Comprehension of embedded clauses in schizophrenia with and without formal thought disorder

*Derya Cokal, PhD[1]; Vitor Zimmerer, PhD[2]; Rosemary Varley, PhD[2]; Stuart Watson, MRCPsych, MD[1,3] and Wolfram Hinzen, PhD[4-6]

[1]Institute of Neuroscience, Newcastle University, Newcastle upon Tyne, UK

[2]Department of Language and Cognition, University College London, London, UK

[3]Northumberland Tyne and Wear NHS Foundation Trust, Newcastle, UK

[4]ICREA (Institució Catalana de Recerca i Estudis Avançats) Barcelona, Spain

[5] Department of Translation and Language Sciences, Universitat Pompeu Fabra, Barcelona, Spain

[6] FIDMAG Germanes Hospitalaries Research Foundation, Benito Menni Hospital, Barcelona, Spain

***Corresponding author**: Institute of Neuroscience

Newcastle University, NE1 7RU, United Kingdom

e-mail: deryacokal@gmail.com

Phone: +44 7447 815996 & +90 532 37238 32

Contributors

DC and WH wrote the first draft of the manuscript. DC processed data for statistical analysis and ran statistical analyses. SW managed and carried out recruitment and assessment. WH and VZ conceived the study. VZ, RV, and WH designed the experiment. All authors offered suggestions and edits to subsequent drafts and gave final approval for submission.

Acknowledgements

Competing Interests

The authors have no competing interests in relation to the work described in this manuscript.

Disclosures

The authors declare no conflicts of interest.

Title: Comprehension of embedded clauses in schizophrenia with and without formal thought disorder

Abstract

Formal thought disorder (FTD) in schizophrenia (SZ) is clinically manifest primarily through language production, where linguistic studies have reported numerous anomalies including lesser use of embedded clauses. Here we explored whether problems of language may extend to comprehension and clause embedding in particular. A sentence-picture matching task was designed with two conditions in which embedded clauses were either presupposed as true (factive) or not. Performance across these two conditions was compared in people with SZ and moderate-to-severe FTD (SZ+FTD), SZ with minimal FTD (SZ-FTD), first degree relatives of people with SZ, and neurotypical controls. The SZ+FTD group performed significantly worse than all others in both conditions, and worse in the non-factive than the factive one. These results demonstrate language dysfunction in comprehension specific to FTD in a critical aspect of grammatical complexity and its associated meaning, which has been independently known to be cognitively significant as well.

Keywords: Schizophrenia, formal thought disorder, comprehension, syntax, language dysfunction

Introduction

Neurocognitive studies of people with schizophrenia (SZ) have long revealed a number of language deficits (Kuperberg, 2010a, b). This has been shown in studies involving undifferentiated schizophrenia groups as compared with neurotypical controls (Morice & McNicol, 1986; Tavano et al., 2008), in first degree relatives (FDR) of people with SZ as compared with neurotypical controls (Docherty et al., 1998; Docherty & Gordinier, 1999; Condray et al., 1992), and in patients with SZ that have the specific symptom of formal thought disorder (FTD) (Rochester & Martin, 1979; McKenna & Oh, 2005). Specifically, people with SZ produce utterances with reduced syntactic complexity and more syntactic errors (Hoffman & Sledge, 1988; Fraser et al., 1986; Morice & McNicol, 1986; Tavano et al., 2008). In line with that, they also underperform on standardized tests for aphasia (Morice & McNicol, 1985; Landre et al., 1992; Oh et al., 2002; Rodriguez-Ferrera et al., 2001; Barrera et al., 2005; Little et al., 2018). However, many studies of syntactic complexity (e.g. Tavano et al., 2008; Marini et al., 2008) have used composite measures of syntactic complexity, making it difficult to isolate the effect of particular syntactic construction types on the impairment seen. It is thus currently unclear which specific forms of syntactic complexity distinguish SZ groups with and without FTD (Çokal et al., 2018; Oh et al., 2002; Stirling et al., 2006). Spontaneous speech in FTD exhibits derailment, tangentiality, and incoherence, among other forms of disorganization, and it is the primary basis for clinically identifying and rating FTD (Andreasen, 1979). An objective linguistic marker could therefore be clinically useful for early detection, diagnosis, and monitoring of SZ and symptoms (Bedi et al., 2015; Zimmerer et al., 2017). A linguistic impairment involving specific aspects of linguistic organization could also mediate FTD neurophysiologically

(McKenna & Oh, 2005), in line with the fact that the neural correlates of FTD relate to classical language areas (Wensing et al., 2017).

Several studies have reported impairments in language comprehension in SZ as well, both in tasks where patients answer questions of the meanings of syntactically complex sentences (Condray et al., 1996; Condray et al., 2002; Bagner et al., 2003) and perform sentence-picture matching tasks (Lelekov et al., 2000; Tavano et al., 2008). Of these, Lelekov et al. (2000), found an effect of complexity on comprehension, which however was not specific to syntactic processing. Bagner et al. (2003) also found such a complexity effect, again in undifferentiated SZ, which correlated with working memory measures. More differentiated and fine-grained linguistic measures are currently needed to illuminate the role of language in the neurocognition of SZ and FTD, and its translational significance.

Here we aimed to explore comprehension of a specific construction type in patients with FTD as compared to patients with SZ without FTD, FDRs, and controls: sentences with clausal embedding (subordinated clausal arguments), as illustrated in sentences such as *The man thinks it is cold outside*. Such constructions involve a particular kind of content: they convey a particular thought (of the speaker), which itself refers to the thought of another person (*it is cold outside*). Any sentence featuring such a subordinated clause (i.e., *it is cold outside*) therefore expresses a 'meta-representation' (Frith, 1992) (i.e. a representation containing a representation of the mental states of others). In a parallel study executed in the same sample (Çokal et al., 2018), we found that the FTD group used significantly fewer embedded clauses compared to neurotypical speakers and FDRs in their spontaneous speech production. This motivated the present study directed at embedded clauses in comprehension. Numerous studies have independently shown that embedded clauses are cognitively significant, which is

presumably due to their meta-representational content. In particular, they relate to performance on explicit false belief tasks (DeVilliers, 2007; DeVilliers & DeVilliers, 2012; Durrleman & Franck, 2015; Durrleman et al., 2017).

Meta-representation is a natural feature of sentences involving embedded clauses, since such sentences (e.g. *He thought she liked him*) can be true, even when the embedded clause is false (i.e., she didn't like him). This is what makes such sentences so apt to reason about mental states (i.e., representations of the world, as opposed to the world itself). However, not all sentences with subordination have this feature. Thus, a speaker who says *He knew she liked him* takes it for a fact that she liked him – this fact is presupposed to obtain, which is why such subordinated clauses are said to be 'factive' (Hinzen & Sheehan, 2011). Factive vs. non-factive interpretations of embedded clauses can therefore be used as a critical test of the comprehension of clausal subordination.

