Discourse in Psychosis Working Group:

Clinical Data Harmonization

Vanteemar S. Sreeraj, Alban Voppel, Sunny Tang and Frederike Stein

Discourse in Psychosis Satellite Meeting
Pavia, Italy 2024

Working Group Objective & Members

- The key objective of this group is to facilitate cross-site and cross-linguistic clinical (i.e. human-rated) assessment of thought disorder and speech and language disturbance.

- As a key step in this process, we need a way of connecting clinical ratings that have been made using different scales.

Working Group Objective & Members

Members:

Sunny Tang: co-lead
Frederike Stein: co-lead
Rosa Ayesa Arriola
Zarina Bilgrami
Emre Bora
Sylvia Ciampelli
Anthony Deo
Tilo Kircher
Víctor Ortiz García de la Foz
Peter Liddle
Raffael Massuda
Natália Mota
Emre Mutlu
Lena Palaniyappan
Susan Rossell
Sarah Sales
Vanteemar Sreeraj
Katharina Stegmayer
Phil Sumner
G. Venkatasubramanian
Alban Voppel
Sebastian Walther
Berna Yalınçetin
…

Outline

Dr. Sreeraj V S

Current landscape of clinical rating scales related to formal thought disorder

Dr. Alban Voppel

Using semantic information from item descriptions to cluster formal thought disorder symptoms

Dr. Frederike Stein & Dr. Sunny Tang

Data-driven dimensions of formal thought disorder across scales

Clinical Characterization of Thought Disorder in Schizophrenia

VANTEEMAR S. SREERAJ , KIRAN BAGALI, SONIKA NICHENAMETLA, NILANJAN DUTTA, SWARNABUDDHA NAYOK, VENKATARAM SHIVAKUMAR, LENA PALANIYAPPAN, GANESAN VENKATASUBRAMANIAN

Dr. Sreeraj VS

Associate Professor of Psychiatry

National Institute of Mental Health and Neurosciences (NIMHANS),

Bengaluru, India

Identification of tools to assess FTD

Systematic survey: Pubmed and Psynet (APA)

("formal thought disorder" OR "thought disorder" OR ("disorg*" AND ("thought" OR "speech" OR "concept*" OR "cognitive"))) AND ("psychosis" OR "schizo*" OR "Bipolar Disorder" OR "Mania" OR "Affective disorder" OR "Depress*" OR "Anxiety" OR "Obsess*" OR "Personality Disorder")

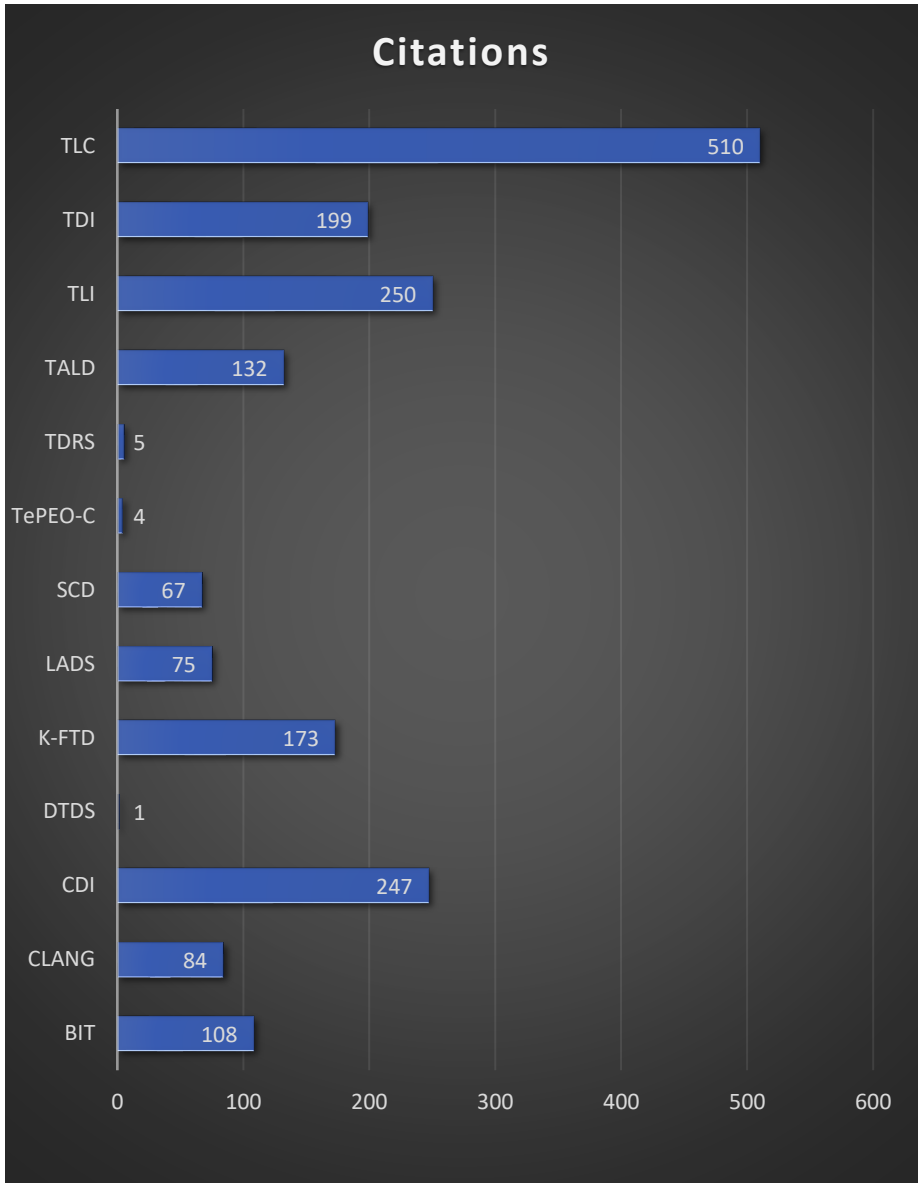
Filters: Abstract, Clinical Study, Clinical Trial, Corrected and Republished Article, Evaluation Study, Historical Article, Multicenter Study, Observational Study, Randomized Controlled Trial, Validation Study, English

