Modeling Acute and Compensated Language Disturbance in Schizophrenia

Uli Grasemann[1], **Ralph Hoffman**[2], **and Risto Miikkulainen**[1]

[1]Department of Computer Sciences, The University of Texas at Austin
[2]Department of Psychiatry, Yale University School of Medicine

Abstract

No current laboratory test can reliably identify patients with schizophrenia. Instead, key symptoms are observed via language, including derailments, where patients cannot follow a coherent storyline, and delusions, where false beliefs are repeated as fact. Brain processes underlying these and other symptoms remain unclear, and characterizing them would greatly enhance our understanding of schizophrenia. In this situation, computational models can be valuable tools to formulate testable hypotheses and to complement clinical research. This work aims to capture the link between biology and schizophrenic symptoms using DISCERN, a connectionist model of human story processing. Competing illness mechanisms proposed to underlie schizophrenia are simulated in DISCERN, and are evaluated at the level of narrative language, i.e. the same level used to diagnose patients. The result is the first simulation of abnormal storytelling in schizophrenia, both in acute psychotic and compensated stages of the disorder. Of all illness models tested, hyperlearning, a model of overly intense memory consolidation, produced the best fit to the language abnormalities of stable outpatients, as well as compelling models of acute psychotic symptoms. If validated experimentally, the hyperlearning hypothesis could advance the current understanding of schizophrenia, and provide a platform for developing future treatments for this disorder.

Introduction

Stories are more than a useful social construct – they are a crucial part of who we are. We make sense of the world by fitting our experience into a coherent narrative. In schizophrenia, this ongoing narrative breaks down. Disturbances in the perception and expression of reality can be observed through the stories a patient tells. Indeed, narrative language is the primary diagnostic tool, and clinicians use it every day to observe and evaluate its manifestations. The purpose of clinical interviews, then, is to use narrative language as a window into the schizophrenic mind. The main idea behind this research is that neural network models of storytelling can provide mechnistic explanations of what is seen through that window. These explanations can then be evaluated through narrative language – at the same level used to diagnose real patients.

The principal strength of neural network models lies in their ability to bridge the gap between complex mental states and behavior on the one hand and underlying neural information processing on the other. This ability is precisely what is needed in schizophrenia research, where the central challenge for decades has been to explain how underlying illness mechanisms could cause altered the behavior.

Consequently, the main goal of this research is to demonstrate that a neural network model can be used meaningfully to simulate possible illness mechanisms in schizophrenia. The different illness models will result in different language behavior, which can then be used to generate predictions about the underlying causes. Thus computational illness models have the potential to complement and guide future medical research.

On the other hand, progress in understanding schizophrenia is likely to lead to progress in basic cognitive science as well. Global, emergent faculties like understanding and telling stories, processing emotions, and forming long-term memories of real or imagined events are difficult to account for computationally, and mental illnesses where these faculties break down offer a unique opportunity to investigate how they emerge from their neural substrate. Developing methods to model such high-level behavior computationally is a second major goal of the research presented in this paper.

Schizophrenia

Schizophrenia is a common and disabling psychiatric disorder. Symptoms include hallucinations, bizarre behavior, delusions, and disorganized language that is hard for listeners to follow. These *psychotic* symptoms tend to wax and wane over time, and in later stages often give way to *negative* symptoms, including blunted emotions and reduced language output. This paper focuses on symptoms that are observed directly via language, most prominently

1. **Delusions**, which are pathological false beliefs. Delusions often share common themes, like being watched by the CIA or being controlled by outside forces. Patients with schizophrenia tend to insert themselves or persons they know into imaginary narratives. Such agent-slotting errors are thought to be the cause of the plots and conspiracies that characterize persecutory delusions.

2. **Disorganized speech**, which refers to fluent spoken language that fails to communicate effectively. It is believed to reflect impaired verbal thought (thought disorder). One of the most prominent signs are derailments, i.e. jumps from one topic or story to another without apparent cause.

Symptoms in schizophrenia vary among patients. Clinical subtypes of schizophrenia include the *paranoid type*, where symptoms include delusions and hallucinations but not prominent language disorganization, and the *disorganized type*, where symptoms are dominated by disorganized language and behavior.

Treatments for schizophrenia mainly rely on medication that can help manage psychotic symptoms. However, these drugs often have severe and dangerous side-effects, and do not help all patients or address all symptoms (Kapur and Mamo, 2003). These shortcomings make a better understanding of schizophrenia an important goal, because it would likely lead to more effective drugs, and might suggest new ways to treat or even prevent schizophrenia (Pearlson, 2000).

