Body Word Count: 2499

Thoughts about Disordered Thinking:

Measuring and Quantifying the Laws of Order and Disorder

Brita Elvevåg[1, 2], Peter Foltz[3, 4], Mark Rosenstein[4], Ramon Ferrer-i-Cancho[5],

Simon De Deyne[6], Eduardo Mizraji[7] & Alex Cohen[8]

[1]Department of Clinical Medicine, University of Tromsø - The Arctic University of Norway, Norway
[2]Norwegian Centre for eHealth Research, University Hospital of North Norway, Tromsø, Norway
[3]Institute of Cognitive Science, University of Colorado, Boulder, Colorado, USA
[4]Pearson, Boulder, Colorado, USA
[5] Complexity and Quantitative Linguistics Lab, Departament de Ciències de la Computació, Universitat Politècnica de Catalunya, Barcelona, Spain
[6] Computational Cognitive Science Lab, School of Psychology, University of Adelaide, Australia
[7] Group of Cognitive Systems Modeling, Biophysics Section, Facultad de Ciencias, Universidad de la República, Montevideo, Uruguay
[8] Department of Psychology, Louisiana State University, Baton Rouge, Louisiana, USA

Submitted as an Invited Commentary to Schizophrenia Bulletin

Correspondence concerning this article should be sent to
Brita Elvevåg, Department of Clinical Medicine, University Hospital of North Norway – Asgård, Postbox 6124, Tromsø, Norway. Phone: (+47) 776-27753.
E-mail: brita@elvevaag.net
Key Words: Serious Mental Illness, Computational Semantics, Quantitative Linguistics, Network Theory

Thought disorder is a major component in the presenting phenomenology in serious mental illness and is important in differential diagnosis. Disturbances in the structure, organization, and coherence of thought are inferred from the reduced intelligibility and increased disorganization of speech that is difficult, and sometimes impossible, for the listener to comprehend. However, despite decades of characterizing and charting the nature and severity of disordered thought, the absence of strong theory-driven approaches about how order in language is tied to illness has limited its value in psychiatric research. The current status of clinical investigations of disordered thought is reviewed by Hart and Lewine [this issue][1] and Sass and Parnas describe the phenomenology [this issue][2]. Current investigative approaches for patient level studies are discussed in a commentary by Cohen et al [this issue][3]. To study the fundamental mechanisms involved in psychopathology the field often employs animal models. This is challenging for thought disorder pathology which is manifest in language. In this commentary we describe three statistical and mathematical approaches to linguistics and semantics - (i) computational natural language, (ii) complex networks and (iii) quantitative linguistics - that address fundamentals of speech communication that may be applied in preclinical investigations of mechanisms of thought disorder.

When considering animal studies, a promising cross-disciplinary approach to operationalizing and interpreting non-human vocal sequences across diverse taxa defines ‘information’ content and ‘meaning’ such that the signals can be examined mathematically to determine the level of structure (entropy), where information theory can be used to determine and assess various transmission and reception errors.[4] However, it is challenging to investigate mechanisms of thought disorder in animals. There are animal models of receptive language impairments and those that display putative erratic or disorganized behavior. However, the case for expressive, semantic-level deficits (i.e., disorganized speech) is quite controversial and the limitations of these models apparent. Such studies often examine whether various manipulations of the sound sequences have meaning to the listening animals[5], while other methods induce changes in vocalizations. For example, a FOXP2 mutation implicated in human speech deficits has been shown to alter the sequencing of mice vocalizations.[6] However, it remains to be established whether such putative errors (transmission and reception) can be characterized further, specifically whether the manipulation resulted in a 'simple' syntactical mixing up of sequences, or resulted in frank changes in information value transmitted. Questions remain whether these are now meaningless sequences, are sequences that mean something different than 'intended,' and so on. Though there is currently an absence of compelling animal models of thought disorder, in this commentary we suggest broadening the information theoretic approaches taken in these studies as well advocating for language specific methods of study that are directly applicable to humans in preclinical investigations.