We therefore chose a sentence-picture matching (SPM) paradigm to manipulate this distinction. To explore the nature of the language dysfunction involved in FTD, we also investigated the relationship between comprehension of factive and non-factive sentences and (a) semantic/syntactic competencies, (b) executive function, (c) general intelligence (current IQ), and (d) non-verbal reasoning. People with FTD have long been shown to display a wide range of cognitive impairments (Schaefer et al., 2013), specifically in executive functioning (Barrera et al., 2005), semantic memory (Barrera et al., 2005; Stirling et al., 2006), and general IQ (Basso et al., 1998; Cuesta & Peralta 1995; Dibben et al., 2008; O'Leary et al., 2000). It has been claimed that the origins of FTD are more closely linked to deficits in executive functioning and semantic processing than to impairments in general syntactic function or general cognition (Stirling et al., 2006).

Since people with SZ, and particularly those with FTD, tend to underperform on a range of cognitive tasks in general (Fioravanti *et al.* 2005; Reichenberg & Harvey 2007; Schaefer *et al.* 2013), we expected participants with SZ (both with and without FTD) to perform less well on the SPM task than neurotypical controls (NC) and first-degree relatives of people with SZ (FDR). Moreover, given that FTD measured as conceptual disorganization is graded across the SZ spectrum, we expected participants without FTD (SZ-FTD) to perform in between non-clinical controls and participants with FTD (SZ+FTD). In addition, since FDRs have been shown to have subclinical linguistic impairments (Condray et al., 1992; Docherty et al., 1998; Docherty & Gordinier, 1999), they should differ from neurotypical controls as well. We further predicted that participants with SZ, and particularly those with FTD, would perform better on factives than on non-factives, since the former differ from the latter in referring to facts (the world) rather than beliefs (representations of the world).

Methods

A favorable ethics opinion was obtained from NRES Committee North East - Newcastle & North Tyneside 2. All participants provided written, informed consent. The experiment was part of a larger test protocol of the AHRC-funded Language and Mental Health project.

Participants. 25 participants with SZ were recruited from a UK secondary care mental healthcare trust (Northumberland, Tyne and Wear (NTW) NHS Foundation Trust). All participants with SZ met DSM-IV diagnostic criteria and scored at least 60 on the Positive and Negative Symptom Scale for Schizophrenia (PANSS) (Kay et al., 1987). Thirteen neurotypical controls (NC) were recruited via an advert placed in hospital and university buildings. Twelve first-degree relatives (FDRs) of people with schizophrenia were also recruited via carer groups and via patients in NTW Trust and

study participants. For all participants, exclusion criteria were substance dependence or abuse, pervasive developmental disorder interfering with language skills (such as dyslexia), significant head injury, stroke, brain tumor and epilepsy. In neurotypical control participants and FDRs, past or current psychotic disorder was also an exclusion criterion. Participants with SZ were dichotomized on the basis of their score on question 2 ('Conceptual Disorganization' [CD]) of the PANSS. Those who scored 3 (minimal FTD) or less were assigned to the SZ-FTD group (N=13), while those who scored 4 (moderate FTD) or more were assigned to the SZ+FTD group (N=12). All subjects were native speakers of English. Table 1 summarizes demographic and clinical data. (Supplementary Table 1 for independent t-test comparisons of age, education, and WASI IQ).

Table 1

Demographic and clinical data (mean and standard deviation) for neurotypical controls (NC), first degree relatives (FDR), and participants with (SZ+FTD) and without (SZ-FTD) formal thought disorder.

	NC	FDR	SZ-FTD	SZ+FTD	F(df) P[1]	F(df) P[2]
Mean age (years)	45.31(12.7)	47.67 (9.7)	39.15 (7.09)	47.92 (13.8)	1.7 (3,49) .170	4.081 (1,24) 0.055
Gender (M/F)	(6/7)	(6/6)	(9/4)	(10/2)	------------	-------------
Years of education	15.84(3.7)	17 (4.12)	13 (4.2)	4.54 (3.5)	2.38 (3,49) 0.081	0.854 (1,24) 0.365
NART	109 (8.9)	104 (10.55)	96.23 (14.41)	95(15.25)	.622 (3,49) .604	0.025 (1,23) 0.876
IQ*	105.9(8.1)	104.5 (9.80)	90.5 (18.20)	77.75 (14.4)	12.19 (3,49) 0.001	3.739 (1,24) 0.066
Illness duration (months)			212.31 (87.80)	266.50 (143.84)	1.317 (1,24) 0.263	1.317 (1,24) 0.263
PANSS item 2	1.00 (0.00)	1.33 (0.65)	1.54 (0.66)	5.17 (0.94)	108.953 (3,49) <0.0005	126.819 (1,24)< 0.0005
PANSS positive subscale	8.31 (0.95)	12.33 (6.21)	20.77(4.62)	30.50 (5.13)	55.971 (3,49) <0.0005	24.921 (1,24)< 0.0005
PANSS negative subscale	8.77 (1.42)	11.75 (6.59)	18.77 (5.99)	25.00 (7.65)	19.138 (3,49) <0.0005	5.186 (1,24) 0.032
PANSS general psychopathology Subscale	22.46(4.65)	31.83 (10.68)	46.38 (7.99)	57.75 (10.71)	39.378 (3,49) <0.0005	9.144 (1,24) 0.006
PANSS total	39.54(5.52)	55.92 (21.89)	86.00 (14.60)	113.25 (19.55)	49.105 (3,49) < 49.105	15.765 (1,24) 0.001

Note: NART refers to the IQ estimate generated from the National Adult Reading Test (Nelson, 1982). IQ* refers to the Wechsler Abbreviated Scale of Intelligence (WASI) full scale IQ4 (Wechsler, 2011). [1] One-way ANOVA; ALL 4 GRPS; [2]SZ-FTD vs SZ+FTD

Materials. Our experiment was a sentence-picture-matching task (SPM) using a within-subject design with two conditions in which the embedded clauses were either 'factive' (required to be true for the sentence as a whole to be true) or 'non-factive' (not required to be true, and in fact they were false with respect to their best-matching pictures). We manipulated the type of subordinated clause by using different types of embedding verbs (see Supplementary Table 2 for all stimuli). For a factive subordinating verb we used either *knows that* or *it is clear that*, whereas for non-factive clauses we used either *thinks that* or *it only seems…that*:

There were 35 sets of items in total, each consisting of an array of two distractor pictures and a target picture, presented vertically on the same A4 page. Pictures were based on ten different scenarios (or event types). Each picture depicted: (a) one experiencer who was the subject of the test sentence (e.g., *the man thinks that…*) and was represented in a way that allowed inferences regarding his beliefs, and (b) a situation (e.g., *…it is warm outside*) described by the embedded clause depicting his beliefs to be correct or not. The experiencer's mental state was represented via his or her clothing, facial expression or body pose. For example, in the picture matching *The man thinks that it is warm* outside, the man is wearing summer clothing (see top picture in Figure 1). In all pictures except those matching *It is clear that…* sentences, the experiencer could not see the situation because of a wall or other obstacle, allowing for false beliefs. Participants had to check whether test sentences matched the apparent beliefs of the experiencer, and whether these beliefs matched the situation (e.g., whether it is warm or cold in the picture). In Figure 1, the test sentence *The man thinks that it is cold outside* matches the middle picture (the top and bottom pictures are distractors), which fits the experiencer's state, and although it is not actually cold outside, this is irrelevant for the truth value of the sentence as a whole (experiencer match + world