However, our understanding of schizophrenia is far from complete. What is known is that schizophrenia is a physical disease. Structural brain abnormalities, genetics, and neurochemistry are key components, and virtually every brain area and major neurotransmitter system has been implicated (Pearlson and Marsh, 1999; Bogerts et al., 2009; Glenthoj

et al., 2009). This wealth of experimental findings has led to many hypotheses regarding the causes of schizophrenia, but not to much clarity about its precise nature.

After a century of research, why is schizophrenia is not yet better understood? One reason is certainly that schizophrenia is both complex and heterogeneous, making the question of cause and effect extremely difficult. Furthermore, experimental techniques often cannot establish causality.

However, a more basic obstacle (we argue) is the lack of an adequate formal language in which to express hypotheses. In other words, new and advanced ways to investigate the schizophrenic brain experimentally should be complemented by new and equally powerful theoretical tools. Contributing to the development of these tools is the principal motivation behind this research.

Approach

This paper aims to demonstrate that connectionist models have the potential to provide these tools. Even though they are not usually intended as physiologically accurate simulations, they nevertheless tend to exhibit characteristics of "brain-like" information processing, and high-level cognition can thus be modeled using plausible analogs of the real neural substrate. A central working hypothesis is that neural networks not only function in a brain-like manner, but can also break down in the same way, creating an opportunity to advance our understanding of both the healthy and the disordered brain.

Computational neural networks have been used previously to simulate research findings related to schizophrenia, including altered working memory (Braver et al., 1999; Cohen and Servan-Schreiber, 1992; Monchi et al., 2000), hyperarousal states (Grossberg and Pepe, 1970; Grossberg, 1999), excessive semantic priming (Spitzer, 1997), alterations of functional connectivity between brain regions (Winder et al., 2007), attention (Wang and Fan, 2007), impaired facial affect recognition (Carter and Neufeld, 2007), and hallucinations and delusions (Hoffman and McGlashan, 1997; Ruppin et al., 1996; Loh et al., 2007). This work is the first, however, to use a model of human storytelling, and therefore the first where candidate illness mechanisms can be evaluated on a clinically relevant level.

The model, called DISCERN (Miikkulainen, 1993), is a complex, multi-modular neural network simulation of human story learning and memory. In this work, it was extended to model complex stories consisting of multiple scripts, as well as emotions and a cognitive filter function (Fidelman et al., 2005; Grasemann et al., 2007). These extensions make it possible to model the complex changes in language behavior observed in patients with schizophrenia.

The DISCERN Model

DISCERN processes stories via a chain of modules, each building on the results of the previous module and providing input for the next (Fig. 1A). The modules communicate using patterns of neuron activations that represent word meanings. Word representations are stored in a central semantic memory, and are learned based on their roles in sentences using the FGREP algorithm (Miikkulainen, 1993). Semantically similar words are reflected by similar activation patterns.

Based on a corpus of 28 stories, DISCERN first learned a lexicon of word representations in this manner. The other

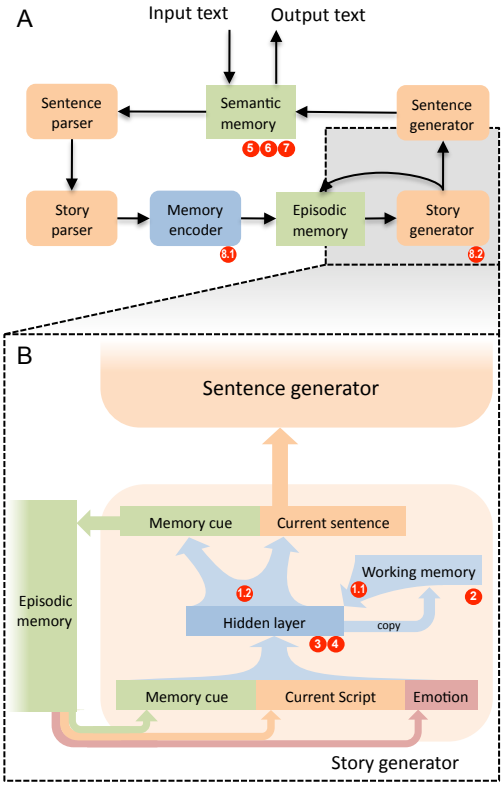

Figure 1: A schematic representation of the DISCERN story processing architecture. (A) Remembering and reproducing a story in DISCERN is achieved by a chain of modules. Tan modules are simple recurrent networks (Elman, 1990), the memory encoder is a RAAM network (Pollack, 1990), and the green modules are content-addressable memory modules. (B) The story generator module shown in more detail. A new memory cue is produced together with each sentence, making the transitions from one script to the next explicit. The numbers indicate implementations of the eight simulated illness mechanisms evaluated as models of schizophrenic language.

