in first-episode schizophrenia: a meta-analytic review. *Neuropsychology* 23, 315–336. 10.1037/a0014708 [PubMed: 19413446]
- Mikolov T, Chen K, Corrado G, Dean J, 2013. Efficient Estimation of Word Representations in Vector Space. arXiv:1301.3781 [cs].
- Minor KS, Marggraf MP, Davis BJ, Mehdiyoun NF, Breier A, 2016. Affective systems induce formal thought disorder in early-stage psychosis. *J Abnorm Psychol* 125, 537–542. 10.1037/abn0000156 [PubMed: 26999283]
- Minor KS, Willits JA, Marggraf MP, Jones MN, Lysaker PH, 2019. Measuring disorganized speech in schizophrenia: automated analysis explains variance in cognitive deficits beyond clinician-rated scales. *Psychol Med* 49, 440–448. 10.1017/S0033291718001046 [PubMed: 29692287]
- Moore DJ, Savla GN, Woods SP, Jeste DV, Palmer BW, 2006. Verbal fluency impairments among middle-aged and older outpatients with schizophrenia are characterized by deficient switching. *Schizophr Res* 87, 254–260. 10.1016/j.schres.2006.06.005 [PubMed: 16854567]
- Oeztuerk OF, Pigoni A, Antonucci LA, Koutsouleris N, 2021. Association between formal thought disorders, neurocognition and functioning in the early stages of psychosis: a systematic review of the last half-century studies. *Eur Arch Psychiatry Clin Neurosci*. 10.1007/s00406-021-01295-3
- Pauselli L, Halpern B, Cleary SD, Ku BS, Covington MA, Compton MT, 2018. Computational linguistic analysis applied to a semantic fluency task to measure derailment and tangentiality in schizophrenia. *Psychiatry Res* 263, 74–79. 10.1016/j.psychres.2018.02.037 [PubMed: 29502041]
- Piras F, Piras F, Banaj N, Ciullo V, Vecchio D, Edden R. a. E., Spalletta G, 2019. Cerebellar GABAergic correlates of cognition-mediated verbal fluency in physiology and schizophrenia. *Acta Psychiatr Scand* 139, 582–594. 10.1111/acps.13027 [PubMed: 30887499]
- Piskulic D, Addington J, Auther A, Cornblatt BA, 2011. Using the global functioning social and role scales in a first-episode sample. *Early Interv Psychiatry* 5, 219–223. 10.1111/j.1751-7893.2011.00263.x [PubMed: 21481198]
- R Core Team, 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Robert PH, Lafont V, Medecin I, Berthet L, Thaubly S, Baudu C, Darcourt G, 1998. Clustering and switching strategies in verbal fluency tasks: comparison between schizophrenics and healthy adults. *J Int Neuropsychol Soc* 4, 539–546. 10.1017/s1355617798466025 [PubMed: 10050358]
- Roche E, Segurado R, Renwick L, McClenaghan A, Sexton S, Frawley T, Chan CK, Bonar M, Clarke M, 2016. Language disturbance and functioning in first episode psychosis. *Psychiatry Res* 235, 29–37. 10.1016/j.psychres.2015.12.008 [PubMed: 26699880]
- RStudio Team, 2020. RStudio: Integrated Development Environment for R. RStudio, PBC, Boston, MA.
- Ryan JO, Pakhomov S, Marino S, Bernick C, Banks S, 2013. Computerized analysis of a verbal fluency test, in: Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics. Presented at the Association for Computational Linguistics, Sofia, Bulgaria, pp. 884–889.
- Shao Z, Janse E, Visser K, Meyer AS, 2014. What do verbal fluency tasks measure? Predictors of verbal fluency performance in older adults. *Front Psychol* 5, 772. 10.3389/fpsyg.2014.00772 [PubMed: 25101034]
- Solovay MR, Shenton ME, Gasperetti C, Coleman M, Kestnbaum E, Carpenter JT, Holzman PS, 1986. Scoring manual for the Thought Disorder Index. *Schizophr Bull* 12, 483–496. 10.1093/schbul/12.3.483 [PubMed: 3764364]