In December 2023

FTD Specific Tools

Full Name	Author	Year	Scales	Citations
1.Bizzare Idiosyncratic Thinking	Marengo	1986	BIT	108
2.Clinical Language Disorder Rating Scale	Chen	1996	CLANG	84
3.Communication Disturbance Index	Docherty	1996	CDI	247
4.Dokuz Eylül University Thought Disorder Scale	*Yalicentin*	*2020*	*DTDS*	*1*
5.Kiddie FTD	Caplan	1989	K-FTD	173
6.Lossening of Association and Disordered Speech patterns	Reilly	1975	LADS	75
7.Schizophrenic Communication Disorders	*Bazin*	*2005*	*SCD*	*67*
8.Test Psychique Expérimental Opérationnalisé pour le diagnostic de Cataphasie	Mainberger	2015	TePEO-C	4
9.Thinking Dysfunction rating scale	Karasau	1979	TDRS	5
10.Thought and Language Disorder	Kircher	2014	TALD	132
11.Thought and Language Index	Liddle	2002	TLI	250
12.Thought Disorder Index	Johnston & Holtzman	1979	TDI	199
13. Thought Language and Communication	Andreasen	1979	TLC	510

Formal Thought Disorder- patient and caregiver versions	Barrera	2008	FTD-p/c	76
Cognitive Slippage Scale (CSS)	Miers and Raulin	1985	CSS	40
Communication Awareness Scale (CAS)	McGrath	2000	CAS	20
Whitaker Index of Schizophrenia Thinking (cognitive test)	Whitaker	1973	WIST	41
Communication Deviance for TAT (CDT)	Wynne	1963	CDT	270



Mapping of items

All items of any FTD specific tool

Relevant items of non-specific tools with >10 citations for assessing FTD (BPRS ~ PANSS, so only PANSS items chosen)

So, a total of 13 tools: 189 items were finally used

Item-Concept conflicts

Each item of each scale was mapped with other-item:

- Composite items: Single item of one scale mapped to two or more items concepts of other scales.

  e.g., PANSS conceptual disorganisation = Tangentiality + rupture of thought + derailment + incoherence + illogicality.

- Non-matching items: Same term used in 2 scales, but mapped to different concepts.

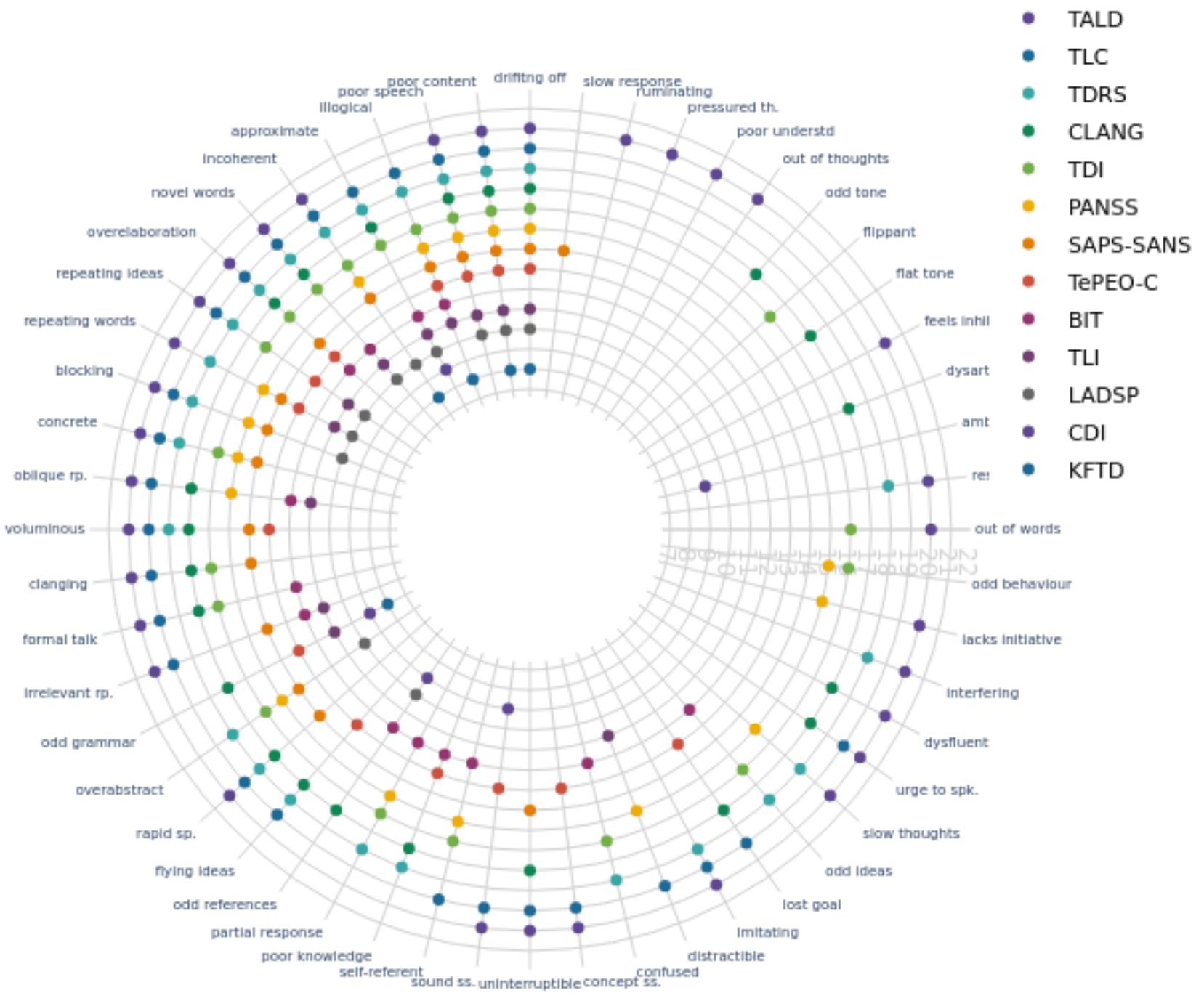
  e.g, Tangentiality of TLC ≠ Tangentiality of TALD