DISCERN modules were then trained in their tasks, i.e. understanding and paraphrasing the stories.

For each story, the words are presented to the sentence parser one at a time as a sequence of activation patterns. The sentence parser builds a representation of each sentence by concatenating the word representations that correspond to agent, predicate, indirect object, modifier, and direct object. At the end of each sentence, the sentence representation is passed on to the story parser. The story parser transforms sequences of sentences into script representations. Such a representation consists of the name of the script and the words and emotion filling its slots. The sequence of script representations that constitute the story is stored in the episodic memory module in a compressed form, which is created by the memory encoder using the Recursive Auto-Associative Memory (RAAM; Pollack 1990) architecture.

Story recall reverses the process, transforming episodic memories back into the original words: The story generator translates the episodic memory representation into a sequence of sentences, and the sentence generator then produces the word sequence for each sentence. Additionally, the story generator produces a memory cue with each sentence, thus making the script transitions explicit and allowing the system to process multi-script stories (Fig. 1B). Based on evidence of an editor function in human speech (Fox Tree, 2000), an out-

put sentence filter was incorporated into the story generator, reducing errors at the cost of reduced successful recall.

DISCERN learned fourteen autobiographical stories describing a first-person character ("I"), his relationships with his boss, girlfriend, etc., and activities such as going to a wedding and getting a traffic ticket. A second group of fourteen impersonal crime stories described Mafia and police characters and their activities. Each story consisted of a sequence of three to seven scripts whose slots were filled with a coherent set of words. An emotional valence code, ranging from very positive to very negative (++, +, +-, -, - -) was also included, biasing memory recall in a fashion analogous to human recall (Bower, 1981). Script organization was shared between autobiographical and crime stories, but story characters were not. For instance, DISCERN learned the following crime story about Vito, a Mafia boss, consisting of $driving and $pulled-over scripts with neutral emotion (slot-fillers in brackets and underlined in the story):

```
$driving [Vito,LA,airport,scared,recklessly,+-]
Vito wanted to go to LA.
Vito entered his car.
Vito drove to the airport.
Vito was scared.
Vito drove recklessly.
$pulled-over [Vito,cop,arrested,murder,+-]
Vito was pulled-over by a cop.
The cop asked Vito for his license.
Vito gave his license to the cop.
The cop checked the license.
The cop arrested Vito for murder.
```

An autobiographical story incorporated the same scripts with different slot-fillers and a strongly negative emotion:

```
$driving[I,home,drunk,home,recklessly,--]
I wanted to go home.
I entered my car.
I drove home.
I was drunk.
I drove recklessly.
$pulled-over [I,cop,arrested,DUI,--]
I was pulled-over by a cop.
The cop asked me for my license.
I gave my license to the cop.
The cop checked the license.
The cop arrested me for DUI.
```

Such common structure provides opportunities for schizophrenia-like language behavior to emerge, including derailments and delusion-like narratives.

Starting from different initial weights, 30 independent DISCERN systems were trained using baclpropagation. After training, average sentence-level recall was 95.6% (SD 0.8%). The "healthy" DISCERN systems were then used to simulate a range of candidate illness mechanisms:

1. *Working memory (WM) disconnection*, prompted by neuroimaging and other studies suggesting cortical disconnection, (especially involving WM networks) in schizophrenia patients (Kim et al., 2003; Karlsgodt, 2008; Sakkalis et al., 2006)). Implemented in DISCERN by cutting context-to-hidden-layer connections in the story generator (SG);
2. *WM noise*, prompted by reports of excessive cortical noise, and reduced efficiency in frontal cortical systems linked to WM in schizophrenia (Potkin et al., 2009; Tan et al., 2007; Winterer and Weinberger, 2004). Simulated by adding Gaussian noise to the context layer of the SG network;
3. *WM gain reduction* has been used previously to simulate hypo-dopaminergic neuromodulation of cortical networks in schizophrenia (Cohen and Servan-Schreiber, 1992). Implemented in DISCERN by changing the slope of the sigmoid activation functions in the SG's hidden layer;
4. *Excessive arousal* could produce both the under- and over-activation at a neuronal level seen in schizophrenia (Kuperberg et al., 2007; Loh et al., 2007), and was the basis of one of the earliest models of schizophrenia (Grossberg and Pepe, 1970). Simulated by increasing hidden layer activations in the SG;
5. *Semantic blurring*, based on the hypothesis that excessive spreading activation in semantic maps contributes to disorganized speech in schizophrenia (Spitzer, 1997; Leeson et al., 2005). Simulated by combining input word representations with their neighbors in the lexicon;
6. *Semantic memory noise*, intended to simulate disorganization of semantic memory suggested by studies demonstrating altered word association and fluency in schizophrenia (Goldberg et al., 1998; Tallent et al., 2001). Implemented by degrading semantic memory with Gaussian noise;
7. *Semantic memory overactivation*, motivated by neuroimaging studies suggesting increased cortical activation during semantic association tasks (Kuperberg et al., 2007; Assaf et al., 2006). Implemented by adding a constant bias to the lexicon's output;
8. *Hyperlearning*, a simulation of aberrant memory consolidation, intended to model memory effects of dopamine-driven, pathologically intense experience (Kapur, 2003; Kapur et al., 2005; Maher, 1974). Simulated by performing additional network training at increased learning rates; applied to the memory encoder and/or the generator modules (i.e. story and sentence generators).