- Troyer AK, Moscovitch M, Winocur G, 1997. Clustering and switching as two components of verbal fluency: evidence from younger and older healthy adults. *Neuropsychology* 11, 138–146. 10.1037//0894-4105.11.1.138 [PubMed: 9055277]
- Unsworth N, Spillers GJ, Brewer GA, 2011. Variation in verbal fluency: a latent variable analysis of clustering, switching, and overall performance. *Q J Exp Psychol (Hove)* 64, 447–466. 10.1080/17470218.2010.505292 [PubMed: 20839136]
- van Beilen M, Pijnenborg M, van Zomeren EH, van den Bosch RJ, Withaar FK, Bouma A, 2004. What is measured by verbal fluency tests in schizophrenia? *Schizophr Res* 69, 267–276. 10.1016/j.schres.2003.09.007 [PubMed: 15469198]

Highlights

- Speech abnormalities are evident in those with early-stage psychosis (EP)
- EP participants had greater response deficits in category than letter verbal fluency
- Semantic similarity was lower in EP than controls yet phonetic similarity was intact
- Fluency semantic similarity was linked to speech organization and social functioning

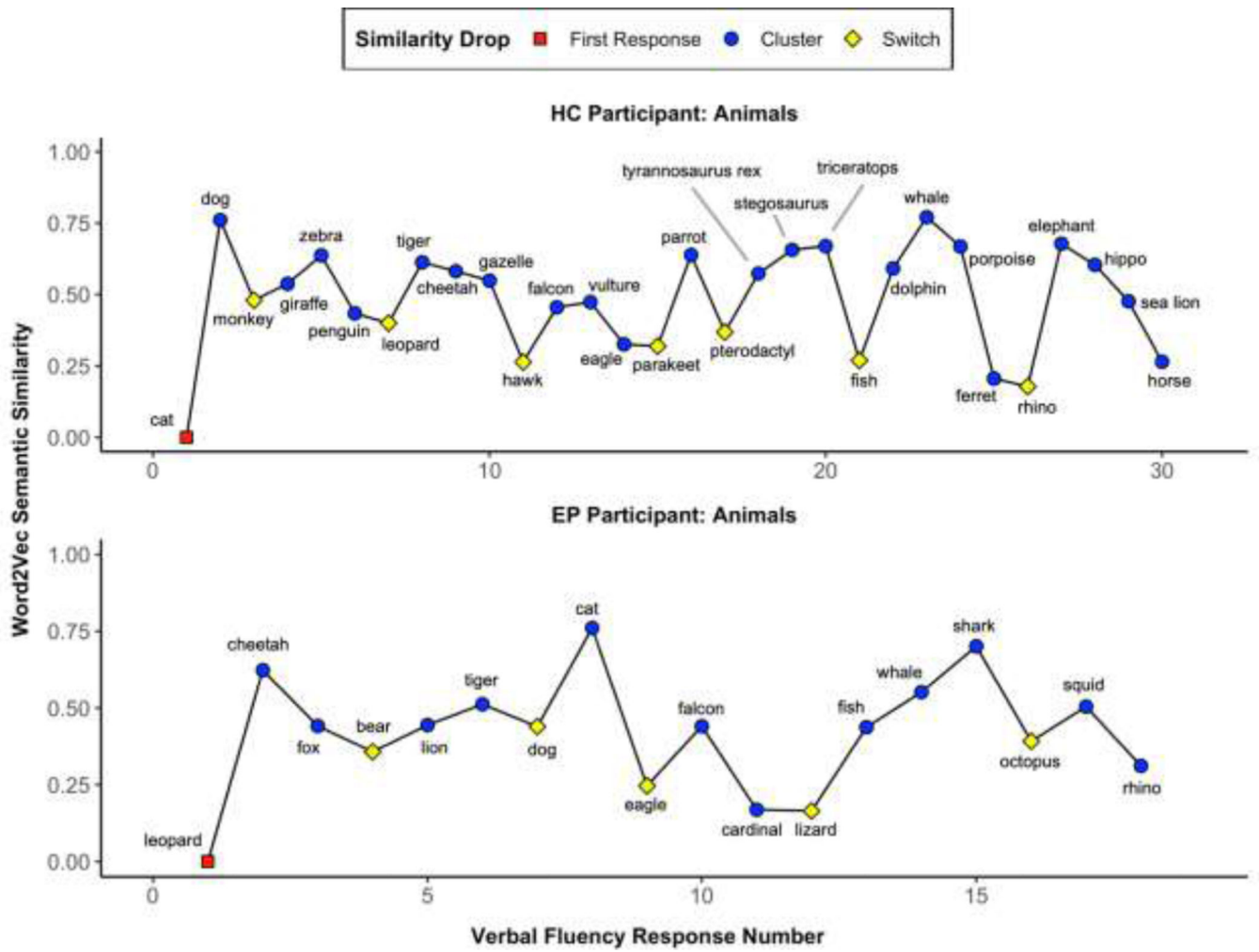


Figure 1. Category verbal fluency responses from selected example participants with corresponding Word2Vec semantic similarity values and similarity drop cluster or switch designations. HC = healthy control participant; EP = participant with early-stage psychosis.