10

mismatch). The test stimulus *The man knows it is warm outside* matches only the top picture, in which the man clearly assumes the weather is warm, and the weather is, indeed, warm. The middle [experiencer mismatch, world match] and bottom [experiencer match, world mismatch] pictures are distractors. In all arrays, pictures that matched in the factive condition had "true" embedded clauses, while for the non-factive condition embedded clauses were always false (i.e., counter-factive).
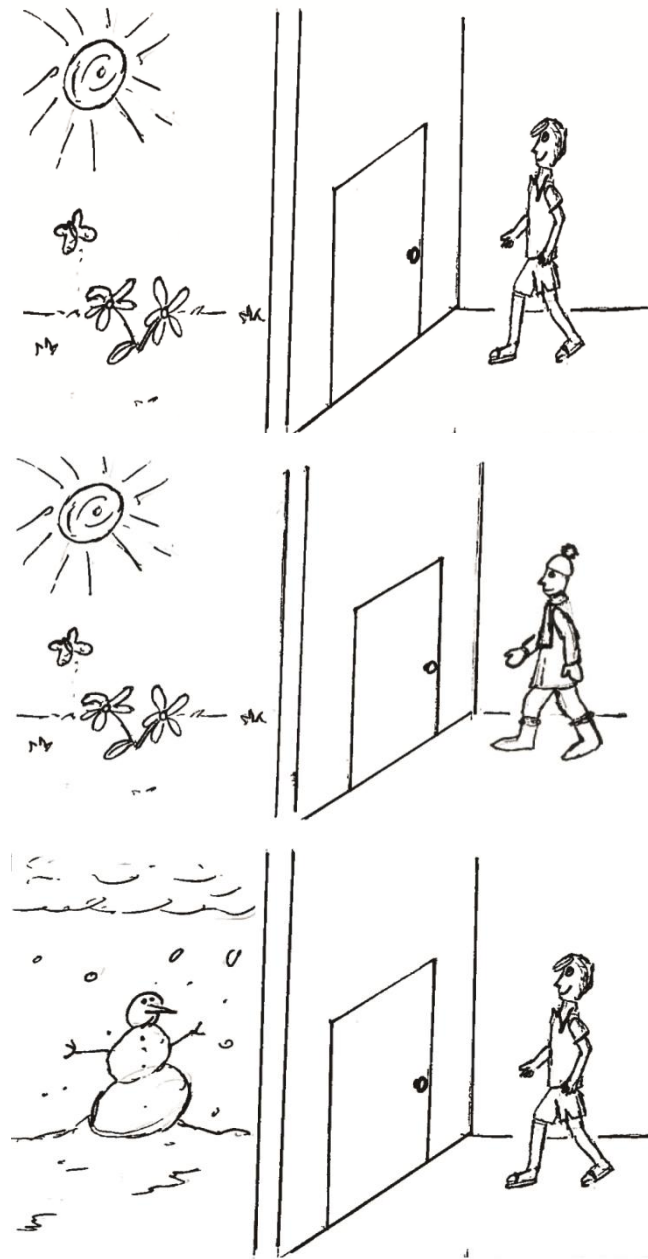
Figure 1. Example of a picture array. We used this particular combination of pictures for two sentences: *The man thinks that it is cold outside* (middle picture correct) and *The man knows that it is warm outside* (top picture correct). Target picture position was balanced across trials.

Neuropsychological Measures

To explore connections between comprehension of embedded clauses and other neurocognitive domains, the following tests were administered, all of which have been previously used to reveal cognitive dysfunction in FTD:

      1. **Sentence comprehension**: Test for Reception of Grammar (TROG-2) (Bishop, 2003).

      2. **Semantic memory.** The Pyramids and Palm Trees test (PPT) (Howard & Patterson, 1992).

      3. **Non-verbal reasoning/executive function.** The Brixton Spatial Anticipation Test (Burgess & Shallice, 1997).

      4. **Verbal and Non-Verbal intelligence.** The Wechsler Abbreviated Scale of Intelligence (WASI-II) was used to measure current intelligence (using the Full Scale IQ-4 measure) which incorporates non-verbal intelligence (Perceptual Reasoning Index (PRI); calculated from Block Design and Matrix Reasoning sub-test scores) and verbal intelligence (Verbal Comprehension Index (VCI); calculated from Vocabulary and Similarities sub-test scores).

Procedures

      For both the SPM experiment and neuropsychological tests, participants were evaluated at their home or the UK secondary care mental healthcare trust (Northumberland, Tyne and Wear (NTW) NHS Foundation Trust). To test sentence picture matching, the experimenter showed the participant one picture array at a time and then a stimulus sentence was read aloud by a native English-speaking experimenter. Participants were asked to point at the picture that best matched the stimulus. Participants could ask for a single repetition. Immediate self-corrections were also allowed, and the corrected response was recorded. Picture order was randomized so that

each position was the target for approximately the same number of trials (top = 12 trials, center = 11 trials, bottom = 12 trials). Each neuropsychological test took 20 - 40 minutes, depending on clinical conditions, and was arranged, if necessary, on different days.

Data Analysis

Data analysis consisted of three stages. In stage one, to address SPM performance across groups, we firstly calculated the overall SPM accuracy score for each group, which is the sum of each participant's correct responses for all trials divided by the total number of test items. Since all overall SPM accuracy scores were not normally distributed, Mann-Whitney U pairwise comparisons were carried out when Kruskal-Wallis showed a significant effect of group (Table 2). In addition, Bonferroni corrections were applied. Since our first hypothesis was directional (i.e., SZ groups – especially SZ+FTD – would have lower performance), we report one-tailed p-values in Table 2.

Then, we calculated each participant's accuracy score for factive and non-factive conditions by summing correct responses for each condition and dividing by total number of factive or non-factive trials. In order to determine whether there was a significant difference across groups, if Kruskal-Wallis showed a significant group effect, we performed Mann-Whitney U pairwise comparisons with Bonferroni corrections (Table 3). In addition, in order to compare group performance differences across factive and non-factive conditions, as well as subordinate clauses (i.e., know, think, it only seems, and it is clear), Friedman's ANOVA and Wilcoxon tests were performed.

In stage two, we explored whether other clinical scores (i.e., PANSS positive/negative subscales, PANSS General Psychopathology Subscale) correlated

14

with overall accuracy-scores for SPM and factive/non-factive conditions by conducting Kendall's tau-b correlation coefficient analyses (Table 4).

In stage three, since according to Shapiro-Wilk tests the distributions for our outcome variables were non-parametric, we ran Kendall's tau-b correlation coefficient analyses between overall SPM accuracy scores and participants' performance on standardized tests of sentence comprehension (TROG-2), visual semantic memory (Pyramids and Palm Trees – 3 picture version [PPT]), IQ [WASI]), and executive functioning (Brixton) (Table 5). Based upon our directional hypotheses, we ran one-tailed correlations. Instead of adjusting p values using the number of comparisons n, we used a different solution (Sankoh et al., 1997) and adjusted it using $n^{1-r(.k)}$, where r(.k) is the average correlation between outcomes.

Results

Stage 1: Overall SPM accuracy

A Kruskal-Wallis test revealed that groups differed significantly in their overall SPM accuracy scores. The SZ+FTD group returned the lowest scores compared to NC, FDR, and SZ-FTD (Tables 2 & 3).

Table 2

Between-group differences in overall accuracy in SPM, and factive and non-factive conditions.