Language affords a spectacular window into the brain and its pathologies. This is not to say that thought disorder is some sort of speech disorder, but rather that the verbal expressions of a patient permit us a unique lens on this extraordinary manifestation of rich and complex inner thought processes in humans. Language 'distortion' is a sign - in the medical sense - that is potentially measurable but currently without a universally accepted measure. An analogy to the equally non-invasive thermometer is that language provides an index into processes inside the body. An abnormal temperature from a cold does not indicate a disease of temperature regulation, but rather provides an indirect pathway toward measuring the internal processes contributing to the observed deviation. In the case of language measures need to be established and calibrated. Our rapidly evolving high tech era with the availability of digitalization of nearly all aspects of written and spoken discourse when combined with current conceptual, mathematical and computational language instruments represents both a challenge and an enormous opportunity. We present three sets of methods that induce order and structure on language and therefore enable the detection of subtle changes associated with the disorder and thus offer promise of revealing the relationship with underlying etiology.

Computational natural language methods

Many statistical natural language processing and associated machine learning applications have been successful at simulating human data and solving extremely difficult problems in the domains of speech (e.g., speech recognition) and language understanding, but have done so without simulating human cognition *per se*. Put differently, the success of these methods does not require that the algorithms are neurobiologically motivated. Several lexico-semantic modeling approaches have been adopted within cognitive neuroscience, notably latent semantic analysis (LSA), a natural language processing approach which derives language meaning from text corpora[7,8] by estimating the semantic relatedness of word sets as a function of the contexts in which they co-occur through the use of probabilistic inference or singular value decomposition. Other related techniques similarly generate distributed representations of language such as Topic Models,[9] Independent Component Analysis,[10] and Neural Networks, notably Deep Learning.[11] Thus, by analyzing very large sets of documents, the resulting statistical models allow us to operationalize the definitions of the structure of semantics so as to detect disordered thought. With this analysis, for any sets of texts it then becomes possible to assay how semantically similar the texts are. For instance, this similarity measure allows deriving the sentence to sentence coherence in a paragraph or determining how closely discourse turns semantically follow each other, thus facilitating comparisons across different populations or within an individual over time. Of note, these statistical-based tools are not simply agnostic assessment tools, rather, they can be used to test hypotheses and build theories and models. For example, LSA has been used to derive semantic coherence scores of discourse from patients with schizophrenia. Importantly these computationally derived metrics complement the traditional human clinical ratings of thought disorder but furthermore provide a framework to experimentally parse the incoherence in a theory-driven manner.[12] Some further successful examples of such approaches in the early stages of the diagnostic process use discourse to differentiate those at high risk of psychosis from unrelated (and presumed) healthy volunteers,[13] as well as predicting from those at risk who will eventually transition to psychosis.[14] It seems realistic to anticipate that natural language processing and machine learning approaches will provide the framework for establishing whether these computationally derived measures are sufficiently sensitive to monitor subtle but important clinical fluctuations.[15]

Thus far we have argued that computational natural language methods are valid methods for measuring the order of language and therefore for detecting the disorder as well. Additionally, the pioneering work of Hoffman tantalizingly suggests that in the near future computational simulations of patients are viable for pre-clinical purposes.[16] Computational methods can be applied to create speech patterns including modifications that mimic aspects of pathological speech. To the best of our knowledge, Hoffman built three decades ago the very first artificial intelligence model of disordered associations and applied this model to account for speech in mania and schizophrenia.[17] Now finally the time is ripe for this computational approach to gather momentum due to the convergence of three factors, namely (i) the availability of large corpora with which to model language; (ii) new computational algorithms that mimic inductive learning in a manner that is neurobiologically plausible (e.g., neural networks); (iii) and the availability of sufficient distributed computing power.[18-20] However, for such methods to be routinely adopted in research they require critical peer scrutiny in terms of how the comparison group or case is set up, reliable validation techniques, reproducible methods, and proper clinical calibration to link language features to underlying etiology.[15]