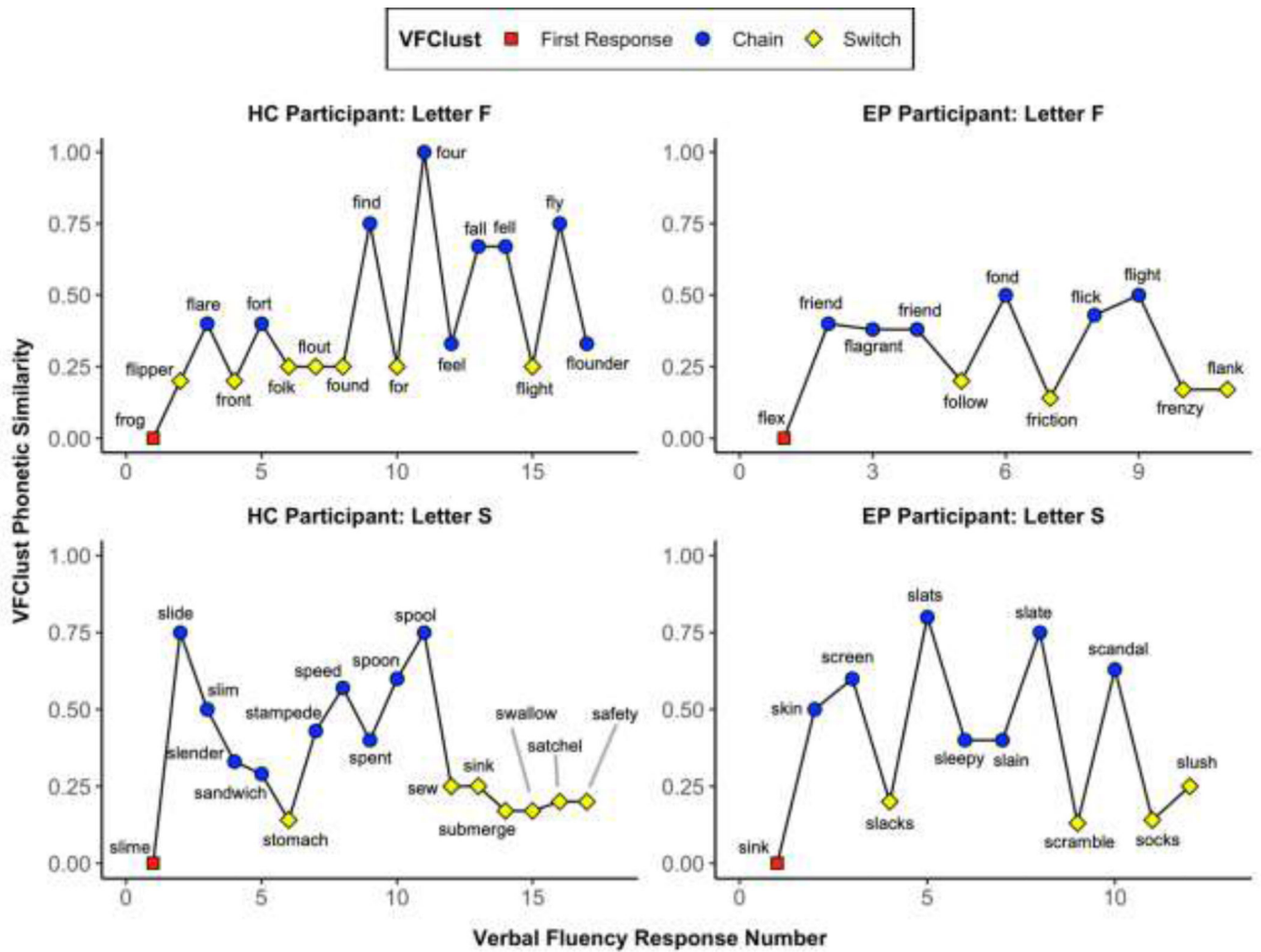


Figure 2. Letter verbal fluency responses from selected example participants with corresponding VFClust phonetic similarity values and cluster or switch designations. HC = healthy control participant; EP = participant with early-stage psychosis. Responses “for” and “four” were differentiated by verbal clarification by the participant (e.g., “four the number”).