Variable	Model Type	Statistics	P[1]	d_{cohen}
Overall accuracy in SPM	Kruskal-Wallis	H (3) = 17.144	.001*	1.191
Factive condition	Kruskal-Wallis	H (3) = 11.656	.009*	0.874
Non-factive condition	Kruskal-Wallis	H (3) = 15.162	.002*	1.078

*Mean difference significant at .05 level. [1] one-tailed p-values

Table 3

Pairwise comparisons of neurotypical controls (NC), first degree relatives (FDR), participants with (SZ+FTD), and without formal thought disorder (SZ-FTD) on overall accuracy, accuracy in factive, and non-factive conditions with Mann Whitney U Tests.

Accuracy	Group	Mean (SD)	Comparisons	U	Z	P^1	Effect size (η^2)	Bonferroni correction
Overall	NC	.97 (.038)	NC vs. SZ+FTD	20.000	-3.200	.005*	.39	<.008*
	FDR	.97 (.030)	NC vs. SZ-FTD	49.000	-1.871	.0395	.12	> .008
	SZ-FTD	.92 (.068)	NC vs. FDR	81.500	.204	>.05	None	> .008
	SZ+ FTD	.71 (.201)	FDR vs. SZ+ FTD	17.000	-3.227	.005*	.42	<.008*
			FDR vs. SZ-FTD	42.500	-1.991	.026*	.14	> .008
			SZ+ FTD vs. SZ-FTD	33.000	-2.460	.007*	.24	<.008*
Factive	NC	.97 (.039)						
	FDR	.98 (.025)	NC vs. SZ+FTD	29.500	-2.749	.003*	.28	<. 008*
	SZ-FTD	.95 (.060)	NC vs. SZ-FTD	70.000	-.820	>.05	None	> .008
	SZ+ FTD	.84 (.154)	NC vs. FDR	80.500	.157	>.05	None	> .008
			FDR vs. SZ+ FTD	24.500	-2.873	.003*	.31	<.008*
			FDR vs. SZ-FTD	61.000	-1.025	>.05	None	> .008
			SZ+ FTD vs. SZ-FTD	39.500	-2.156	.0155*	.18	> .008
Non-factive	NC	.96 (.064)	NC vs. SZ+FTD	22.000	-3.188	.005*	.37	<.008*
	FDR	.94, (.12)	NC vs. SZ-FTD	46.000	-2.122	.025*	.15	> .008
	SZ-FTD	.85 (.19)	NC vs. FDR	70.000	-.509	>.05	None	> .008
	SZ+ FTD	.56, (.35)	FDR vs. SZ+ FTD	23.000	-2.914	.002*	.33	< .008*
			FDR vs. SZ-FTD	50.500	-51.578	>.05	None	> .008
			SZ+ FTD vs. SZ-FTD	42.000	-1.959	.026*	.15	> .008

*A Bonferroni correction was applied by dividing alpha value by the number of group comparisons for each linguistic variable. The significance threshold was set at .008. [1] one-tailed p-values.

Accuracy on factive vs. non-factive trials.

Kruskal-Wallis test showed a significant difference across groups for factive and non-factive conditions. In both factive and non-factive trials, SZ+FTD performed poorly compared to NC, FDR, and SZ-FTD (Tables 2 & 3). Only in the case of SZ-FTD vs. SZ+FTD the differences were not significant after Bonferroni correction.

Friedman's ANOVA and Wilcoxon tests revealed that type of subordinated clauses (i.e., factive vs. non-factive) had a significant effect on SZ+FTD's performance: SZ+FTD performed significantly worse in the non-factive condition relative to the factive condition, $\chi^2(1) = 4.455$, p = .035; Z = -2.401; p = .007.

Item analysis on each type of subordinated clause. Friedman's ANOVA did not demonstrate a main effect of trial type for NC, FDR, and SZ-FTD: NC: NC: $\chi^2(3) = .412$, p > .05, FDR: $\chi^2(3) = 3.245$, p > .05; SZ-FTD: $\chi^2(3) = 7.306$, p = .058 (see Supplementary Tables 3a for mean and standard deviation in trials for groups). Within SZ+FTD participants, Friedman's ANOVA revealed significant differences between trial types, $\chi^2(3) = 8.874$, p = .026. In Wilcoxon tests, SZ+FTD scores in *know* trials were better than *think* trials, as well as in *it only seems* trials. In addition, performance in *it is clear* trials was significantly better than in *think* and *it only seems* trials (see Supplementary Tables 3b for performance comparisons of (SZ+FTD) with Wilcoxon tests).

Stage 2: Correlations between SPM scores and clinical test scores

There was no negative or positive correlation between participants' overall accuracy scores on SPM, factive and non-factive conditions, and clinical-tests (Table 4), other than a negative correlation between participants' CD score and their performances in the SPM experiment.

Table 4

Kendall's tau-b correlation coefficient of clinical tests with overall SPM score, and factive/non-factive comprehension accuracy of joined SZ group (participants with [SZ+FTD] and without thought disorder [SZ-FTD])

		CD score	P[2]. Total positive scale	P. Total negative scale	P. General psychopathology scale	P. T[...] scor[...]
	CD[1] score					
Joined Clinical Group n= 25	P[2]. Total positive scale	τ = .666** p[3] = .001				
	P. Total negative scale	τ = .240 p = .123	τ = .343* p = .021			
	P. General psychopathology scale	τ = .449** p = .004	τ = .617** p = .001	τ = .497** p = .001		
	P. Total Score	τ = .520** p = .001	τ = .659** p = .001	τ = .628** p= .001	τ = .793** p = .001	
	Overall SPM accuracy score	τ = -.436** p = .005	τ = -.212 p= .156	τ = -.221 p = .137	τ = -.129 p = .384	τ = -[...] p = .[...]
	Accuracy in factive conditions	τ = -.418 p = .011	τ = -.205 p = .190	τ = -.116 p = .453	τ = -.048 p = .753	τ = -[...] p = .[...]
	Accuracy in non-factive conditions	τ = -.348* p = .029	τ = -.148 p = .329	τ = -.257 p = .088	τ = -.117 p = .434	τ = -[...] p = .[...]

** Correlation is significant at the 0.001 level; * Correlation is significant at the level 0.05 level.
[1]Conceptual disorganization; [2] PANSS; [3] measured using two-tailed Kendall's τ.

Stage 3: Correlations with standardized neuropsychological measures

First, using the neurotypical control group mean and SD, we converted all SZ group test scores into z-scores (Supplementary Table 4 for mean scores & z-scores). Because of performance differences between factive and non-factive conditions, we correlated standardized tests against these conditions separately. Accuracy across factive and non-factive trials of participants with SZ was strongly correlated and conceptually related. While in SZ-FTD the correlation between performance on factive and non-factive conditions was not significant, it was significant for SZ+FTD (Table 5).

Table 5

Kendall's tau-b correlation coefficient of standardized tests (Sentence comprehension [TROG], visual semantic memory [PPT], Non-verbal reasoning [Brixton]; WASI [IQ]) with factive/non-factive comprehension accuracy of joined clinical groups, and participants with (SZ+FTD) and without thought disorder (SZ-FTD).