Complex networks

Statistics and the quantification of linguistic properties also play a central role in the new branch of network science that is being applied to linguistics,[21] such that human language is modeled as real-world networks.[22] Graph-theoretic analyses of linguistic and semantic networks are rapidly emerging as a useful framework for the semantic organization of concepts in memory. These can be used to explore individual differences in networks constructed from word association norms and text corpora either aggregated across responses of large samples of individuals or more sensitively by collecting associative network data within individuals studied over long periods of time.[23] Speech graphs have been demonstrated to provide useful quantitative measures of numerous features of disordered speech in patients with schizophrenia and mania that complement standard psychometric scales. Applications include measuring semantic coherence by computing disorganization;[24] representing specific thought features such as divergence and recurrence as graph measures;[25] graphically visualizing discourse as a trajectory in which the degree of disorder can be measured,[26] and also charting the flow of thought associations at an individual level in patients with serious mental illness as compared with a normative network[27]. These tools have also been applied to study the putative rigidity of thought in those with Asperger's Syndrome,[28] as well as to explore creative thinking,[29] and to chart spoken language in those administered MDMA ("ecstasy") and methamphetamine.[30] These network approaches to the mental lexicon extend the classical Collins and Loftus hierarchical taxonomic knowledge model[31] and since they include most words used in language result in an alternative structure that is primarily thematic (versus taxonomic). The sheer scale of these networks - combined with modern computational power - offers a radically new research framework for establishing the precise locus of aberrations in a *dynamic* network.[32] However, it remains to be seen if these methods and results can be reliably replicated.

Quantitative linguistics

Language is a communication method and the main goal of communicating is to convey information efficiently.[33] Based upon this premise, speech and its subsequent disorder can be operationally evaluated in terms of *'communication efficiency.'* [34,35] Such a phenotype affords a quantitative conceptualization. To understand a breakdown in communication it is useful to consider the evolutionary forces that drove the development of language, and then address this both at a phylogeny (i.e., species) and an ontogeny (i.e., developmental) level. Since language has not left fossils we have to infer how signals emerged to label objects and how the combinatorial nature of joining these labels resulted in a nearly infinite repertoire of sentences constructed from a finite number of words. Clearly, the capacity to combine elements to build phrases and sentences (i.e., syntax) was pivotal in language evolution, just as they are in children's language acquisition. An assumption is that a rich lexicon developed contingent upon a cognitive system that was able to exploit this combinatorial power by inducing rules based on exposure to language;[36] e.g., *This is a WUG. Now there is another one. There are two of them. There are two* ___ is an example that shows that children can apply the pluralization rule to words they do not know.[37]

Indisputably, statistical patterns inform us about important facets of language.[38] Quantitative linguistics concerns itself with these statistical properties, some that may be universals (i.e., present in all languages). Scaling laws are not unique to quantitative linguistics, but are found within a variety of cognitive processes.[39] One well-known example is Zipf's law,[34] which posits that the distribution of frequencies of words in any large corpus of language follows a power law such that the frequency of any word's occurrence is inversely proportional to the word's rank in the frequency table. For example, 150 of the most frequent words (e.g., 'the', 'of', 'and', etc.) account for almost half of the words used. The shape of this power function is consistent across languages to a large extent (i.e., universal). Zipf's law has been argued to index efficiency in information communication.[40] It has been shown that a simple word-object association matrix obeying Zipf's law - possibly a precondition for communication - produces syntactical rules as a by-product of a scale-free network architecture (i.e., 'syntax for free').[41] Such regular patterns are highly relevant to understanding the dynamic principles underlying language evolution, as well as how it develops and breaks down.[42]