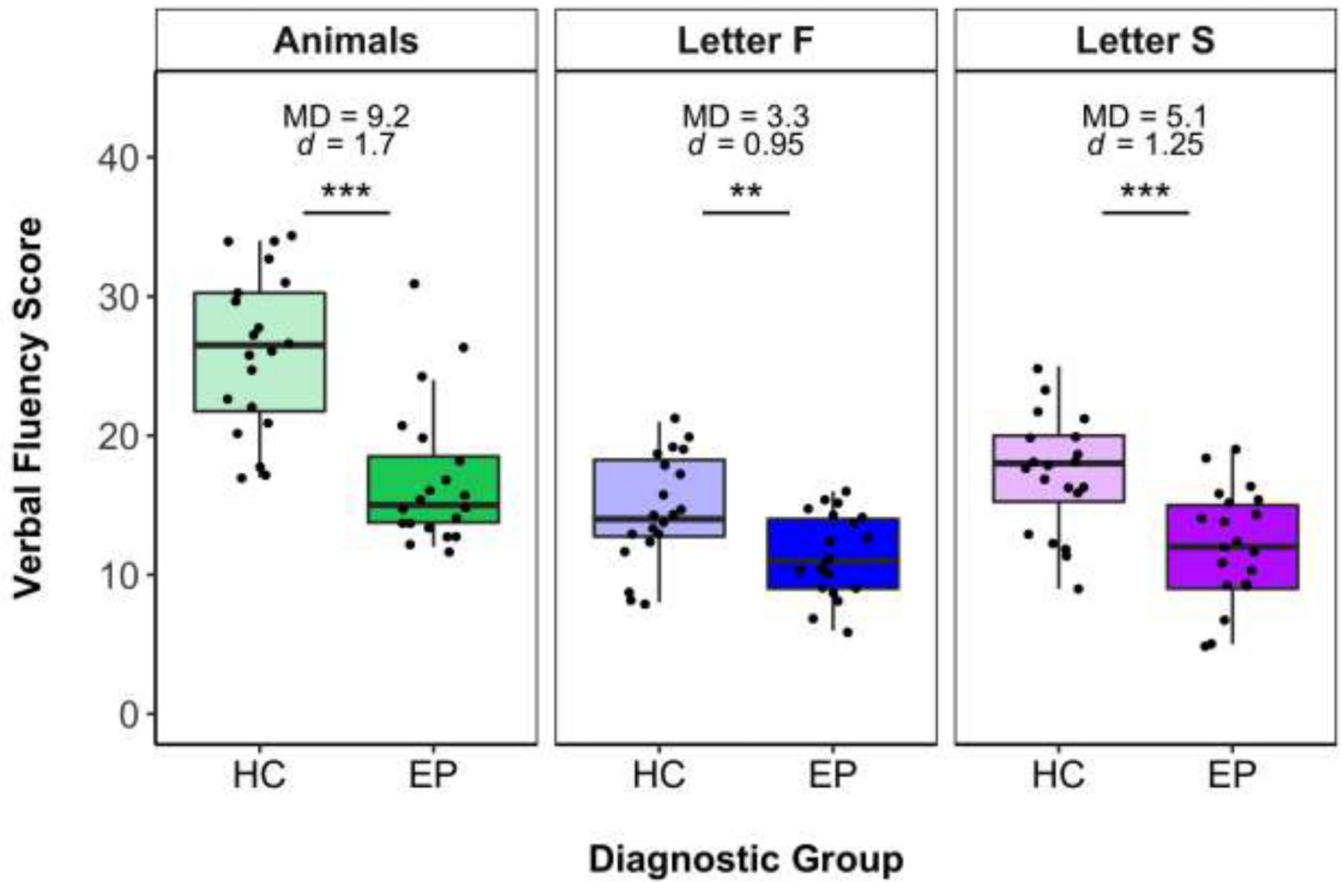


Figure 3.

Boxplots of the total number of correct responses produced in healthy control (HC) and early-stage psychosis (EP) groups in category and letter verbal fluency tasks. MD = mean difference; d = Cohen's d (standardized mean difference).