The amount of distortion or damage introduced by an illness model (or *lesion*) can be controlled in each case by a "severity" parameter. Fig. 1 illustrates the location of the lesions in DISCERN. The lesions each result in characteristic language behavior that can be compared to that of schizophrenic patients.

Experiment I: Matching Human Data

The goal of the first experiment is to perform a rigorous and quantitative comparison of DISCERN and human subjects. Its foundation is a study of human story recall in schizophrenia, conducted at Yale as part of a joint project with this work (Hoffman et al., 2011). In the study, 20 healthy controls and 37 medicated outpatients with schizophrenia recalled three brief, thematically related stories. Several variables measuring different aspects of language and memory performance were scored, resulting in a unique quantitative characterization of language behavior in schizophrenia.

Four measures were used in the comparison: (1) recall success, measuring overall story recall performance; (2) derailment errors, measuring difficulty in maintaining a consistent storyline; (3) agent-slotting errors, measuring delusion-like confusions among story characters; and (4) lexical misfire errors, measuring difficulty in slotting words other than agents. The errors were represented as penetrance scores, i.e. divided by the number of error-free outputs. Patients scored lower on recall success, and produced more derailments and agent-slotting errors than controls. Lexical misfires did not differentiate patients from controls.

To measure how well the language profile of DISCERN simulations matched the experimental data, a novel mean-square deviation metric (Marchiori and Warglien, 2008) based on the four outcome variables was developed, measuring the goodness-of-fit (GOF) to human patients and controls. Lesion damage and output filtering were adjusted to optimize

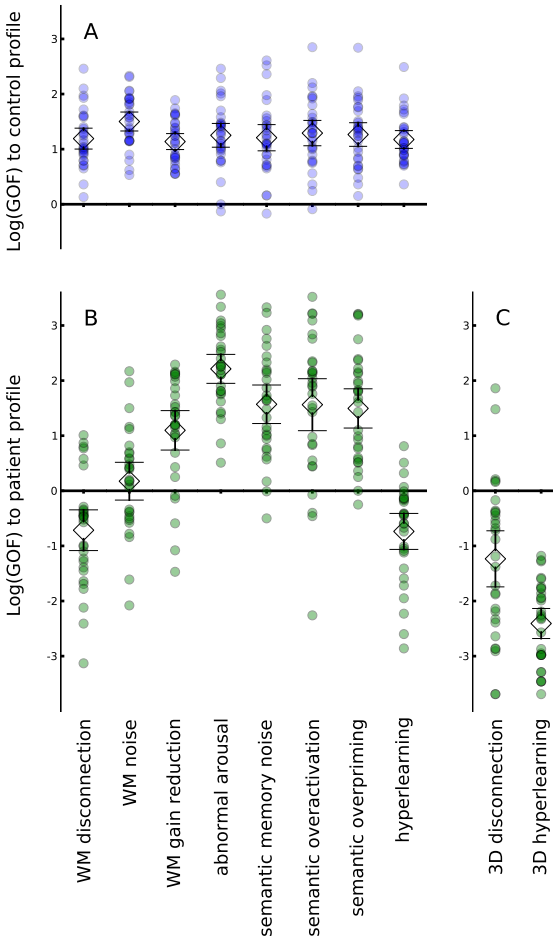

Figure 2: Goodness-of-fit of each illness model to human story-recall data. GOF is measured using a mean-square deviation metric to the language profile of controls (A) and patients (B,C). GOF was log-converted to normalize distributions; smaller values represent a better fit. Illness models are equivalent in matching controls. However, hyperlearning and WM-disconnection fit the patients better than other models. (C) Applying disconnection and hyperlearning in an additional location in DISCERN improves GOF to patients. Hyperlearning fits significantly better than WM-disconnection, suggesting a more likely illness mechanism of schizophrenia.