Group		Factives	Non-factives	TROG	PPT	Brixton
	Factives					
Joined Clinical Group	Non-factives	$\tau = .469$** $p^4 = .002$				
	TROG	$\tau = .388^1$** $p = .008$	$\tau = .405^1$** $p = .005$			
	PPT	$\tau = .674^1$** $p < .001$	$\tau = .375^1$** $p = .008$	$\tau = .468^1$** $p = .002$		
	Brixton	$\tau = .272$* $p = .039$	$\tau = .228$ $p = .065$	$\tau = .331^1$* $p = .015$	$\tau = .450^1$** $p = .001$	
	WASI	$\tau = .566^1$** $p = .001$	$\tau = .458^1$** $p = .001$	$\tau = .537^1$* $p = .001$	$\tau = .560^1$** $p = .001$	$\tau = .309^1$* $p = .016$
SZ+FTD	Non-factives	$\tau = .400$ $p = .040$				
	TROG	$\tau = .229$ $p = .171$	$\tau = .224$ $p = .173$			
	PPT	$\tau = .742^2$** $p = .001$	$\tau = .425^2$* $p = .031$	$\tau = .359$ $p= .067$		
	Brixton	$\tau = .286$ $p = .105$	$\tau = .264$ $p = .120$	$\tau = .167$ $p= .240$	$\tau = .375$** $p= .048$	
	WASI	$\tau = .635^2$** $p = .003$	$\tau = .388$* $p = .042$	$\tau = .404$* $p= .043$	$\tau = .594^2$** $p= .004$	$\tau = .246$ $p= .137$

SZ-FTD					
Non-factives	$\tau = .315$ $p = .099$				
TROG	$\tau = .243$ $p = .156$	$\tau = .346$ $p = .069$			
PPT	$\tau = .559*$ $p = .011$	$\tau = .153$ $p = .229$	$\tau = .582**$ $p = .006$		
Brixton	$\tau = -.078$ $p = .370$	$\tau = - .115$ $p = .306$	$\tau = .395*$ $p = .039$	$\tau = .386*$ $p = .044$	
WASI	$\tau = .518**$ $p = .012$	$\tau = .396*$ $p = .039$	$\tau = .484*$ $p = .014$	$\tau = .561**$ $p = .006$	$\tau = .146$ $p = .249$

$*p < .05.$ $**p < 0.01.$

[1] In our joined schizophrenia data, we adjusted p values using $n^{1-r(.k)}$: $(.k) = 21 - 0.47 = 1.39$. These were significant after we adjusted the significance threshold for the two outcomes ($p = .05/1.39 = 0.036$).

[2] In our SZ+FTD data $(.k) = 21 - 0.40 = 1.32$, the association between performance on factive/non-factive trials and SZ+FTD's PPT scores was significant after adjustment ($p = .05 / 1.32 = 0.038$).

[3] In our SZ-FTD data $(.k) = 2^{1 - 0.32} = 1.25$, the association between performance on factive/non-factive trials and SZ-FTD's PPT scores was significant after adjustment ($p = .05 / 1.25 = 0.04$).

[4] Measured using one-tailed Kendall's τ.

The effect size differences between NC and SZ+FTD were large for TROG, PPT, Brixton, WASI IQ, and factive and non-factive conditions. Similar differences between NC and SZ-FTD were seen except for the factive condition (Supplementary Table 5a/b for comparisons of effect sizes for means and z-score means). Effect size differences between FDR and SZ+FTD were large (negative) only in TROG, WASI IQ, factive and non-factive conditions. In the comparisons of FDR and SZ-FTD, the effect size for factive and non-factive conditions were medium negative, with a large negative effect difference seen only for WASI IQ score.

Sentence comprehension TROG. There was a significant positive correlation between TROG scores and accuracy in the factive/non-factive conditions for the combined SZ group (Table 5). In a separate sub-group analysis, there were no significant correlations between TROG-2 scores and SZ+FTD's and SZ-FTD's accuracy in either conditions. This suggests that although generalized sentence comprehension is associated with accuracy on factive and non-factive subordinate clauses in SZ groups, this association holds irrespective of whether patients had FTD.

Visual semantic memory PPT. There was a significant positive correlation between the PPT scores and the accuracy of the combined SZ group in both conditions Table 5). While a positive correlation between SZ+FTD's and SZ-FTD's PPT scores and their accuracy was found in the factive condition, in the non-factive condition a positive correlation was seen for SZ+FTD but not SZ-FTD. After adjustment, the association between factive trials and SZ+FTD scores was still significant.

Non-verbal executive function. After significant threshold adjustment, there was a positive correlation between Brixton scores and the accuracy of the combined SZ group in factive/non-factive conditions (Table 5). However, a separate analysis of the

two SZ groups revealed no significant positive or negative correlation between Brixton scores and accuracy in either factive or non-factive conditions. This indicates that even though non-verbal executive function is associated with accuracy on factive/non-factive SPM in SZ groups, this association holds irrespective of whether participants had FTD.

WASI (IQ score): There was a significant positive correlation between WASI IQ scores and SZ participants' accuracy in both joined and separate group analyses in both conditions (Table 5). Patients who had lower WASI scores performed worse in both trial types, with the effect of IQ more pronounced in factive trials.

Overall, accuracy of participants with SZ was significantly correlated with TROG and PPT results, as well as IQ scores. Higher accuracy in sentence comprehension (TROG) and non-verbal executive function (Brixton) was positively correlated with both SZ groups' ability to comprehend factive/non-factive subordinating clauses.

Discussion

The main results of this study confirm our hypotheses, that participants with FTD would show a deficit in comprehending embedded clauses: SZ+FTD significantly underperformed on this task relative to all other groups. They did so as well when factive and non-factive conditions were considered separately, though in this case the comparisons with SZ-FTD did not remain significant after Bonferroni correction. Participants with SZ-FTD did not differ from non-clinical controls in their overall scores or in either of the conditions, suggesting specificity of the problem of comprehending embedded clauses to the SZ+FTD group. Finally, also in line with expectations, patients with SZ+FTD performed significantly worse on non-factive embedded clauses than on factives. These results cohere with results from a linguistic production task in our parallel study of the same sample, which showed significantly

lesser use of embedded clauses in the spontaneous narrative speech in the SZ+FTD group as compared with both neurotypical controls and FDRs (Çokal et al., 2018), while the SZ-FTD group did not differ from these groups. Other studies of speech production in SZ, too, have shown a decrease in syntactic complexity (Morice & McNicol, 1986; Tavano et al., 2008; Marini et al., 2008), but there was no specificity to FTD in these cases, and subordinated clauses were not isolated as linguistic measures.

At this point, two main questions arise: (i) What factor makes sentences with embedded clauses difficult for people with SZ+FTD, whether factive or non-factive, and (ii) What makes non-factive embedded clauses harder? Plausible answers to these questions turn out to be related. A sentence with an embedded clause requires the integration of two units of propositional information, corresponding to the main and embedded clauses, respectively. This integration is not a simple *conjunction* of two such units, since one is a *part of* (or *subordinated under*) the other. This syntactic fact has the crucial semantic correlate that the embedded clause is not evaluated independently, as a comment on a fact in the world, but on how a person *represents* a fact in the world. The syntax of clausal subordination thus introduces a difference in semantic complexity, which is cognitively significant due to the kind of thoughts it inherently encodes.