** $p < .01$; *** $p < .001$

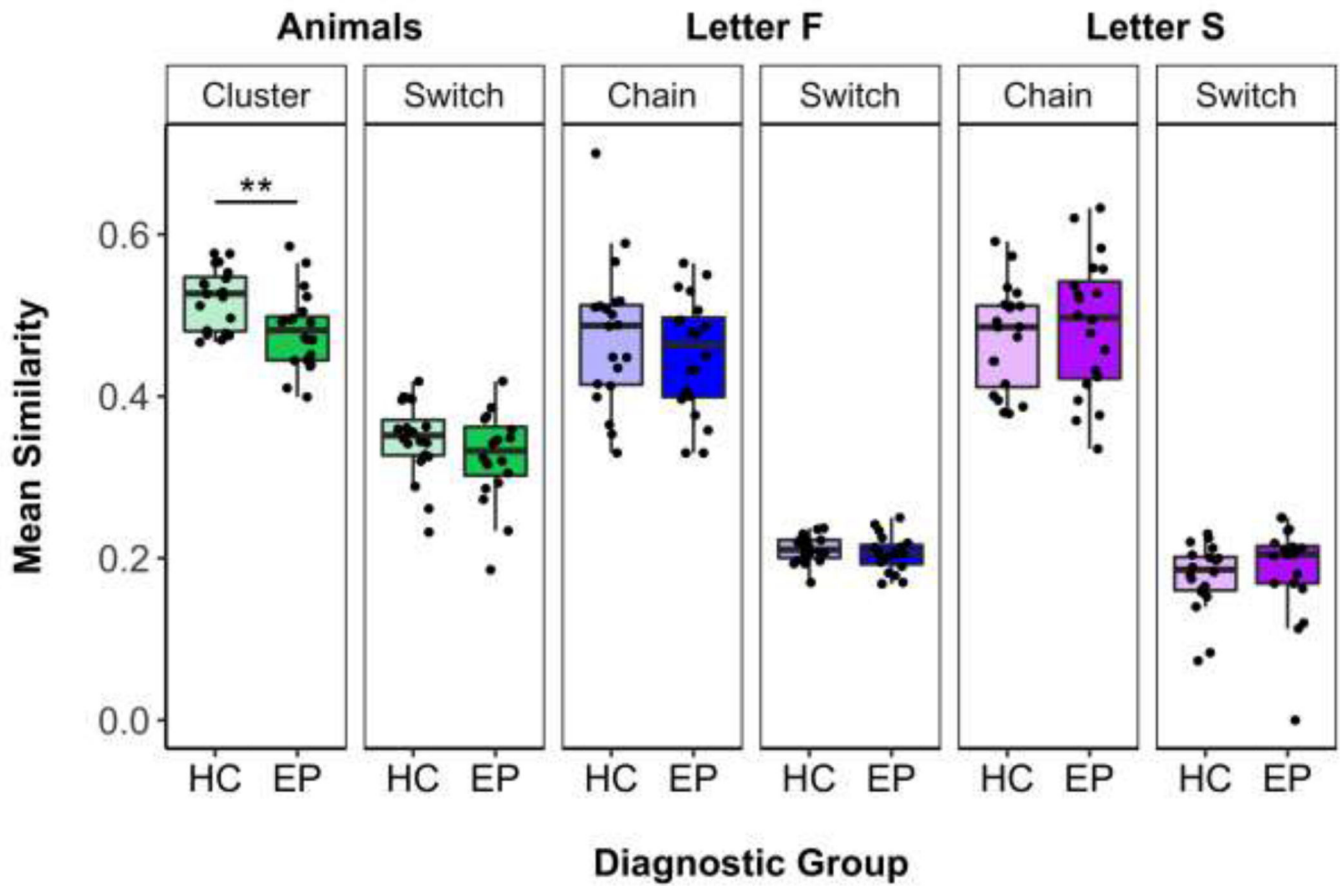


Figure 4. Boxplots of the mean semantic similarity of category verbal fluency responses and mean phonetic similarity of letter fluency responses, averaged across cluster-related and switch-related responses. HC = healthy control group; EP = early-stage psychosis group.
 ** $p < .01$