GOF for each of the eight illness mechanisms applied to each of the 30 independent DISCERN systems.

A mixed-model analysis with best-fit GOF as the response variable revealed no significant difference between lesions in matching healthy controls ($p = 0.07$; Fig. 2A). In contrast, the eight illness models differed significantly in how well they matched the patients' story-recall performance ($p < 0.0001$): WM-disconnection and hyperlearning were robustly superior to the other six models ($p < 0.0005$) but were not different from each other (Fig. 2B).

The two best models, WM-disconnection and hyperlearning, were further studied by applying each lesion in an additional location in DISCERN: hyperlearning was now applied to both smemory encoder and generator modules, and disconnection was applied to both context-to-hidden and output connections of the story generator. As a result, GOF to patients improved overall ($p < 0.0001$), and hyperlearning produced a significantly better fit than disconnection ($p < 0.0001$).

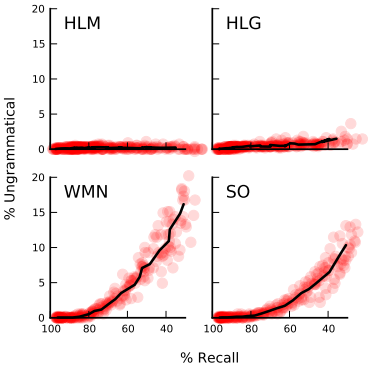

Figure 3: Percentages of ungrammatical sentences caused by four representative lesions is shown as recall performance drops in response to increasing damage. Solid lines are averages over ten DISCERN systems. Dots are individual lesioned systems. All lesions except hyperlearning (HLM, HLG) produce frequent grammar errors, which are not usually seen in schizophrenia.

The main result is that hyperlearning can match the story-recall profile of patients better than other illness models. That all models were equivalent in matching healthy controls suggests that hyperlearning captures specific aspects of schizophrenia, rather than general sources of human error-proneness.

Experiment II: Psychotic Language

Whereas the human subject study (necessarily) focused on medicated outpatients, psychosis is the hallmark of schizophrenia (Kapur, 2003). Its manifestations are the most distinctive signs of schizophrenia, and tend to dominate the early stages where models of intervention could be most useful. The second experiment consequently aimed to recreate language characteristic of acute psychosis, and to evaluate the ability of the illness models to do so.

Ten DISCERN systems were subjected to varying degrees of damage using the same eight lesions investigated earlier. Additionally, hyperlearning was applied to DISCERN's generator modules, for a total of nine lesions. The main goal was to determine whether any lesions are able to produce language abnormalities suggestive of (and consistent with) the psychotic symptoms of schizophrenia. First, syntax and morphology in schizophrenia tend to be normal or nearly so (Covington et al., 2005). However, most lesions (i.e. all except the hyperlearning models) cause frequent grammatical errors (Fig. 3). For example, the following was produced by DISCERN after WM-disconnection (target words in parentheses):

```
The Police thought that *the(Tony) bombed *airport(City-Hall).
The Police wanted to arrest *the(Tony).
The Police found that *to(Tony) *LA(was) *St-Mary's(in).
The Police planned to arrest *I(Tony) in New-York.
```

Distorted constructions like "The Police found that to LA St-Mary's" make the text appear random and non-sensical rather than disorganized or delusional, which is typically not the case in schizophrenia. In contrast, hyperlearning causes grammatical errors very rarely. Furthermore, hyperlearning applied to generator modules (HLG) causes consistent patterns of agent-slotting errors to emerge:

```
*I(Tony) hated *my(his) job.
*I(Tony) was a bad gangster.
*I(Tony) wanted to go to City-Hall.
*I(Tony) entered his car.
*I(Tony) drove to City-Hall.
*I(Tony) was *on-time(scared).
*I(Tony) drove *recklessly(carefully).
*I(Tony) entered City-Hall for *wedding(bombing).
Tony bombed City-Hall.
```

Such stable error patterns are very frequent following HLG, and virtually absent in other lesions. As in the example above, grammar remains intact. Often, meaningful and locally consistent new narratives emerge, suggesting how