  But Tangentiality of TLC = Cross-talk of TALD

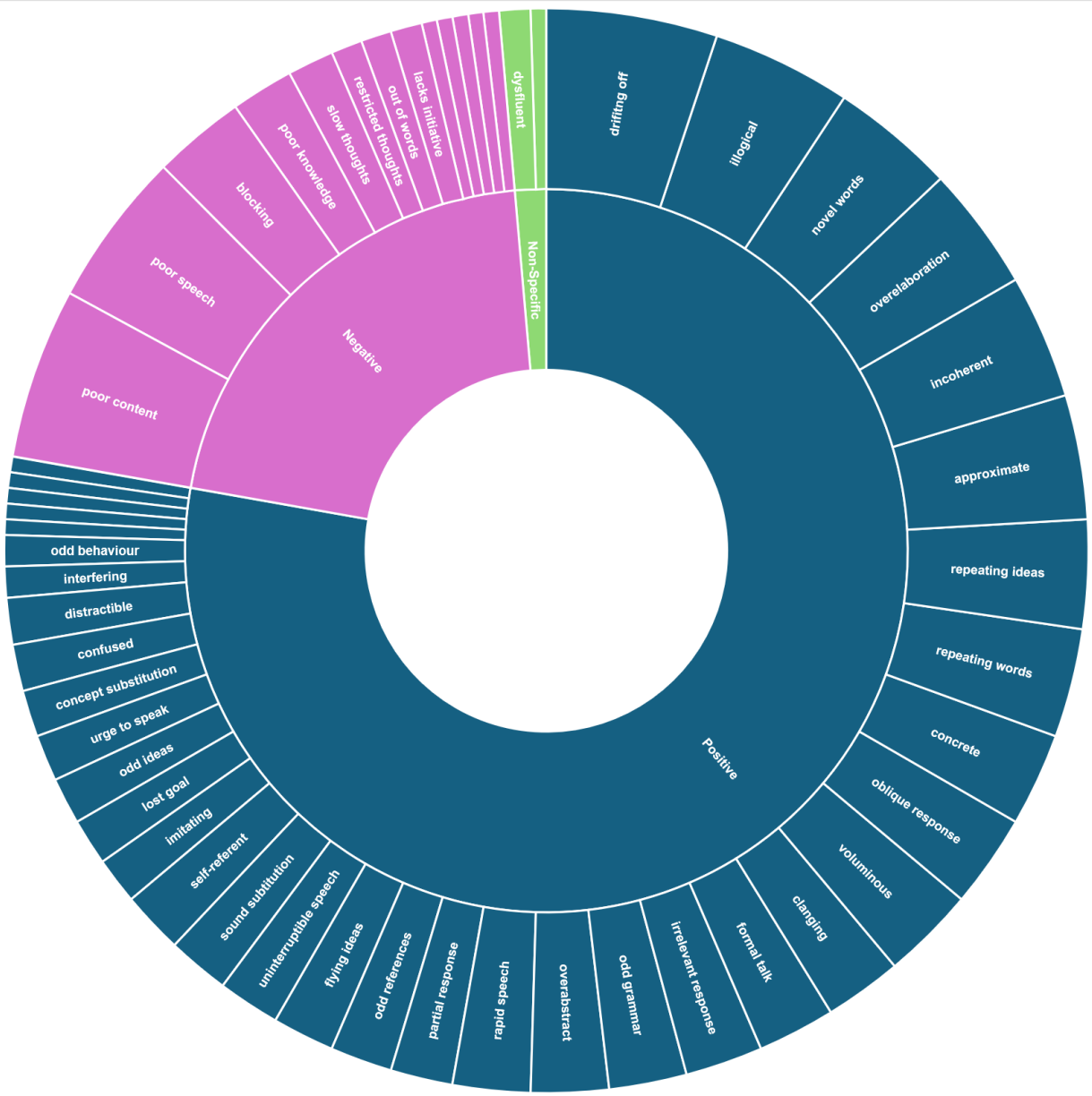
- Non-related items: Items which were considered not to be a FTD were removed

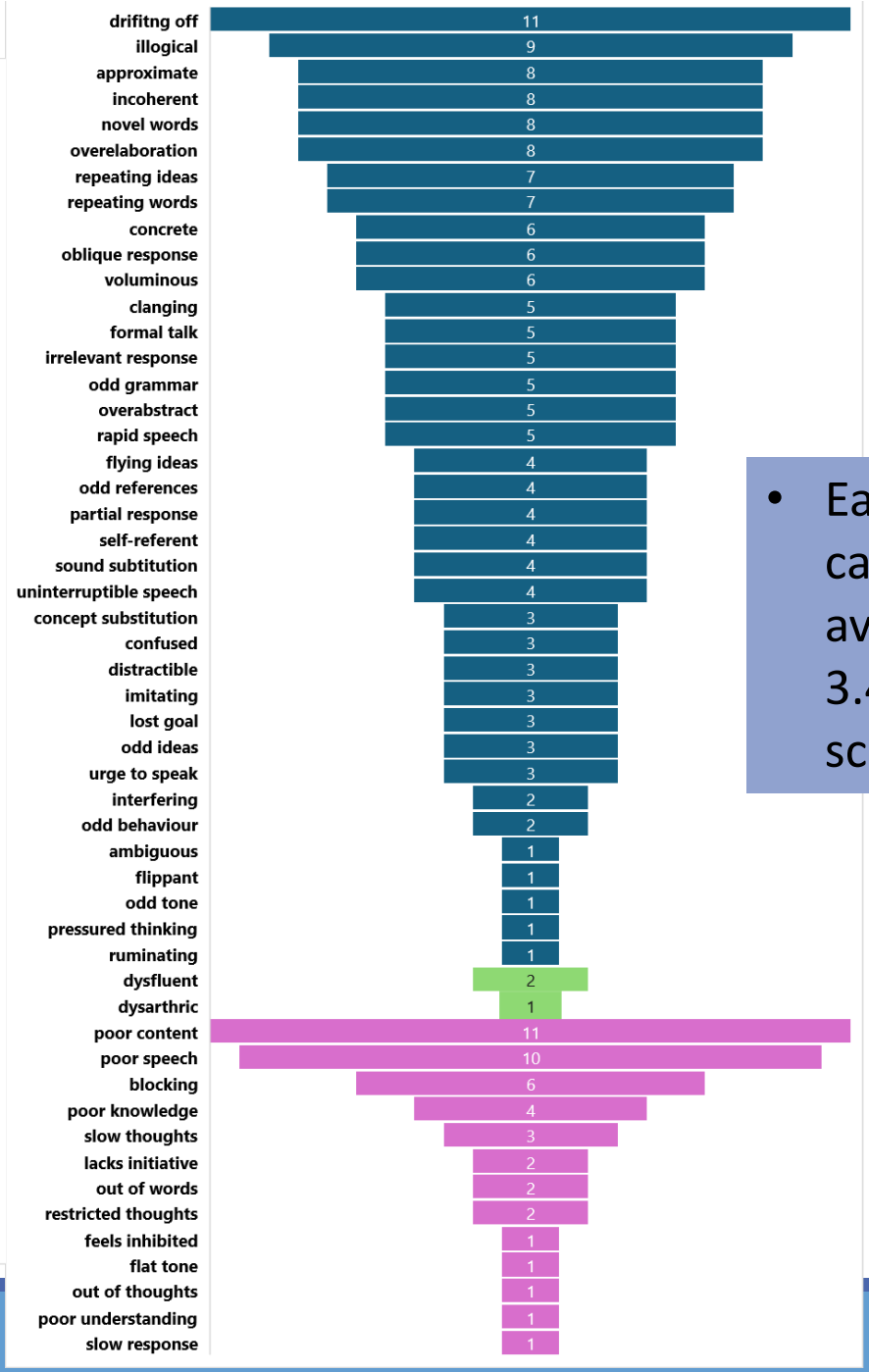
  e.g., Judgement and memory in TDRS

Representation of the features in the scales



- Drifting off & poverty of speech are covered by most scales (11/13)
- A total of 11 items are captured by only 1 of the 13 scales