This is true in both the factive and the non-factive cases, yet with the difference that the embedded clause was always true in the former case and always false (in our design) in the latter case. The cases thus differ only insofar as there was always either *concordance* between beliefs and facts, or else always *discordance*. Therefore, the data plausibly reflect an increase in cognitive demands that discordance entails: representing a proposition as *merely* a thought content, rather than as a fact of the world. There is precedent for this difference between factive and non-factive clausal embedding from a study of people with post-stroke aphasia using the same SPM paradigm, where

participants with aphasia performed equally well as their healthy controls in the factive condition but significantly worse in the non-factive condition (Zimmerer et al., 2018). Differences found in the present study between SZ+FTD and both non-clinical control groups shows that a problem of language comprehension of embedded clauses exists in both aphasia and FTD, yet in FTD, the problem generalizes and is more profound, as it affects the factive case as well. It should be noted that factive clauses differ from non-factive ones syntactically in several respects as well (Sheehan & Hinzen, 2011), but it is not obvious how these syntactic differences would explain that factive clauses appear easier to process.

We expected that in the syntactic comprehension of FDRs a linguistic signal indicating genetic liability might also show, as it has already done in the domain of referential language use (Condray et al., 1992; Docherty et al., 1998; Docherty & Gordinier, 1999). This was not seen in the present study, where FDRs performed very similarly to neurotypical controls, suggesting that this measure of language performance in the domain of syntactic comprehension may not reveal subtle language changes as seen in the domain of the communication disturbances targeted in the above studies.

A final question is whether some non-linguistic neuropsychological deficit could explain the pattern of these results. Correlations between SPM performance and our standardized neuropsychological measures were seen in all cases, from sentence comprehension scores (TROG) to executive function (Brixton), visual semantic memory (PPT), and general IQ (WASI). These measures have all previously shown neurocognitive impairments in FTD. Correlations with the TROG makes good sense, since our task requires general sentence comprehension skills. As for the Brixton, processing non-factive embedded clauses could be argued to require an element of inhibition (suppressing one's own knowledge of the true state of the world, cf.

Pellicano, 2007; Durrleman & Franck, 2015). However, correlations with the Brixton held in both the factive and non-factive cases. As for the PPT, our task also involves semantic decoding of visual arrangements. Beyond that, the PPT requires resisting the association between the two coordinates (e.g., fir trees and palm trees) and determine the more abstract relation with the probe item (pyramid). In short, integration of visual information is necessary – perhaps resembling the way left and right sides of the picture in the SPM task have to be integrated. In addition, low PPT scores indicate an impaired semantic network, which is likely to interfere with task performance. On the other hand, the grammatical mechanism of embedding seems to have little conceptual connection with visual semantics: There seems to be no difference in visual semantic memory terms between *know* and *think*, both of which encode highly abstract, mental state concepts. At a conceptual level, none of our neuropsychological measures seem particularly suited to explain the result of better factive processing over non-factive processing in our data.

Taken together, the multiple correlations seen in our study do not suggest strong explanatory fit between specific neurocognitive domains and comprehension of embedded clauses. Overall, while it is clear that comprehension of embedded clauses does not reflect a problem confined to language in some specific or modular sense, models of how language and cognition connect in neurocognitive pathologies are required to make progress on the neurocognitive deficits underlying SZ and FTD.

Another neurocognitive variable potentially relevant to our results is theory of mind (ToM). A limitation of this study is that no direct standardized measures of ToM were available, raising an important question for future research. While ToM impairment partially independent of general cognitive impairment is well-attested in SZ (Bora et al., 2009; Sprong et al., 2007), it is by no means specific to FTD. Moreover, a

ToM impairment as such does not naturally explain the difference we found between factive and non-factive test stimuli, both of which involve reasoning about mental states. Furthermore, our task, while involving mentalizing, is quite different from the false belief-tasks involved in classical ToM experiments, where one character acquires a false belief about an object by not witnessing a change-of-location undertaken by a different actor (Baron-Cohen et al., 1985). Indeed, in our design, the relevant reported belief in the non-factive condition was *always* false. Finally, numerous studies have shown that performance on standard false belief tasks is related to comprehension of embedded clauses (DeVilliers, 2007; DeVilliers & DeVilliers, 2012; Durrleman & Franck, 2015; Durrleman et al., 2017). This suggests that even in a nonverbal false belief task, language is an 'offline' cognitive tool that enhances ToM performance as participants reason through this task. If so, we would not expect ToM to be an independent explanatory factor for the pattern seen here. On the other hand, the language-dependence of at least some ToM tasks has also been questioned (Dungan & Saxe, 2012; Fouget d'Arc & Ramus, 2011; Onishi & Baillargeon, 2005; Varley & Siegal, 2000). Future work should address this complex issue in people with SZ, with and without FTD. Another limitation of this study is the small sample size, which requires confirmation of the present data in larger groups.

Conclusions

In summary, this study has shown, for the first time, that patients with SZ and FTD but not without FTD show a specific sentence comprehension deficit in the domain of embedded clauses, with performance on non-factive embedded clauses degrading more than on factives. These results cast new light on language impairment in FTD, using embedded clauses as a new window into cognitive dysfunction. This suggests more specific targets for neurophysiological studies. While recent meta-analyses of the

neural correlates of FTD point strongly to classical language territory in the brain (Wensing et al., 2017), the neural basis of clausal embedding deserves more targeted study, in both neurotypical language processing and language processing in FTD. Clausal embedding may also define a target for strategies of cognitive remediation and for attempts to detect, predict, and track disease progression using automated measures (Bedi et al., 2015; Holshausen et al., 2014).

References

Andreasen NC (1979) Thought, language, and communication disorders. Clinical assessment, definition of terms, and evaluation of their reliability. *Arch. Gen Psychiatry*, 36 (12): 1315-21.

Bedi G, Carrillo F, Cecchi GA, Fernandez Slezak D, Sigman M, Mota N., Ribeiro S, Javitt CD, Copelli M, Corcoran CM (2015) Automated analysis of free speech predicts psychosis onset in high-risk youths. *npj Schizophrenia,* 15030..

Bagner D M, Melinder M R, Barch D M (2003) Language comprehension and working memory deficits in patients with schizophrenia. *Schizophrenia Research* 60: 299–309.

Barrera A, McKenna P, Berrios GE (2005). Formal thought disorder in schizophrenia: an executive or a semantic deficit? *Psychological Medicine* 35: 121-132.

Baron-Cohen S, Leslie A M, Frith U (1985) Does the autistic child have a "theory of mind"? *Cognition* 21: 37-46.

Bora E, Yucel M, Pantelis C (2009) Theory of mind impairment in schizophrenia: Metaanalysis. *Schizophrenia Research* 109: 1–9.

Basso MR, Nasrallah HA, Olson SC, Bornstein RA (1998) Neuropsychologial correlates of negative, disorganized and psychotic symptoms in schizophrenia. *Schizophrenia Research* 31: 99-111.

Bishop DVM (2003) *Test for reception of grammar: TROG-2*. PsychCorp, UK: Pearson
  Education Limited.

Burgess PW, Shallice T (1997) *Hayling and Brixton Tests manual*. Bury St. Edmunds,
  UK: Thames Valley Test Company.