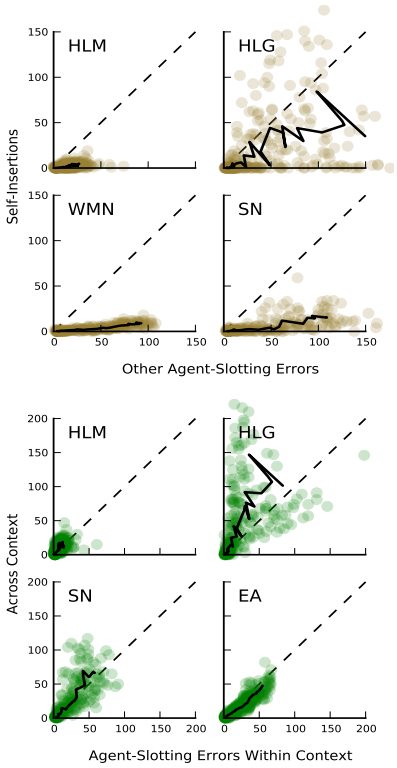

Figure 4: Types of agent-slotting errors for different illness models in DISCERN. Solid lines are averages at increasing levels of lesion severity. Top: Hyperlearning applied to generator modules (HLG) tends to insert the self-character into crime stories. Bottom: HLG frequently causes autobiographical characters to intrude across contexts into crime stories. Note that HLM, the other hyperlearning lesion, causes very few agent-slotting errors, suggesting fundamentally different language behavior.

agent-slotting errors could give rise to the spurious plots that characterize grandiose or persecutory delusions. Interestingly, in the human subject study, patients with narrative delusions made significantly more agent-slotting errors than other groups ($p = 0.015$, corrected post-hoc comparison, $\alpha = 0.05$), which supports this view of how delusions are formed. Furthermore, agent-slotting errors tend to intrude into crime stories and often involve the "I"-character (Fig. 4), as is common in schizophrenia.

When hyperlearning is applied to the memory encoder (HLM), the resulting language is very different: DISCERN makes few agent-slotting errors, but frequently jumps from one story to another, creating the impression of disorganized speech:

```
I went to Four-Seasons.
I sat at a table.
I ordered wine.
I drank the wine.
I met *Stacy(Kate) at
Four-Seasons.
[jumping to story #16]
Stacy was in her 20s.
Stacy had a ponytail.
Stacy was from New-York.
```

Here, DISCERN jumps from a story about meeting his mother Kate to a similar story about his girlfriend Stacy. Such derailments are driven by content rather than random memory errors; as in the example above, they tend to involve another similar story, and are often "foreshadowed" by content intruding from that story. Other lesions (except HLG) often jump between stories as well, but only in HLM are derailments the dominant error pattern (Fig. 5), and thus create a plausible model of disorganized speech in schizophrenia.

Each version of hyperlearning thus simulates one psychotic symptom but not the other. Since delusions and derailments are two hallmark symptoms of (respectively) paranoid-type and disorganized-type schizophrenia, hyperlearing suggests how these clinical subtypes could emerge from a shared underlying brain mechanism occuring at two different locations.

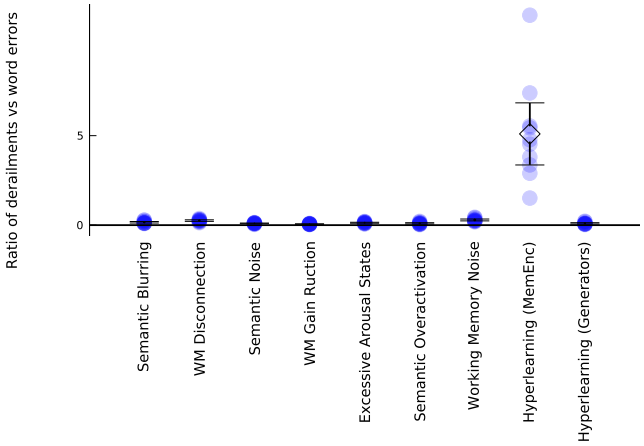

Figure 5: Frequency of derailed language vs. word-level errors. Lesion severity was adjusted so that recall was close to 40%. Hyperlearning causes frequent derailments and few word-level errors. In all other lesions, word-level errors dominate derailments.

Discussion and Future Work

The results show that DISCERN can indeed be used to simulate and compare alternative illness mechanisms in schizophrenia. Creating a computational framework where this is possible was one of the two main goals of this research. A more ambitious goal was to create an illness model that captures important aspects of schizophrenic language. The hyperlearning hypothesis is such a model.

In sum, hyperlearning can match the story-recall profile of human patients with schizophrenia (but not healthy controls) better than other models, and only the hyperlearning models produce compelling simulations of psychotic language. An account of delusion formation based on stable patterns of agency shifts is supported by human data showing that patients with narrative delusions make more agent-slotting errors (Hoffman et al., 2011). Applied to the memory encoder network, hyperlearning causes signs of language disorganization but not delusions. Hyperlearning thus models different symptoms depending on where in the model it is applied, suggesting how clinical subtypes of schizophrenia could emerge through a shared mechanism. Interestingly, although it is a new hypothesis, hyperlearning converges with a number of recent neurobiological findings linking schizophrenia, increased DA transmission and hippocampal activation, and elevated prediction-error response (Schott et al., 2006; Corlett, 2006; van Os and Kapur, 2009; Schobel et al., 2009).