In order to understand thought disorder, the level of analysis can benefit from considering language's evolutionary history across humanity: Quantitative linguistics can map these evolutionary traces and does so by uncovering statistical patterns and universals. Consider the role of phonological short-term memory which is likely crucial in language acquisition[43] and thus probably also in early language evolution. Dependency grammar specifies how words are linked in sentences, and shows that two words are linked if one depends syntactically on the other. Interestingly, in a sentence, the distance between syntactically linked words seems constrained with close to 90% of such relationships occurring at a distance of less than or equal to two words,[44] likely because of biological constraints (e.g., for purposes of breathing for utterance or memory for comprehension). Such cost minimization or least effort principles are found in a large range of universals in quantitative linguistics and general laws in nature.[45]

What implications does this set of regularities have for understanding disordered speech? If a grammatical error is made and the phrase "beautiful car black" is uttered (instead of “beautiful black car”), focusing on the error is of limited value. What is more informative (from an evolutionary perspective) would be to categorize the error in terms of path distance (i.e., shorter head-modifier dependencies for the erroneous phrase and longer for the correct phrase). Put differently, it is probably easier for our biological system to produce the erroneous one, but a peculiar English grammar rule alerts most of us that this would be incorrect. So although examining the type of error is useful, categorizing the speech error in terms of cost minimization can advance our understanding of thought disorder.[46] Importantly, these methods can be used to identify statistical patterns in language that are quite independent from the topic or speaker’s concerns and thus offer power tools to parse the core mechanisms underpinning thought disorder.

Summary and Conclusions

In response to *Schizophrenia Bulletin* articles by Lanin-Kettering and Harrow[47] and by Chaika and Lambe[48] on speech in patients with schizophrenia, Harrod[49] expressed discontent with the prevailing idea that many symptoms considered differential for schizophrenia, were in fact neither a disorder of thought nor linguistic composition but instead were indicative of problems in semiotics (see replies [50,51]). Three decades later Hart and Lewine [this issue][1] conclude their *Schizophrenia Bulletin* review on thought disorder by emphasizing the importance of "revisiting the significance of thought disorder as a core dimension in the study of psychosis" as "research in this area has the potential to elucidate robust etiological links, which, in turn, could inform individualized, effective intervention approaches." (p. xxx).[1] Our commentary agrees with their conclusion and we propose that measuring and quantifying the laws of order and disorder in speech will profit from adopting methods and theories from computational natural language, complex networks and quantitative linguistics to formally test predictions. Such approaches afford testable hypotheses and solutions to the problem raised by Harrod[49] by modeling word associations as well as episodic and semantic associative activations that link related representations in order to produce context-dependent and goal-oriented linguistic behavior.[26] Thus, the path forward is to use methods that are grounded in theory and which leverage modern computational techniques to induce regularities in linguistics, rule patterns, semantics and the structure of information being communicated, allowing us to better understand the etiology and nature of disordered thought.

Funding

This work was supported by a Research Council of Norway FRIPRO award (grant number 231395).

Acknowledgements

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

References

1. Hart M, Lewine RR. Rethinking Thought Disorder. *Schizophr Bull* 2017.

2. Sass L, Parnas J. Thought disorder, subjectivity, and the self. *Schizophr Bull* 2017.

3. Cohen AS, Le TP, Fedechko TL, Elvevåg B. Can RDoC help find order in thought disorder? *Schizophr Bull* 2017.

4. Kershenbaum A, Blumstein DT, Roch MA, et al. Acoustic sequences in non-human animals: A tutorial review and prospectus. *Biol Rev Camb Philos Soc* 2016;91:13-52.

5. Barnes TD, Wozniak DF, Gutierrez J, Han TU, Drayna D, Holy TE. A mutation associated with stuttering alters mouse pup ultrasonic vocalizations. *Curr Biol*. 2016;26(8):1009-1018.

6. Chabout J, Sarkar A, Patel SR, et al. A Foxp2 mutation implicated in human speech deficits alters sequencing of ultrasonic vocalizations in adult male mice. *Front Behav Neurosci*. 2016;10:197.