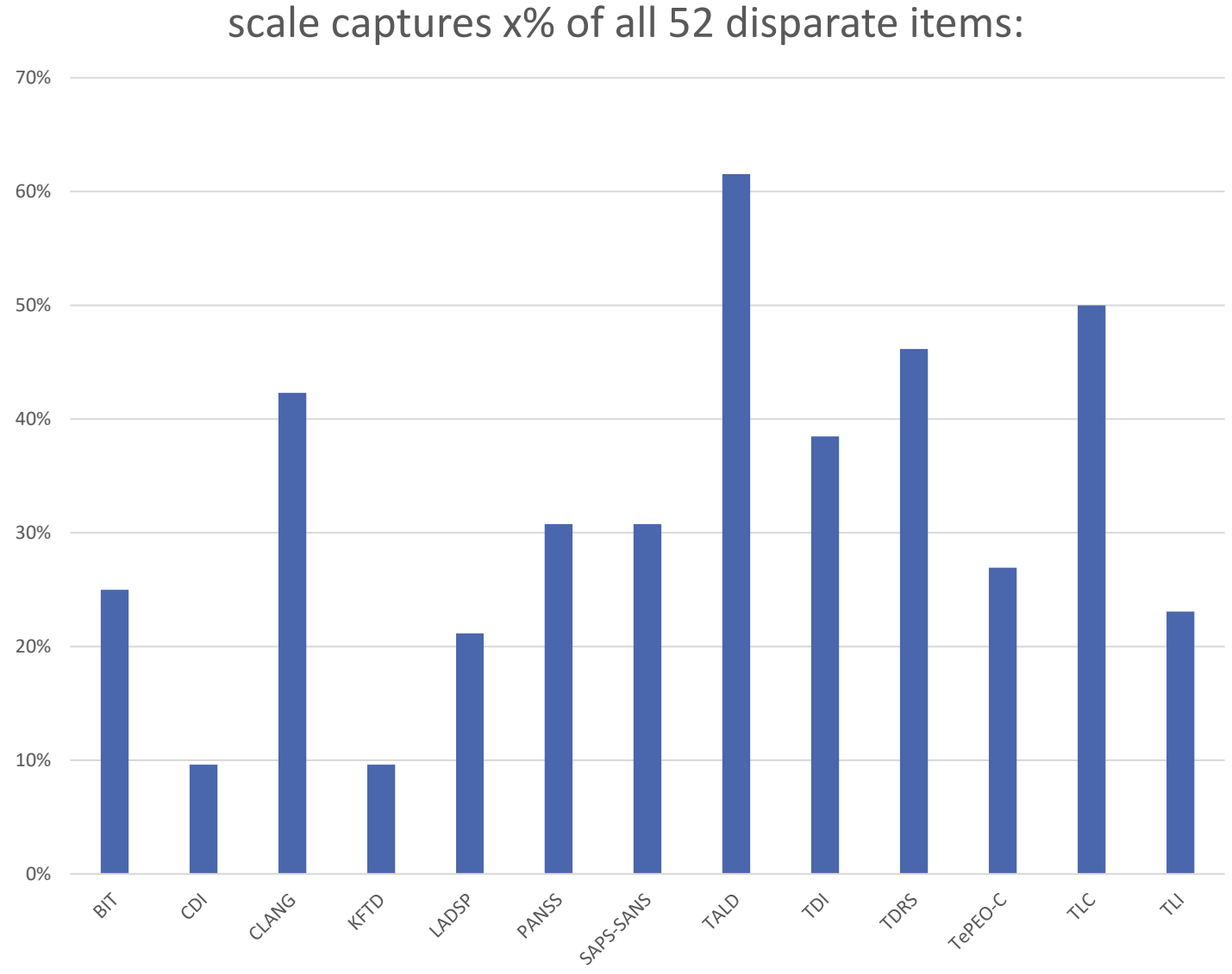
- Each item is captured by an average of only 3.42 (SD=2.2) scales

Scales' ability to quantify features

TALD captures 62% of features

TLC around 50% of features

SANS SAPS/ PANSS only 30% of features



How coherent are our rating scales?

Using Semantic Information to Cluster Formal Thought Disorder Rating Scales

Alban Voppel

Postdoctoral fellow

CEYMH, Douglas mental health hospital, McGill University

Centre of Excellence in Youth Mental Health
Centre d'excellence en santé mentale des jeunes

Douglas INSTITUT UNIVERSITAIRE EN SANTÉ MENTALE MENTAL HEALTH UNIVERSITY INSTITUTE

McGill UNIVERSITY

Douglas CENTRE DE RECHERCHE RESEARCH CENTRE

Centre intégré universitaire de santé et de services sociaux de l'Ouest-de-l'Île-de-Montréal
Québec

Quantitative features

7 scales

103 items

Subset of items for:

- SAPS-SANS
- PANSS

Scale	Average	Min Number of Words	Max Number of Words
BIT	29.2	6	94
CLANG	14.4	2	61
PANSS	37.5	26	53
SAPS-SANS	90.2	43	191
TALD	54.5	6	140
TLC	95.1	31	242
TLI	64	9	184