The most important next step will be to investigate whether hyperlearning really occurs in humans, which should be possible with standard experimental techniques. Neural correlates of exaggerated memory consolidation are accessible to fMRI, and behavioral studies could investigate changes in the speed and intensity of memory consolidation. Further, since hyperlearning was able to simulate the language in both acute and compensated stages of schizophrenia, it may be able to model the transition between the two, which could lead to the first model of antipsychotic drug action.

More generally, hyperlearning serves to demonstrate the strengths of the computational modeling approach: First, hyperlearning was not part of the initial set of illness models, but instead emerged from preliminary experiments. In this

way, computational models can suggest novel, alternative hypotheses through unexpected behavior. Second, several predictions, like the existence of a shared illness mechanism for different clinical subtypes, were equally unexpected, demonstrating how computational models can tie together explanations of seemingly disparate symptoms. Third, hyperlearning demonstrates the conceptual reach of neural network modeling: It makes explicit, ties together, and converges with theories and research findings from psychology, psychiatry, linguistics, and theoretical and experimental neuroscience. In this way, connectionist models can make possible progress that would be difficult to achieve in any single discipline alone.

References

Assaf, M. et al. (2006). Abnormal object recall and anterior cingulate overactivation correlate with formal thought disorder in schizophrenia. *Biol Psychiat*, 59:452–9.

Bogerts, B., Steiner, J., and Bernstein, H. (2009). Brain abnormalities in schizophrenia. In Kasper, S., and Papadimitriou, G., editors, *Schizophrenia*, chapter 9. NY: Informa Healthcare, second edition.

Bower, G. (1981). Mood and memory. *Am Psychologist*, 36.

Braver, T., Barch, D., and J, C. (1999). Cognition and control in schizophrenia: A computational model of dopamine and prefrontal function. *Biol Psychiat*, 46:312–28.

Carter, J., and Neufeld, R. (2007). Cognitive processing of facial affect: Connectionist model of deviations in schizophrenia. *J Abnormal Psychol*, 116:90–305.

Cohen, J., and Servan-Schreiber, D. (1992). Context, cortex and dopamine: A connectionist approach to behaviour and biology in schizophrenia. *Psychol Review*, 99:45–77.

Corlett, P. e. a. (2006). Frontal responses during learning predict vulnerability to the psychotogenic effects of ketamine: linking cognition, brain activity, and psychosis. *Arch Gen Psychiat*, 63:611–21.

Covington, M. et al. (2005). Schizophr and the structure of language: The linguist's view. *Schizophr Res*, 77:85–98.

Elman, J. (1990). Finding structure in time. *Cog Sci*, 14:179–211.

Fidelman, P., Miikkulainen, R., and Hoffman, R. E. (2005). A subsymbolic model of complex story understanding. In *CogSci'05*.

Fox Tree, J. (2000). Coordinating spontaneous talk. In Wheeldon, L., editor, *Aspects of language production*, 375–406. New York, NY: Psychology Press.

Glenthoj, B., Christiansen, L., Rasmussen, H., and Oranje, B. (2009). Biochemical alterations in schizophrenia. In Kasper, S., and Papadimitriou, G., editors, *Schizophrenia*, chapter 11. New York: Informa Healthcare, second edition.

Goldberg, T., Aloia, M., Gourovitch, M., Missar, D., Pickar, D., and Weinberger, D. (1998). Cognitive substrates of thought disorder, I: the semantic system. *Am J Psychiat*, 155:1671–6.

Grasemann, U., Miikkulainen, R., and Hoffman, R. (2007). A subsymbolic model of language pathology in schizophrenia. In *Proc CogSci'07*. Nashville, Tennessee, USA.

Grossberg, S. (1999). How hallucinations may arise from brain mechanisms of learning, attention, and volition. *J Internat Neuropsychol Soc*, 6:579–588.

Grossberg, S., and Pepe, J. (1970). Schizophr: possible dependence of associational span, bowing and primacy versus recency on spiking threshold. *Behav Sci*, 15:359–62.

Hoffman, R., Grasemann, U., Gueorguieva, R., Quinlan, D., Lane, D., and Miikkulainen, R. (2011). Using computational patients to evaluate illness mechanisms in schizophrenia. *Biol Psychiat*. In press.

Hoffman, R., and McGlashan, T. (1997). Synaptic elimination, neurodevelopment and the mechanism of hallucinated 'voices' in schizophrenia. *Am J Psychiat*, 154:1683–1689.

