

Reduced Idea Density in Speech as an Indicator of Schizophrenia and Ketamine Intoxication

Cati Brown Michael A. Covington
University of Georgia

James Semple John Brown
GlaxoSmithKline Plc

Contact: catib@uga.edu

Special thanks to **Sara Weinstein** (U. of British Columbia) for use of her speech data and to **Congzhou He** (U. of Georgia) for help with data analysis.



Research supported by
GlaxoSmithKline
Research &
Development Ltd.

Overview

Idea density is the number of propositions expressed per 100 words (Kintsch 1974). Snowdon et al. (1996) found low idea density in written prose to be a strong predictor of Alzheimer's disease in later life. In this study, we **automate the measurement of idea density** and show that low idea density in speech is correlated with **schizophrenia** and with subanesthetic doses of **ketamine**, which is thought to produce a cognitive impairment similar to schizophrenia.

An idea or proposition is anything that can be true or false. Thus "The brown dog barked in the garden" contains three propositions (dog was brown, dog barked, it happened in the garden). Kintsch (1974) introduced propositional density (idea density) as a measure of text complexity.

Idea density corresponds closely to the ratio of verbs, adjectives, adverbs, prepositions, and conjunctions to the total number of words in a speech sample. Exploiting this fact, we developed a Computerized Idea Density Rater (CIDR), which estimates idea density through part-of-speech tagging.

We used CIDR to analyze two sets of volunteers' descriptions of pictures from the Thematic Apperception Test. A between-groups comparison of schizophrenics and healthy controls revealed that patients' speech tended to contain lower idea density. In a placebo-controlled, double-blind, cross-over study of healthy volunteers given subanesthetic doses of ketamine, idea density was generally lower in the ketamine condition, even though disordered speech was not, in general, evident to the listener.

These results indicate that the automated measurement of idea density is potentially useful for psychiatric diagnosis and accelerated drug development.

Ketamine Experiment

During a placebo-controlled, double-blind experiment at Cambridge University, healthy volunteers were asked to describe some pictures from the Thematic Apperception Test (TAT, Murray 1943/1971) after two different sub-anesthetic doses of ketamine and after placebo.

Each volunteer was recorded on 3 occasions at least a week apart. Volunteers were unaware which dose they were receiving unless they felt the effects of the drug.

This study's "low dose" and "high dose" are much smaller than the usual anesthetic dose, specifically:

Low dose = 0.08 mg/kg over 10 min + 0.04 mg/kg/hr x 90 min
High dose = 0.27 mg/kg over 10 min + 0.14 mg/kg/hr x 90 min
Usual anesthetic dose = 1 to 4.5 mg/kg (*Physicians' Drug Hbk.* 2003).

Speech was recorded 30 minutes into the experiment.

Schizophrenia Experiment

Descriptions of TAT pictures were also recorded at the University of British Columbia in an experiment conducted by Sara Weinstein, who kindly provided us with copies of the transcripts.

The subjects were 11 healthy controls and 11 patients with schizophrenia, matched for age, education, IQ and socio-economic status. Patients with schizophrenia were recruited from the Vancouver Hospital's Schizophrenia Day Program as well as through their primary care physicians. Control subjects were recruited with advertisements posted around the University of British Columbia campus.

Both sets of speech data were also used for other studies.

Computer Speech Analysis

Both sets of speech samples were transcribed by researchers who were unaware of each subject's condition. The samples were then analyzed by computer.

For this study, a computer program was written which we call CIDR (Computer Idea Density Rater). It approximates the number of ideas in a text by counting the number of verbs, adjectives, adverbs, prepositions, and subordinating conjunctions, each of which corresponds to one proposition.

CIDR is implemented in the Python programming language (www.python.org), and parts of speech are identified (tagged) using a shareware Python package called *MontyTagger* (part of the module *MontyLingua*, Liu 2004).

Results

Both ketamine and schizophrenia were associated with significantly reduced idea density.

The ketamine experiment yielded usable speech samples from 11 volunteers (1 lacking one of the three recordings). Almost all of them spoke with lower idea density with ketamine than with placebo, even though the dose was so small that no effects on speech were audible to the untrained ear. With the higher dose (still very low compared to anesthetic doses), the reduction was highly significant ($p=0.006$). The paired nature of the test (measuring each volunteer under 3 conditions) made it very sensitive.