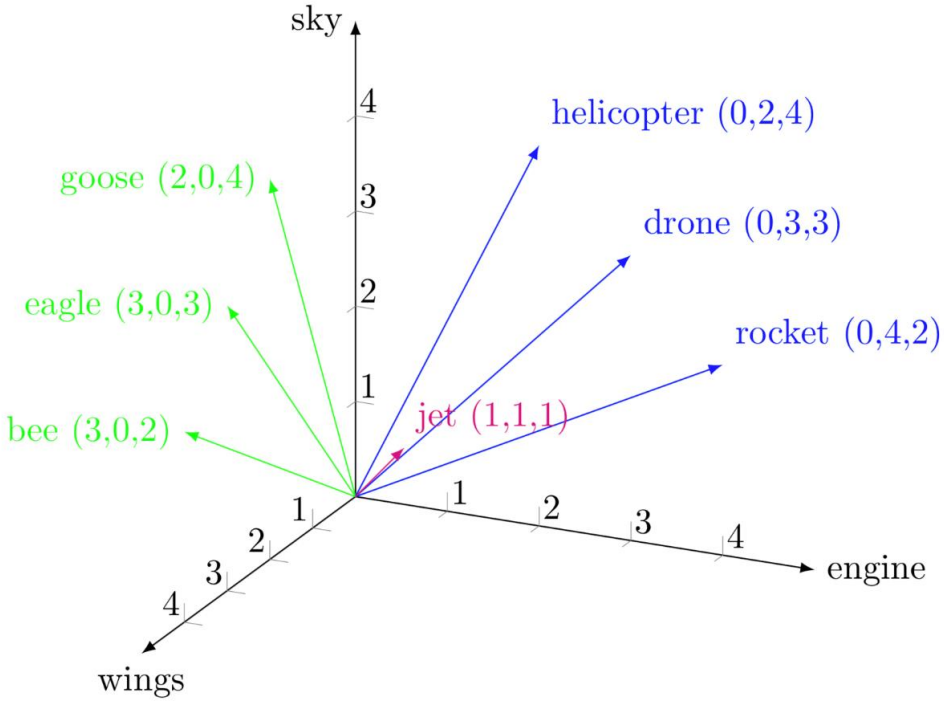
Our approach: use a language model

Language models learn semantic relations

Can be used to measure semantic similarity

[-5 key word +5]

N	Concordance
1	and continue to work with the definitions <w PRP>throughout the essay, relating the different definitions
2	The Dionysiac cult was established <w PRP>throughout Greece and a symbolic reconciliation with
3	t[illegible] dispensers, yet he did not utter a groan <w PRP>throughout the operation. Afterwards she found herself
4	the irony of that struck her. <s n="2414"><w PRP>Throughout the revising, to her surprise and her
5	to make the king their puppet continued <w PRP>throughout his minority. The queen mother remarried
6	and (iii) the trees that are shade-bearers <w PRP>throughout life and are incapable of rapid growth in the
7	Available worldwide THIS TEXT IS AVAILABLE THROUGHOUT THE WORLD only as part of the
8	this situation prevails in all industries <w PRP>throughout the economy (that is, price equals
9	of the most significant features of their lives <w PRP>throughout the disability career. Achieving this
10	Trades Union of February 1834. <s n="13"><w PRP>Throughout his radical and union career Doherty
11	Type C may have been present <w AV0>throughout — from the early to the late-fourth
12	11 Mediation Attempts <s n="1204"><w PRP>Throughout the war, there were constant efforts to
13	the Good News of Christ's love spreading <w PRP>throughout the world. The present imbalance in the
14	a feeling of continued need for defecation <w PRP>throughout the distension period. Only one subject
15	on account of the high-grade circuitry used <w AV0>throughout. The combo offers the ubiquitous passive



Our approach: use a language model

Authors describe their meaning in descriptions of items and scales

Language models can quantify this meaning across scales

TLC-9 **Clanging**: A pattern of speech in which sounds rather than meaningful relationships appear to govern word choice, so that the intelligibility of the speech is impaired and redundant words are introduced. In addition to rhyming relationships, this pattern of speech may also include punning associations, so that a word similar in sound brings in a new thought.

CLANG-1: **Excess phonetic association**: Abnormal association based on phonetic similarity (punning and clang associations)

BERT Sentence Similarity: 0.744

CLANG-9 **Aprosodic speech**: Flat monotonous speech without appropriate inflexion and emotional quality

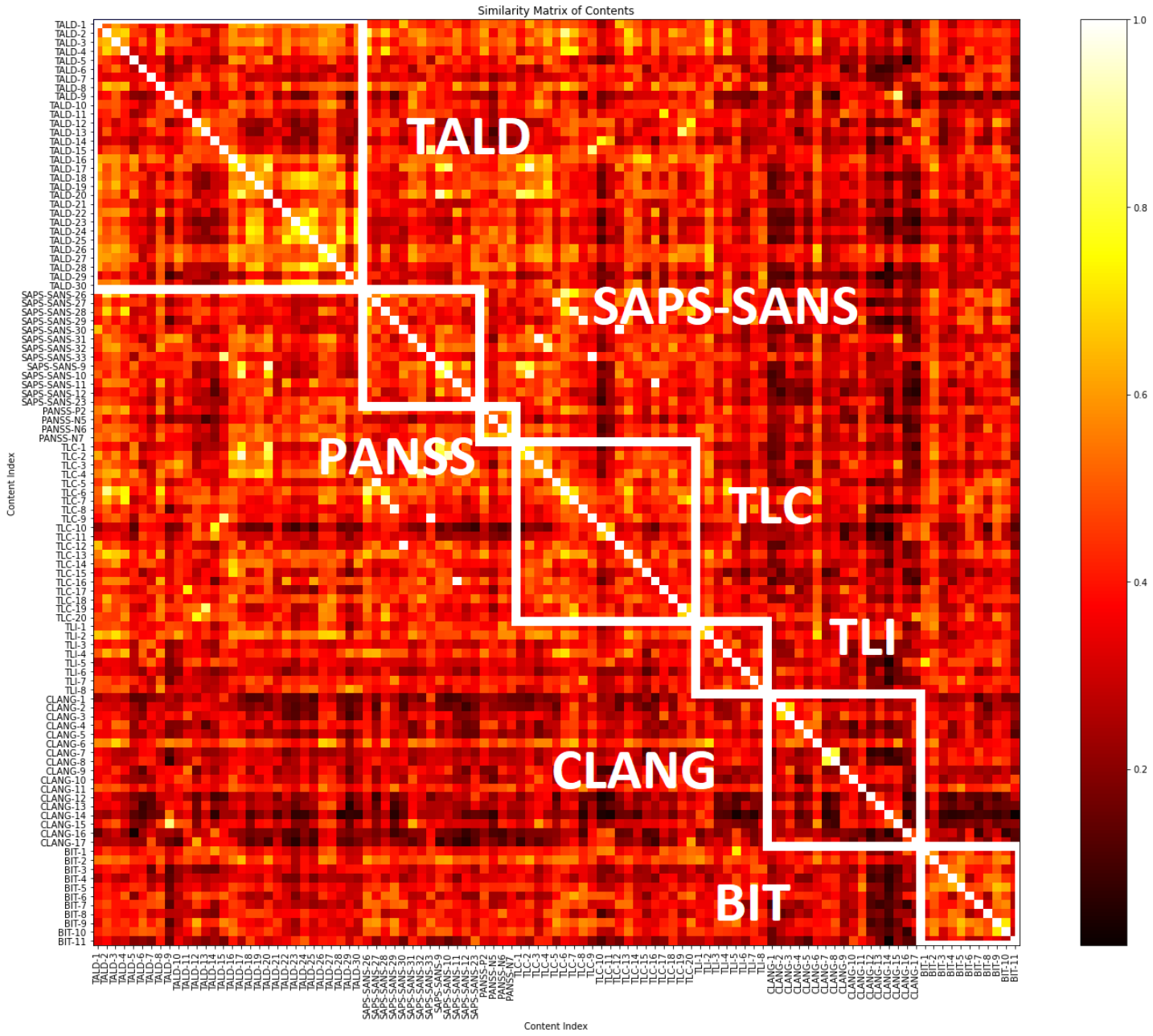
TLC-4: **Distractible speech**: During the course of a discussion or interview, the patient stops talking in the middle of a sentence or idea and changes the subject in response to a nearby stimulus.