Kapur, S. (2003). Psychosis as a state of aberrant salience: a framework linking biology, phenomenology, and pharmacology in schizophrenia. *Am J Psychiat*, 160(1):13–23.

Kapur, S., and Mamo, D. (2003). Half a century of antipsychotics and still a central role for dopamine D2 receptors. *Prog Neuropsychopharmacol Biol Psychiat*, 27(7):1081–90.

Kapur, S., Mizrahi, R., and Li, M. (2005). From dopamine to salience to psychosis – linking biology, pharmacology and phenomenology of psychosis. *Schizophr Res*, 79(1):919.

Karlsgodt, K. et al. (2008). Diffusion tensor imaging of the superior longitudinal fasciculus and working memory in recent-onset schizophrenia. *Biol Psychiat*, 63:512–8.

Kim, J. et al. (2003). Functional disconnection between the prefrontal and parietal cortices during working memory processing in schizophrenia: a [15(O)]H2O PET study. *Am J Psychiat*, 160:919–23.

Kuperberg, G. R., Deckersbach, T., Holt, D. J., Goff, D., and West, W. C. (2007). Increased temporal and prefrontal activity in response to semantic associations in schizophrenia. *Arch Gen Psychiat*, 64:138–151.

Leeson, V., Simpson, A., McKenna, P., and Laws, K. (2005). Executive inhibition and semantic association in schizophrenia. *Schizophr Res*, 74:61–7.

Loh, M., Rolls, E., and Deco, G. (2007). A dynamical systems hypothesis of schizophrenia. *PLOS Comput Biology*, 3:e228.

Maher, B. (1974). Delusional thinking and perceptual disorder. *J Indiv Psychol*, 30:98–113.

Marchiori, D. & Warglien, M. (2008). Predicting human interactive learning by regret-driven neural networks. *Science*, 319:1111–3.

Miikkulainen, R. (1993). *Subsymbolic Natural Language Processing*. Cambridge, MA: MIT Press.

Monchi, O., Taylor, T., and Dagher, A. (2000). A neural model of working memory processes in normal subjects, parkinson's disease and schizophrenia for fMRI design and predictions. *Neural Networks*, 13:953–73.

Pearlson, G. (2000). Neurobiology of schizophrenia. *Ann Neurol*, 48:556–66.

Pearlson, G., and Marsh, L. (1999). Structural brain imaging in schizophrenia: a selective review. *Biol Psych*, 46:627.

Pollack, J. (1990). Recursive distributed representations. *Art Intell*, 46(1):159–216.

Potkin, S. et al. (2009). Working memory and DLPFC inefficiency in schizophrenia: the FBIRN study. *Schizophr Bull*, 35:19–31.

Ruppin, E., Reggia, J., and Horn, D. (1996). Pathogenesis of schizophrenic delusions and hallucinations: A neural model. *Schizophr Bull*, 22:105–23.

Sakkalis, V. et al. (2006). Time-significant wavelet coherence for the evaluation of schizophrenic brain activity using a graph theory approach. In *Proc IEEE Eng in Med and Biol Soc*, volume 1, 4265–8.

Schobel, S. et al. (2009). Differential targeting of the CA1 subfield of the hippocampal formation by schizophrenia and related psychotic disorders. *Arch Gen Psychiat*, 66(9):938–46.

Schott, B. et al. (2006). The dopaminergic midbrain participates in human episodic memory formation: evidence from genetic imaging. *J Neurosci*, 26(5):1407–17.

Spitzer, M. (1997). A cognitive neuroscience view of schizophrenic thought disorder. *Schizophr Bull*, 23(1):29–50.

Tallent, K., Weinberger, D., and Goldberg, T. (2001). Associating semantic space abnormalities with formal thought disorder in schizophrenia: use of triadic comparisons. *J Clin & Exp Neuropsychol*, 23:285–96.

Tan, H., Callicott, J., and Weinberger, D. (2007). Dysfunctional and compensatory prefrontal cortical systems, genes and the pathogenesis of schizophrenia. *Cereb Cortex*, 17 Suppl 1:i171–81.

van Os, J., and Kapur, S. (2009). Schizophrenia. *Lancet*, 374:635.

Wang, H., and Fan, J. (2007). Human attentional networks: A connectionist model. *J Cogn Neurosci*, 19:1678–89.

Winder, R., Cortes, R., Reggia, J., and Tagamets, M. (2007). Functional connectivity in fMRI: A modeling approach for estimation and for relating to local circuits. *Neuroimage*, 34:1093–107.

Winterer, G., and Weinberger, D. (2004). Genes, dopamine and cortical signal-to-noise ratio in schizophrenia. *Tr Neurosci*, 27:683.