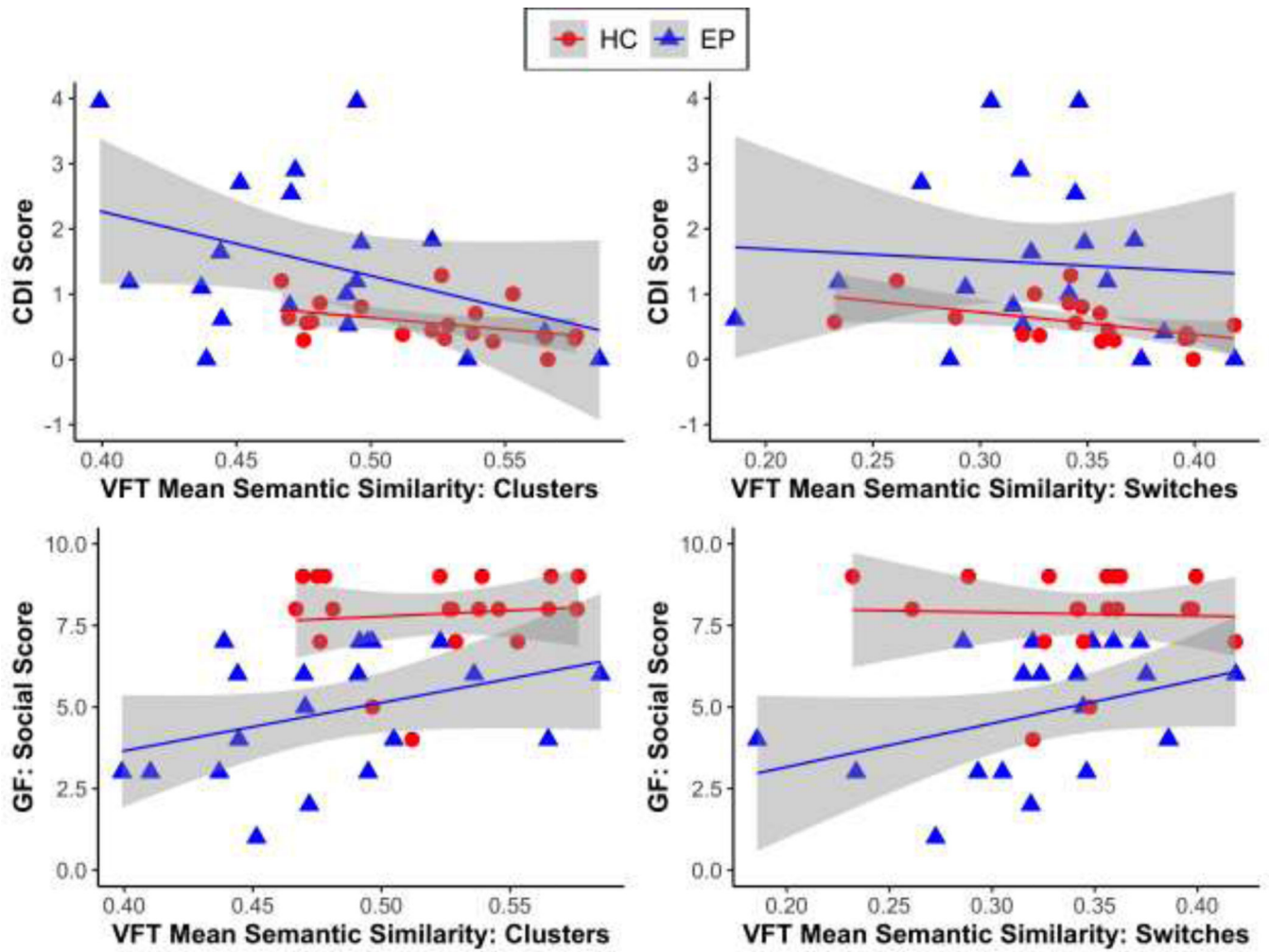


Figure 5. Scatter plots in early-stage psychosis (EP) and healthy control (HC) participants showing the correlations between category verbal fluency test (VFT) semantic similarity for cluster- and switch-related responses and: (1) *Top*: Communication Disturbances Index (CDI) scores and (2) *Bottom*: Global Functioning (GF): Social scores.

Table 1.