BERT Sentence Similarity: 0.316

Workflow

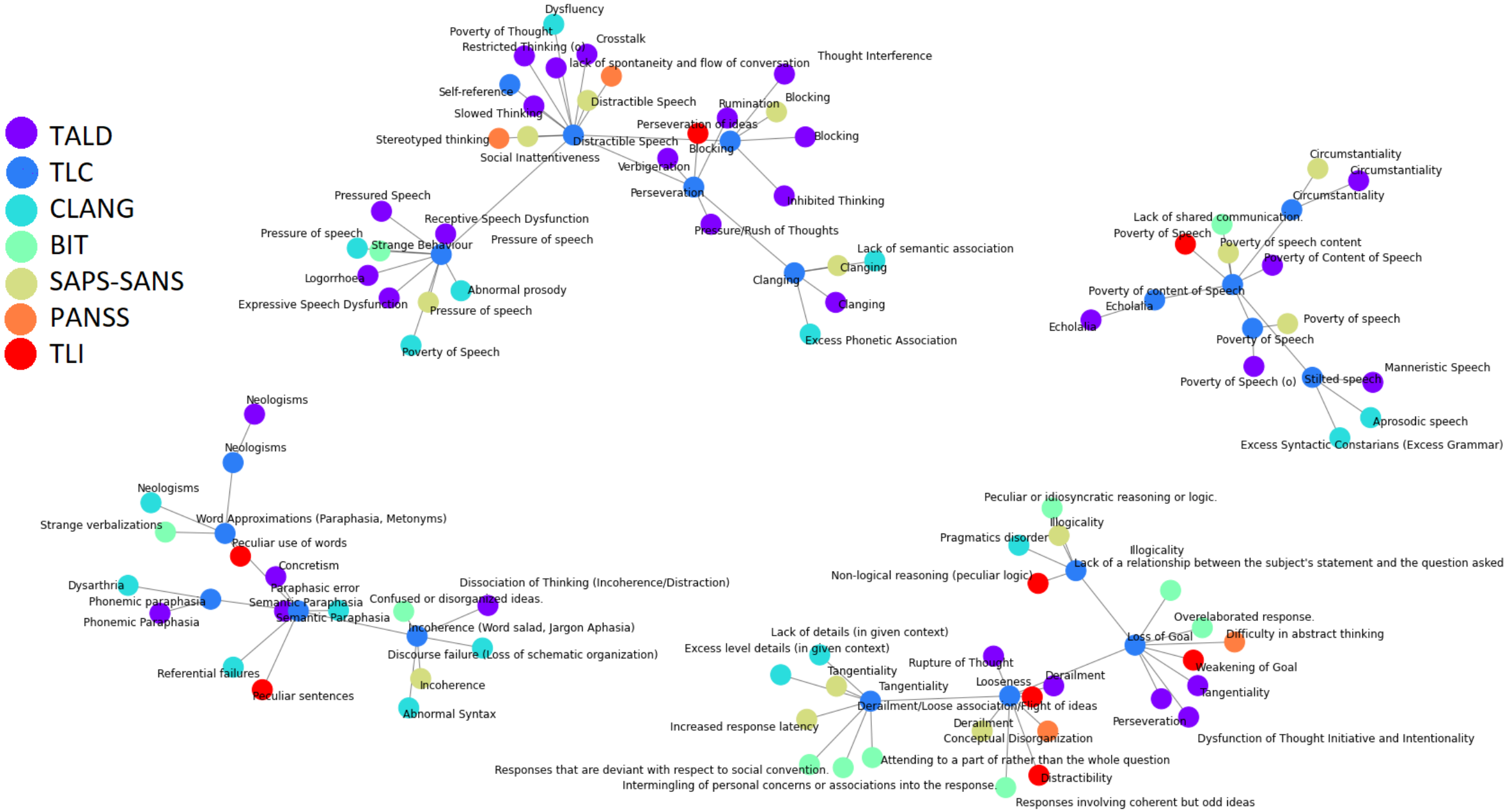
1. Collect English item-level descriptions from rating scales (Thanks to Sreeraj V S)
2. Remove examples and severity level descriptions
3. Sentence-level BERT embedding
    a) **all-mpnet-base-v2**
    b) Produces a 768-length vector embedding of each input sentence or paragraph
4. Calculate item-to-item similarity