Çokal D, Sevilla G, Jones WS, Zimmerer W, Deamer F, Douglas M, Spencer H,
  Turkington D, Ferrier N, Varley R, Watson S, Hinzen W (2018) The Language
  Profile of Formal Thought Disorder. *npj schizophrenia* 4:18.

Condray R, Steinhauer SR, Goldstein G (1992) Language Comprehension in
  Schizophrenics and Their Brothers. *Biological Psychiatry* 32: 790–802.

Condray R, Steinhauer SR, van Kammen DP, Kasparek A (1996) Working memory
  capacity predicts language comprehension in schizophrenic patients.
  Schizophrenia Research 20:1-13.

Condray R, Steinhauer SR, van Kammen DP, Kasparek A (2002)The language system
  in schizophrenia: Effects of capacity and linguistic structure. *Schizophrenia
  Bulletin* 28: 475–90.

Cuesta MJ, Peralta V (1995) Cognitive disorders in the positive, negative and
  disorganization syndromes of schizophrenia. *Psychiatry Research* 58: 227-235.

de Villiers JG (2007) The interface of language and theory of mind. *Lingua* 117: 1858–
  1878.

De Villiers, P., J. de Villiers (2012). Deception dissociates from false belief reasoning in
  deaf children. *British Journal of Developmental Psychology* 30, 188–209.

Dibben CRM, Rice C, Laws K, McKenna PJ (2008) Is executive impairment associated
  with schizophrenic syndromes? A meta-analysis. *Psychological Medicine* 39:
  381-392.

Docherty NM, Rhinewine J, Labhart R, Gordinier S (1998) Communication

Disturbances and Family Psychiatric History in Parents of Schizophrenic Patients. *The Journal of Nervous & Mental Disease* 186: 12, 761768.

Docherty NM, Gordinier SW (1999) Immediate memory, attention and communication disturbances in schizophrenia patients and their relatives. *Psychological Medicine* 29: 189-197.

Dungan J., Saxe R. (2012). Matched false-belief performance during verbal and nonverbal interference. Cognitive Science 36 (6), 1148-1156.

Durrleman S, Franck J (2015) Exploring links between language and cognition in autism spectrum disorders: Complement sentences, false belief, and executive functioning. *Journal of Communication Disorders* 54:15–31.

Durrleman S, Hinzen W, Franck J (2017) False belief and relative clauses in Autism Spectrum Disorders. *Journal of Communication Disorders* 74: 35-44.

Fioravanti M, Carlone O, Vitale B, Cinti ME, Clare L (2005) A meta-analysis of cognitive deficits in adults with a diagnosis of schizophrenia. *Neuropsychol Rev*. 15: 73-95. doi:10.1007/s11065-005-6254-9.

Fouget d'Arc, B., Ramus, F. (2011). Belief attribution despite verbal interference. Quarterly Journal of Experimental Psychology 64 (5), 975–990.

Fraser WI, King KM, Thomas P, Kendell RE (1986) The diagnosis of schizophrenia by language analysis. *British Journal of Psychiatry* 148: 275–278.

Frith C (1992) *The cognitive neuropsychiatry of schizophrenia*. Hove, UK: Psychology Press.

Hinzen W and M Sheehan (2011) Moving towards the edge. *Linguistic Analysis* 37: 405-458.

Hoffman RE, Sledge W (1988) An analysis of grammatical deviance occurring in spontaneous schizophrenic speech. *Journal of Neurolinguistics* 3: 89–101.

Holshausen K, Harvey PD, Elvevag B, Foltz PW, Bowie CR (2014) Latent semantic variables are associated with formal thought disorder and adaptive behavior in older inpatients with schizophrenia *Cortex* 55: 88e96.

Howard D, Patterson K (1992) *Pyramids and Palm Trees: A test of semantic access from pictures and words*. Bury St. Edmunds, UK: Thames Valley Test Company.

Kay SR, Fiszbein A, Opler LA (1987) The Positive and Negative Syndrome Scale for Schizophrenia. *Schizophrenia Bulletin* 13: 261–276.

Kuperberg GR, McGuire PK, David AS (1998) Reduced sensitivity to linguistic context in schizophrenic thought disorder: evidence from on-line monitoring for words in linguistically anomalous sentences. *J Abnorm Psychol*.107: 423-434. doi:10.1037//0021-843X.107.3.423.

Kuperberg GR (2010a) Language in schizophrenia Part 1: Introduction. *Language and Linguistics Compass* 4: 576-589.

Kuperberg GR (2010b) Language in schizophrenia Part 2: What can psycholinguistics bring to the study of schizophrenia...and vice versa? *Language and Linguistics Compass* 4: 590-604.

Landre NA, Taylor MA, Kearns KP (1992) Language functioning in schizophrenic and aphasic patients. *Neuropsychiatry, Neuropsychol Behav Neurol* 5:7–14.

Lelekov T, Franck N, Dominey PF, Georgieff N, Bernard ÂC (2000) Cognitive sequence processing and syntactic comprehension in schizophrenia. *Neuroreport* 11: 2145-2149.

Little B, Gallagher P, Zimmerer V, Varley R, Douglas M, Spencer H, Cokal D, Deamer, F, Turkington D, Ferrier IN, Hinzen W, Watson S (2018) Language comprehension in schizophrenia and aphasia: the relationship with non-verbal

cognition and thought disorder, under review.

Marini A, Spoletini I, Rubino IA, Ciuffa M, Bria P, Martinotti G, Banfi G, Boccascino
R, Strom P, Siracusano A, Caltagirone C, Spalletta G (2008) The language of
schizophrenia: an analysis of micro and macrolinguistic abilities and their
neuropsychological correlates. *Schizophr. Res.* 105: 144–155.

McKenna PJ, Oh TM (2005). *Schizophrenic Speech: Making Sense of Bathroots and
Ponds that Fall in Doorways*. Cambridge: Cambridge University Press.

Morice R, McNicol D (1985) The Comprehension and Production of Complex Syntax
in Schizophrenia. *Cortex* 21: 567-580.

Morice R, & McNicol D (1986) Language Changes in Schizophrenia: A Limited
Replication. *Schizophrenia Bulletin* 12: 239-250.

O'Leary DS, Flaum M, Kesler ML, Flashman LA, Arndt S, Andreasen NC (2000).
Cognitive correlates of negative, disorganized and psychotic symptom
dimensions of schizophrenia. *Journal of Neuropsychiatry and Clinical
Neuroscience* 12: 4-15.

Oh TM, McCarthy RA, McKenna PJ (2002) Is there a schizophasia? A study applying
the single case approach to formal thought disorder in schizophrenia. *Neurocase,*
8: 233–244.

Onishi, K. H., Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs.
Science 308 (5719), 255-258.

Pellicano E (2007) Links between theory of mind and executive function in young
children with autism: Clues to developmental primacy. *Developmental
Psychology* 43: 974–990.

Reichenberg A, Harvey PD (2007) Neuropsychological impairments in schizophrenia: Integration of performance-based and brain imaging findings. *Psychol Bull*. 133:833-858. doi:10.1037/0033-2909.133.5.833.

Rochester S, Martin R (1979) *Crazy Talk: A Study of the Discourse of Schizophrenic Speakers*. New York: Plenum Publishing.