The schizophrenia experiment was not an individually paired test, and its results are somewhat less clear, but idea density is nonetheless lower in schizophrenia than in healthy controls ($p=0.04$). As the histogram shows, although the means and ranges of idea density for the two groups were not far apart, the distributions were very different, a fact which should be explored with a larger sample.

Discussion

These results show that idea density of speech is a useful neuropsychological measurement and that the approximate method used in CIDR is accurate enough to reliably show differences between ketamine-influenced speech, schizophrenic speech, and normal speech.

Further, these results show that the effect of ketamine on idea density is similar to the effect of schizophrenia – both cause a small but distinct decrease.

We have not yet explored whether these decreases are *qualitatively* alike. In different situations, a high or low idea density can mean different things, ranging from stylistic choices to discourse planning ability and limits on short-term memory.

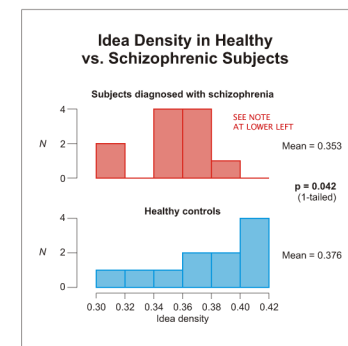
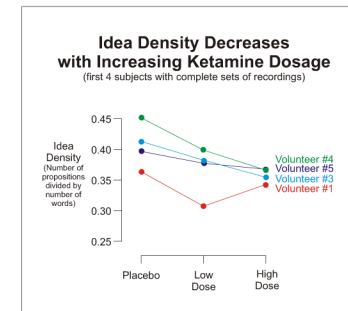
Because of the ease of making the measurement, CIDR is a useful tool for exploring the psychology of speech.

NOTE ADDED JULY 2005: Subsequent analysis of the same data set has failed to replicate the reduction of idea density in schizophrenia, and we suspect an error in handling the data. The reduction in ketamine intoxication is confirmed.

Idea density in ketamine experiment

Volunteer	Placebo	Low Dose	High Dose
1	0.364	0.307	0.344
2	0.402	0.365	
3	0.422	0.383	0.358
4	0.454	0.400	0.366
5	0.394	0.378	0.366
6	0.343	0.390	0.353
7	0.408	0.369	0.278
9	0.414	0.414	0.356
10	0.447	0.320	0.423
11	0.361	0.422	0.357
12	0.362	0.367	0.348

$p=$ 0.0835 0.0063
(one-tailed t-test, relative to placebo)



What are ideas and idea density?

Idea density = number of ideas / number of words where ideas (propositions) are whatever can be true or false.

Example:

When the dog barked, the old man heard it in the garden.

Propositions:

- (1) Dog barked
- (2) Man heard dog
- (3) Man was old
- (4) Proposition (2) occurred in garden
- (5) Proposition (2) occurred at same time as (1)

5 propositions, 12 words: idea density 5/12 = 0.42 (quite high).

Ideas correspond almost exactly to *verbs, adjectives, adverbs, prepositions, and conjunctions* (not nouns, pronouns, or determiners).

When the dog barked, the old man heard it in the garden.
(5) (1) (3) (2) (4)

This is a shortcut for determining idea density by computer.

Snowdon et al. (1996) found that Alzheimer's disease could be predicted from the idea density of prose written 50 years before onset of symptoms. We believe that idea density is a sensitive measure not only of linguistic style, but also of neuropsychological status.

References

- Kintsch, Walter (1974). *The representation of meaning in memory*. Hillsdale, N.J.: Erlbaum.
- Liu, Hugo (2004). *MontyLingua: An end-to-end natural language processor with common sense*. Available at: web.media.mit.edu/hugo/montylingua.
- Murray, H. A. (1971). *Thematic Apperception Test: Manual*. Cambridge, MA: Harvard University Press. (Original work published 1943.)
- Snowdon, D. A.; Kemper, S. J.; Mortimer, J. A.; Greiner, L. H.; Wekstein, D. R.; and Markesbery, W. R. (1996) Linguistic ability in early life and cognitive function and Alzheimer's disease in late life: findings from the Nun Study. *Journal of the American Medical Association* 275:528-532.