https://huggingface.co/sentence-transformers/all-mpnet-base-v2



103 selected items

TALD	30
SAPS-SANS	16
PANSS	4
TLC	20
TLI	6
CLANG	17
BIT	11
Total	103


TALD
TLC
CLANG
BIT
SAPS-SANS
PANSS
TLI
Dysfluency
Poverty of Thought
Restricted Thinking (o)
Crosstalk
lack of spontaneity and flow of conversation
Thought Interference
Self-reference
Slowed Thinking
Distractible Speech
Rumination
Blocking
Stereotyped thinking
Social Inattentiveness
Distractible Speech
Perseveration of ideas
Blocking
Blocking
Verbigeration
Perseveration
Inhibited Thinking
Pressure/Rush of Thoughts
Pressured Speech
Receptive Speech Dysfunction
Pressure of speech
Strange Behaviour
Pressure of speech
Logorrhoea
Abnormal prosody
Expressive Speech Dysfunction
Pressure of speech
Poverty of Speech
Lack of semantic association
Clanging
Clanging
Clanging
Excess Phonetic Association
Circumstantiality
Circumstantiality
Circumstantiality
Lack of shared communication.
Poverty of Speech
Poverty of speech content
Poverty of Content of Speech
Poverty of content of Speech
Echolalia
Echolalia
Poverty of speech
Poverty of Speech
Poverty of Speech (o)
Stilted speech
Manneristic Speech
Aprosodic speech
Excess Syntactic Constarians (Excess Grammar)
Neologisms
Neologisms
Neologisms
Strange verbalizations
Word Approximations (Paraphasia, Metonyms)
Peculiar use of words
Concretism
Dysarthria
Paraphasic error
Phonemic paraphasia
Semantic Paraphasia
Phonemic Paraphasia
Semantic Paraphasia
Dissociation of Thinking (Incoherence/Distraction)
Confused or disorganized ideas.
Incoherence (Word salad, Jargon Aphasia)
Discourse failure (Loss of schematic organization)
Referential failures
Peculiar sentences
Incoherence
Abnormal Syntax
Peculiar or idiosyncratic reasoning or logic.
Illogicality
Pragmatics disorder
Illogicality
Lack of a relationship between the subject's statement and the question asked
Non-logical reasoning (peculiar logic)
Overelaborated response.
Loss of Goal
Difficulty in abstract thinking
Weakening of Goal
Tangentiality
Perseveration
Dysfunction of Thought Initiative and Intentionality
Lack of details (in given context)
Excess level details (in given context)
Tangentiality
Tangentiality
Rupture of Thought
Looseness
Derailment
Derailment/Loose association/Flight of ideas
Increased response latency
Derailment
Conceptual Disorganization
Responses that are deviant with respect to social convention.
Attending to a part of rather than the whole question
Intermingling of personal concerns or associations into the response.
Distractibility
Responses involving coherent but odd ideas

Ask the experts for their attributions

- 103 items
- Assign to one of 4 groups based on descriptions
- 6 experts
- Compare their responses

Results from the experts: Substantial agreement

Fleiss' kappa

Ratings	Fleiss' kappa	SE	95% CI	
			Lower	Upper
Overall	0.617	0.016	0.585	0.648
1	0.636	0.028	0.582	0.690
2	0.476	0.028	0.422	0.530
3	0.589	0.028	0.535	0.643
4	0.716	0.028	0.662	0.770

Note. 103 subjects/items and 6 raters/measurements. Confidence intervals are asymptotic.

Group 2 features involve abrupt topic shifts in response to stimuli, repeated words or subjects, speech interruptions, and memory lapses regarding previous speech.

Discussion & future steps

- Rating scales for FTD show a wide range of conceptualizations and wordings
- Expert raters agree on the semantic division using TLC as backbone
- The issue of antonyms
  - Sentence 1: responses are too slow
  - Sentence 2: responses are too fast.
  - BERT Sentence Similarity: 0.891
  - Repeat clustering with each scale and compute common connections?
- Repeat with all psychotic symptoms?

Data-driven approaches for clinical harmonization

Frederike Stein
Sunny Tang

Previous work on dimensions of formal thought disorder

- **explorative and confirmatory factor analyses of FTD symptoms** (SANS+SAPS) in two samples with a total *N*=1041 patients with affective and psychotic disorders

explorative

3 factors explaining 50.58% of total variance

Factor	Symptom	Cronbach´s Alpha
Disorganization	Tangentiality	.857
	Derailment	
	Circumstantiality	
	Pressure of speech	
Emptiness	Poverty of speech	.757
	Poverty of content	
	Increased latency of response	
	Blocking	
Incoherence	Incoherence	.728
	Illogicality	
	Distractibility	

MDD *n*=887 (571 f) Ø 38.62 years, SD 13.23
BD *n*=151 (81 f) Ø 36.56 years, SD 12.13
SSD *n*=144 (65 f) Ø 34.7 years, SD 11.93

confirmatory

- sample 2: χ^2=44.88, CFI=0.909, RMSEA=0.046
- total sample: χ^2=66.097, CFI=0.928, RMSEA=0.045
- age- and sex-matched sub-sample with same n per diagnosis: χ^2=51.43, CFI=0.924, RMSEA=0.045

Cross-validated 3-factor model

Extension of existing factor models in SZ

FTD not only a core symptom of SZ, but also **prevalent in other disorders**

Stein et *al*., 2022, Schizophrenia Bulletin

Reproducible?

Very consistent with results from another language, independent sample, independent raters

- Stein et al., 2022
  - SANS/SAPS scale, German language
  - Disorganization
    - Tangentiality
    - Circumstantiality
    - Derailment
    - Pressure of Speech
  - Incoherence
    - Incoherence
    - Illogicality
    - Distractibility
  - Emptiness
    - Poverty of Speech
    - Poverty of Content of Speech
    - Increased latency of response
    - Blocking
- Our Study:
  - TLC scale, English language
  - Inefficiency
    - Tangentiality
    - Circumstantiality
    - Derailment
    - Pressure of Speech
    - Distractible Speech
    - Perseverations
  - Incoherence
    - Incoherence
    - Illogicality
    - Neologism
  - Impaired Expressivity
    - Poverty of Speech
    - Increased Latency
    - Decreased vocal inflections

Data-driven approaches for clinical harmonization

23 studies/samples
10 countries
8 languages:

FOR2107
TALD_Study
FOR2107_SPAPS project
Remora
LPOP
ACES
PsyCL
In-PsyCL
TLI_DEU_comb
TLI_DEU2
Early_schizobipolar_DEU
Discourse_lena
CAFLIP
TOPSY
IMPLEMENT
OCoPS-P
BrAGG-SoS
Gesture study
Discouse Santander
Discourse_BR-PR
Turkish adaptaion of TALD
RAPSODI
SCAANS
IPS

PIs involved:
Tilo Kircher
Frederike Stein
Sunny Tang
Berna Yalincetin
Emre Bora
Lena Palaniyappan
Tim Crowe
Sebastian Walther
Rosa Ayesa Aricola
Raffael Massuda
Emre Mutlu
Iris Sommer
Sreeraj Venkatasubramanian
Katharina Stegmayer

DISCOURSE IN PSYCHOSIS

Basic Information

- Total sample size: **6,376** subjects
- Groups included:
  - Healthy controls
  - Help seeking
  - First degree relation
  - Major Depression
  - Bipolar Disorder
  - Schizophrenia spectrum disorders
- Dx methods:
  - DSM-IV / V and SCID I
  - ICD-10
- Study exclusion criteria:
  - verbal IQ<80 or <70
  - history of head trauma or unconsciousness
  - severe medical illnesses (cancer, autoimmune diseases, and infections)
  - neurological illness
  - current substance dependence
  - substance induced disorders or disorders secondary to general medical condition
  - disorders affecting speech or language
  - movement disorders affecting speech
  - ECT
- Type of FTD data acquisition:
  - semi-structured clinical interview or discourse interview

HC	MDD	BD	SSD	help seeking	other psychotic dx	first degree relation
2090	1398	380	2241	150	40	77

No. of sites with:

SANS	SAPS	TLI	TLC	TALD	PANSS	CDI
10	**7**	**8**	**9**	**7**	**14**	**1**

Basic Information

Combination	Sites	Total N	HC	MDD	BD	SSD
PANSS + SANS	PsyCL IPS	169	0	0	0	169
PANSS + SAPS	PsyCL	84	0	0	0	84
PANSS + TLI	TLI_DEU_comb Discourse_lena TOPSY IMPLEMENT	557	158	0	0	399
PANSS + TLC	PsyCL CAFLIP Gesture study IPS	311	0	0	0	311
PANSS + TALD	OCoPS-P BrAGG-SoS Turkish adaptaion of TALD RAPSODI	334	35	24	20	255
PANSS + TLC + SANS	PsyCL IPS	169	0	0	0	169

Basic information

Combination	Sites	Total N	HC	MDD	BD	SSD
SANS + SAPS	FOR2107 TALD_Study FOR2107_SPAPS project PsyCL TLI_DEU2 SCAANS	3770	1651	1374	285	460
SANS + TLC	TALD_Study Remora LPOP ACES PsyCL IPS	540	156	63	20	301
SAPS + TLI	FOR2107_SPAPS project TLI_DEU2 Early_schizobipolar_DEU	572	202	117	101	152
SANS + SAPS + TALD	FOR2107 TALD_Study FOR2107_SPAPS project German	918	403	392	69	54

Basic Information

Combination	Sites	Total N	HC	MDD	BD	SSD
SANS + SAPS + TLI + TALD	FOR2107_SPAPS project	392	202	117	26	47
SAPS + SANS + TLC + TALD	TALD_Study	210	64	63	20	63

Other scales assessed include:
HAM-D (6)
YMRS (8)
SPQ-B (1)
BPRS (4)
BNSS (2)
CLANG (1)
CASH (4)
CAINS (2)
BPS (1)

Top combinations

Combination	Sites	Total N	HC	MDD	BD	SSD
SANS + SAPS	FOR2107, TALD_Study FOR2107_SPAPS project PsyCL TLI_DEU2 SCAANS	**3770**	1651	1374	285	460
SANS + SAPS + TALD	FOR2107 TALD_Study FOR2107_SPAPS project German	**918**	403	392	69	54
SAPS + TLI	FOR2107_SPAPS project TLI_DEU2 Early_schizobipolar_DEU	**572**	202	117	101	152
PANSS + TLI	TLI_DEU_comb Discourse_lena TOPSY IMPLEMENT	**557**	158	0	0	399
SANS + TLC	TALD_Study Remora LPOP ACES PsyCL IPS	**540**	156	63	20	301

Analysis approach

Deidentified data will be processed by Dr. Frederike Stein and team, University of Marburg Germany, and aggregated for analysis:

- **Step 1: Poly-Correlations**
  - Identify items/symptoms that can be collapsed, i.e., multi-colinearity
- **Step 2: Pairwise Correlation Matrix**
- **Step 3: Random Forest Imputation**
  - Impute missing data based on existing information (decide which variables should be used for that)
- **Step 4: Factor analysis**
  - Exploratory
  - Confirmatory (could be of interest for sites that will not share single items)
  - Subsamples
- **Step 5: Method for computing factor scores**

All analyses will be performed using R and/or Python

Variables for analysis

- Diagnosis
- Age
- Sex
- Medication (CPZ)
- Age of Onset
- Duration of illness
- Years of education
- Number of hospitalizations
- At least two FTD rating scales:
  - SANS
  - SAPS
  - TALD
  - TLC
  - TLI
  - PANSS
  - BPRS

SubjID	Subject ID should match the format of the IDs in other files
Dx	Diagnosis (0=controls, 1=patients).
Age	Age in years at time of scan (please use integers)
Sex	Males=1, Females=2
Race_Ethnicity	1=White/caucasian, 2=Black/African, 3=Asian 4=other, NA=no info
Site_name	Site name (categorical variable)
CPZ	CPZ is chlorpromazine equivalents according to Woods et al. 2005
AO	AO is age of onset in years, i.e. age at first psychiatric episode
DURILL	DURILL is duration of illness or Age - AO
HAND	HAND is handedness (0=Right, 1=Left, 2=Ambidextrous, NA=Not Available)
EDU	Years of Education
NHOSP	NHOSP is number of hospitalizations for psychiatric episodes and should be an integer; NA = no information available
Major_Depressive_Disorder_Current	1 = yes, 0 = no, and NA = no information available.
Major_Depressive_Disorder_Lifetime	1 = yes, 0 = no, and NA = no information available.
Depression_With_Psychotic_Features_Current	1 = yes, 0 = no, and NA = no information available.
Depression_With_Psychotic_Features_Lifetime	1 = yes, 0 = no, and NA = no information available.
Bipolar_Disorder_Current	1 = yes, 0 = no, and NA = no information available.
Bipolar_Disorder_Lifetime	1 = yes, 0 = no, and NA = no information available.
Bipolar_Disorder_Type	BD-I=1; BD-II=2; NA = no information available.
Bipolar_Disorder_With_Psychotic_Features_Current	1 = yes, 0 = no, and NA = no information available.
Bipolar_Disorder_With_Psychotic_Features_Lifetime	1 = yes, 0 = no, and NA = no information available.
Psychotic_Disorder_Current	1 = yes, 0 = no, and NA = no information available.
Psychotic_Disorder_Lifetime	1 = yes, 0 = no, and NA = no information available.

https://docs.google.com/spreadsheets/d/1rMlSYiAGoLqRmBzScq1hbZBAQOyiHLSvw9J2ikSxj4k/edit?usp=sharing

Data transfer agreement

- DTA is prepared
- In case you need a DTA please let us know
  - Adapt if necessary using tracked-mode

DATA TRANSFER AGREEMENT

between

Universität Marburg
represented by the President, Prof. Dr. Thomas Nauss
Biegenstraße 10, 35037 Marburg
Germany

Performed by:
Department of Psychiatry and Psychotherapy
Systems Neuroscience in Psychiatry and Psychotherapy
Rudolf-Bultmann-Str.8, 35039 Marburg
Germany

Responsible Scientists: Dr. Frederike Stein

(Hereinafter collectively referred to as "**Recipient Institution**")

And

[Contact information for your site]

University of XXX
Address

Responsible Scientists: XXX

(Hereinafter collectively referred to as "**Provider Institution**")

Next Steps

- **Start data collection (!)**
  - If you are interested in participating: let us know!
- Set-up the code for analyses
- Preliminary analyses in a smaller sample to test the study approach

Thank you for your attention!

Discussion Comments Questions