7. Deerwester SC, Dumais ST, Furnas GW, Harshman RA. Indexing by latent semantic analysis. *J Am Soc Inf Sci*. 1990;41(6):391-407.

8. Landauer TK, Foltz PW, Laham D. An introduction to latent semantic analysis. *Discourse Process*. 1998;25(2):259-284.

9. Blei DM, Ng AY, Jordan MI, Latent Dirichlet Allocation. *J Mach Learn Res*. 2003;3:993-1022.

10. Hyvärinen A, Karhunen J, Oja E. *Independent component analysis.* Vol 46: John Wiley & Sons; 2004.

11. Hinton GE, Osindero S, Teh Y-W. A fast learning algorithm for deep belief nets. *Neural Comput*. 2006;18(7):1527-1554.

12. Elvevåg B, Foltz PW, Weinberger DR, Goldberg TE. Quantifying incoherence in speech: an automated methodology and novel application to schizophrenia. *Schizophr Res*. 2007;93(1-3):304-316.

13. Rosenstein M, Foltz PW, DeLisi LE, Elvevåg B. Language as a biomarker in those at high-risk for psychosis. *Schizophr Res*. 2015;165(2-3):249-250.

14. Bedi G, Carrillo F, Cecchi GA, et al. Automated analysis of free speech predicts psychosis onset in high-risk youths. *NPJ Schizophr*. 2015;1:15030. doi:10.1038/npjschz.2015.30.

15. Foltz PW, Rosenstein M, Elvevåg B. Detecting clinically significant events through automated language analysis: Quo imus? *NPJ Schizophr*. 2016;2:15054. doi:10.1038/npjschz.2015.54.

16. Hoffman RE, Grasemann U, Gueorguieva R, Quinlan D, Lane D, Miikkulainen R. Using computational patients to evaluate illness mechanisms in schizophrenia. *Biol Psychiatry*. 2011;69(10):997-1005.

17. Hoffman RE. Computer simulations of neural information processing and the schizophrenia-mania dichotomy. *Arch Gen Psychiatry* 1987; 44: 178-188.

18. Corlett PR, Fletcher PC. Computational psychiatry: A Rosetta Stone linking the brain to mental illness. *Lancet Psychiatry*. 2014;1(5):399-402.

19. Friston KJ, Stephan KE, Montague R, Dolan RJ. Computational psychiatry: The brain as a phantastic organ. *Lancet Psychiatry*. 2014;1(2):148-158.

20. Wang XJ, Krystal JH. Computational psychiatry. *Neuron*. 2014;84(3):638-654.

21. Baronchelli A, Ferrer-i-Cancho R, Pastor-Satorras R, Chater N, Christiansen MH. Networks in cognitive science. *Trends Cogn Sci*. 2013;17(7):348-360.

22. Cong J, Liu H. Approaching human language with complex networks. *Phys Life Rev*. 2014;11(4):598-618.

23. Morais AS, Olsson H, Schooler LJ. Mapping the structure of semantic memory. *Cogn Sci* 2013;37:125–145.

24. Cabana Á, Valle-Lisboa JC, Elvevåg B, Mizraji E. Detecting order-disorder transitions in discourse: implications for schizophrenia. *Schizophr Res*. 2011;131(1-3):157-164.

25. Mota NB, Vasconcelos NAP, Lemos N, et al. Speech graphs provide a quantitative measure of thought disorder in psychosis. *PLoS One*. 2012;7(4):e34928. doi:10.1371/journal.pone.0034928.

26. Valle-Lisboa JC, Pomi A, Cabana Á, Elvevåg B, Mizraji E. A modular approach to language production: models and facts. *Cortex*. 2014;55(1):61-76.

27. De Deyne S, Elvevåg B, Hui C, Poon V, Chen E. Rich semantic networks applied to schizophrenia: A new framework. *Schizophr Res* 2016;176:454-455.

28. Kenett YN, Gold R, Faust M. The hyper-modular associative mind: A computational analysis of associative responses of persons with Asperger Syndrome. *Lang Speech*. 2016;59(3):297-317. doi:10.1177/0023830915589397.