Sample characteristics

Variable	Healthy controls	Early-stage psychosis	<i>t</i> or χ^2
<i>n</i>	20	20	--
Age (years)	23.65 (5.58)	24.85 (4.53)	-0.747
Sex (F/M)	3/17	4/16	0.173
Race (AA/C/M)	12/7/1	13/6/1	0.117
Hispanic or Latino (yes/no)	1/19	2/18	0.36
Education level (HS/SC/BD)	3/13/4	12/6/2	8.65*
PANSS positive total score	8.2 (1.28)	11.55 (4.77)	-3.03**
PANSS negative total score	8.45 (2.14)	14.5 (6.03)	-4.23***
GF: Social score	7.85 (1.35)	4.85 (1.9)	5.76***
CDI score	0.57 (0.33)	1.48 (1.23) [†]	-3.13**

Sex, race, ethnicity, and education level are frequency values. Remaining values are presented as mean (standard deviation). Sex: F = female; M = male; Race: AA = African American or Black; C = Caucasian; M = Multiracial; Education level: HS = high school diploma or lower; SC = some college or associate degree; BD = bachelor's degree or higher; PANSS = Positive and Negative Syndrome Scale; GF-Social = Global Functioning: Social scale; CDI = Communication Disturbances Index.

[†]Missing value for one participant.

*
 $p < .05$

**
 $p < .01$

 $p < .001$

Table 2.

Computational measures of verbal fluency test behavior

Variable	Healthy controls	Early-stage psychosis	<i>t</i>	<i>d</i>
VFT score				
Animals	26.15 (5.7)	16.95 (5.08)	5.39 ^{***}	1.7
Letter F	14.7 (3.89)	11.4 (2.96)	3.02 ^{**}	0.95
Letter S	17.2 (4.2)	12.1 (3.99)	3.94 ^{***}	1.25
Percentage switches				
Animals	28.59 (5.41)	29.22 (6.14)	-0.34	0.11
Letter F	55.84 (17.61)	52.96 (20.11)	0.48	0.15
Letter S	52.83 (17.43)	50.08 (15.86)	0.52	0.16
Switch count				
Animals	7.4 (2.19)	4.75 (1.8)	4.18 ^{***}	1.32
Letter F	7.7 (2.87)	5.7 (2.03)	2.55 [*]	0.81
Letter S	8.3 (2.64)	6 (2.97)	2.59 [*]	0.82
Mean cluster or chain size				
Animals	3.27 (0.55)	3.05 (0.49)	1.32	0.42
Letter F ⁺	1.8 (0.57)	1.96 (1.06)	-0.6	0.19
Letter S	1.93 (0.61)	1.82 (0.37)	0.66	0.21
Mean cluster-related similarity				
Animals	0.52 (0.04)	0.48 (0.05)	2.95 ^{**}	0.93
Letter F	0.48 (0.09)	0.45 (0.07)	0.94	0.3
Letter S	0.47 (0.06)	0.49 (0.08)	-0.6	0.19
Mean switch-related similarity				
Animals	0.35 (0.05)	0.33 (0.05)	1.24	0.39
Letter F	0.21 (0.02)	0.21 (0.02)	0.83	0.27
Letter S ⁺	0.18 (0.04)	0.19 (0.06)	-0.69	0.22

Values are presented as mean (standard deviation). VFT score = total number of correct responses in the verbal fluency test; *d* = Cohen's *d* effect size. Similarity values are Word2Vec semantic similarity (animal VFT) and VFClust phonetic similarity (F and S VFT).

⁺ Variables had non-normal distributions; group differences remained non-significant when using Mann Whitney U tests.

* $p < .05$

** $p < .01$

*** $p < .001$

Table 3.

Correlations between verbal fluency response semantic similarity, speech disorganization, and social functioning

	CDI Score			GF: Social Score		
	Group	<i>n</i>	Correlation	Group	<i>n</i>	Correlation
VFT cluster-related	All	39	-0.54 ^{***}	All	40	0.43 ^{**}
semantic similarity	HC	20	-0.45 [*]	HC	20	0.04
	EP	19	-0.25	EP	20	0.36
VFT switch-related	All	39	-0.46 ^{**}	All	40	0.28
semantic similarity	HC	20	-0.61 ^{**}	HC	20	0.01
	EP	19	-0.16	EP	20	0.38

Correlations are Spearman's rho. HC = healthy control group; EP = early-stage psychosis group; VFT = verbal fluency test; GF-Social = Global Functioning: Social scale; CDI = Communication Disturbances Index.

*
 $p < .05$

**
 $p < .01$

 $p < .001$