29. Kenett YN, Anaki D, Faust M. Investigating the structure of semantic networks in low and high creative persons. *Front Hum Neurosci* 2014;8:407.

30. Bedi G, Cecchi GA, Slezak DF, Carrillo F, Sigman M, de Wit H. A window into the intoxicated mind? Speech as an index of psychoactive drug effects. *Neuropsychopharmacology* 2014;39:2340-2348.

31. Collins AM, Loftus EF. Spreading activation theory of semantic processing. *Psychol Rev*. 1975;82(6):407-428.

32. Elvevåg B, Cohen AS, Wolters MK, et al. An examination of the language construct in NIMH's research domain criteria: Time for reconceptualization! *Am J Med Genet B Neuropsychiatr Genet*. 2016;171(6):904-919. doi:10.1002/ajmg.b.32438.

33. Grice HP, White AR. Symposium: The causal theory of perception. *Proc Aristot Soc Suppl Vol*. 1961;35:121-168.

34. Zipf GK. *Human Behavior and the Principle of Least Effort.* Cambridge, MA: Addison-Wesley Press; 1949.

35. Ferrer-i-Cancho R, Hernández-Fernández A, Lusseau D, Agoramoorthy G, Hsu MJ, Semple S. Compression as a universal principle of animal behavior. *Cogn Sci*. 2013;37(8):1565-1578.

36. Bickerton D. *Language and species*. Chicago, IL: University of Chicago Press; 1990.

37. Berko J. A child's learning of English morphology. *Word.* 1958;14:150-177.

38. Ferrer-i-Cancho R, Solé R V. Least effort and the origins of scaling in human language. *Proc Natl Acad Sci U S A*. 2003;100(3):788-791.

39. Kello CT, Brown GDA, Ferrer-i-Cancho R, et al. Scaling laws in cognitive sciences. *Trends Cogn Sci*. 2010;14(5):223-232.

40. Powers DMW. Applications and explanations of Zipf’s law. *Proceedings of the joint conferences on new methods in language processing and computational natural language learning*. *ACL* 1998:151-160.

41. Ferrer-i-Cancho R, Riordan O, Bollobás B. The consequences of Zipf's law for syntax and symbolic reference. *Proceedings of the Royal Society of London B* 2005;272:561-565.

42. Ferrer-i-Cancho R. When language breaks into pieces: A conflict between communication through isolated signals and language. *BioSystems.* 2006;84:242-253.

43. Gathercole SE, Baddeley AD. Phonological memory deficits in language disordered children: Is there a causal connection? *J Mem Lang*. 1990;29(3):336-360.

44. Ferrer-i-Cancho R. Euclidean distance between syntactically linked words. *Phys Rev E* 2004;70:056135.

45. Gustison ML, Semple S, Ferrer-i-Cancho R, Bergman TJ. Gelada vocal sequences follow Menzerath's linguistic law. *Proc Natl Acad Sci U S A* 2016: 113 (19):E2750-2758.

46. Ferrer-i-Cancho R. The placement of the head that minimizes online memory. *Language Dynamics and Change.* 2015;5(1):114-137.

47. Lanin-Kettering I, Harrow M. The thought behind the words: a view of schizophrenic speech and thinking disorders. *Schizophr Bull* 1985;1:1-7.

48. Chaika E, Lambe R.The locus of dysfunction in schiozphrenic speech *Schizophr Bull* 1985:11;8-15.

49. Harrod JB. Schizophrenia as a semiotic disorder. *Schizophr Bull* 1986;12:12-13.

50. Chaika E, Lambe R. Is schizophrenia a semiotic disorder? Replies to Harrod. *Schizophr Bull* 1986;12:14-15.

51. Harrow, M., Prather, P., Lanin-Kettering, I. Is schizophrenia a semiotic disorder? Replies to Harrod. *Schizophr Bull* 1986:12; 15–19.