Rodriguez-Ferrera S, McCarthy RA, McKenna, PJ (2001) Language in schizophrenia and its relationship to formal thought disorder. *Psychol. Med* 31: 197–205.

Sankoh AJ, Huque MF, Dubey SD (1997) Some comments on frequently used multiple endpoint adjustment methods in clinical trials. *Statistics in Medicine*, *16*: 2529–42.

Schaefer J, Giangrande E, Weinberger DR, Dickinson D (2013) The global cognitive impairment in schizophrenia: Consistent over decades and around the world. *Schizophr Res*. 150:42-50. doi:10.1016/j.schres.2013.07.009.

Sprong M, Schothorst P, Vos E, Hox J, van Engeland H (2007) Theory of mind in schizophrenia: Meta-analysis. *British Journal of Psychiatry* 191: 5–13.

Stirling J, Hellewell J, Blakey A, Deakin W (2006) Thought disorder in schizophrenia is associated with both executive dysfunction and circumscribed impairments in semantic function. *Psychol Med*. 36:475. doi:10.1017/S0033291705006884

Thomas P, King K, Fraser WI, Kendell RE (1990) Linguistic performance in schizophrenia: a comparison of acute and chronic patients. *British Journal of Psychiatry* 156: 204–210.

Thomas P, King K, Fraser WI (1987) Positive and negative symptoms of schizophrenia and linguistic performance. *Acta Psychiatrica Scandinavica*, 76(2): 144-151.

Tavano A, Sponda S, Fabbro F, Perlini C, Rambaldelli G, Ferro A, Cerruti S, Tansella M, & Brambilla P (2008) Specific linguistic and pragmatic deficits in Italian patients with schizophrenia. *Schizophr. Res.* 102**:** 53–62.

Varley, R., Siegal, M. (2000). Evidence for cognition without grammar from causal reasoning and 'theory of mind' in an agrammatic aphasic patient. Current Biology 10 (12), 723-726.

Wensing T, Cieslik EC, Müller VI, Hoffstaedter F, Eickhoff SB, Nickl-Jockschat T (2017) Neural correlates of formal thought disorder: An activation likelihood estimation meta-analysis. *Human Brain Mapping* 38: 4946-4965.

Wechsler D (2011) *Wechsler Abbreviated Scale of Intelligence - Second Edition (WASI-II)*. San Antonio, TX: NCS Pearson.

Zimmerer V, Watson S, Turkington D, Ferrier N, Hinzen W (2017) Deictic and propositional meaning – new perspectives on language in schizophrenia. *Frontiers in Psychiatry* 8:17 doi: 10.3389/fpsyt.2017.00017.

Zimmerer V, Varley R, Deamer F, Hinzen W. (in press). Factive and counterfactive interpretations in aphasia and their relationship with lexical, syntactic and conceptual capacities. *Journal of Neurolinguistics*.

Supplementary Table 1.

Independent t-test comparisons of age, education, and WASI IQ for neurotypical controls (NC), first degree relatives (FDR), and participants with (SZ+FTD) and without formal thought disorder (SZ-FTD)

Group comparisons	Age		Education		IQ/WASI	
	t (df)	p	t (df)	p	t (df)	p
NC vs. SZ+ FTD	.952 (28)	> .05	1.142 (28)	> .05	-5.763* (28)	.001
NC vs. SZ-FTD	1.836 (28)	> .05	-2.386* (28)	.024	-3.584* (28)	.001
NC vs. FDR	.042 (28)	> .05	-.330 (28)	> .05	1.151 (28)	>.05
FDR vs. SZ+ FTD	.952 (28)	> .05	1.403 (28)	> .05	-4.738* (28)	.001
FDR vs. SZ-FTD	1.787 (28)	.085	-2.568*(28)	.016	2.815* (28)	.009
SZ+ FTD vs. SZ-FTD	-2.825* (28)	.009	-1.324 (28)	> .05	1.008 (28)	>.05

[1] Sig. (2-tailed)

Supplementary Table 2.

A set of items used in the sentence-picture matching experiment.

Conditions	Sentence
Factive[1]	The mother knows that the child is naughty.
	The mother knows that the child is nice.
	The man knows that it is cold outside.
	The man knows that it is warm outside.
	The man knows that the dog is harmless.
	The woman knows that dinner is ready.
	The woman knows that dinner is not ready.
	The man knows that the bathroom is clean.
	The man knows that the dog is dangerous.
	The man knows that the bathroom is dirty.
	It is clear to the woman that the computer is fixed.
	It is clear to the man that the dog is small.
	It is clear to the woman that the stall is free.
	It is clear to the woman that the man is tall.
	It is clear to the woman that the man is small.
	It is clear to the woman that the stall is occupied.
	It is clear to the man that the pool is safe.
	It is clear to the woman that the computer is broken.
	It is clear to the man that the pool is dangerous.
	It is clear to the man that the dog is big.
Non-factive	The man thinks that the dog is dangerous.
	The man thinks that the bathroom is clean.
	The man thinks that the dog is harmless.
	The woman thinks that dinner is not ready.
	The woman thinks that dinner is ready.
	The man thinks that the bathroom is dirty.
	The mother thinks that the child is naughty.
	The man thinks that it is cold outside.
	The man thinks that it is warm outside.
	The mother thinks that the child is nice.
	It only seems to the man that the pool is safe.
	It only seems to the man that the dog is big.
	It only seems to the woman that the man is tall.
	It only seems to the woman that the stall is occupied.
	It only seems to the woman that the computer is fixed.

[1]The pictures and digital scoring sheet for the experiment can be found in the following link: https://www.cognitionandgrammar.net/s/Factivity-SPM.zip.

Supplementary Table 3a.

Mean and standard deviation in trials for neurotypical controls (NC), first degree relatives (FDR), and participants with (SZ+FTD) and without (SZ-FTD) formal thought disorder.

	NC	FDR	SZ-FTD	SZ+FTD
Trials	M (s.d)	M (s.d)	M (s.d)	M (s.d)
Know	.97 (.06)	.97 (.05)	.95 (.09)	.87 (.16)
Think	.96 (.06)	.94 (.14)	.88 (.19)	.54 (.38)
It only seems	.97 (.08)	.93 (.09)	.78 (.23)	.58 (.40)
It is clear	.98 (.04)	.98 (.04)	.96 (.06)	.81 (.18)

Supplementary Table 3b.

Performance comparisons of (SZ+FTD) with Wilcoxon tests.

Trial comparisons	Wilcoxon tests	p	r
Know vs. think	Z = -2.313	.0115*	.66
Know vs. it only seems	Z = -2.019	.0215*	.58
It is clear vs. think	Z = -2. 439	.008*	.70
It is clear vs. it only seems	Z = -1. 725	.047*	.50

*Mean difference significant at .05 level.
[1] p(1-tailed)

Supplementary Table 4.

Mean scores, Z-scores, and standard deviations of neurotypical controls (NC), first-degree relatives (FDR), participants with SZ and FTD (SZ+FTD) and without FTD (SZ-FTD) across tests (TROG, PPT, Brixton).

	NC		FDR		*SZ-FTD*		*SZ+FTD*	
	M (SD)	z (SD)	M(SD)	Z(SD)	M (SD)	Z (SD)	M (SD)	Z (SD)
Sentence comprehension (TROG)	18.62 (1.55)	.4981 (.441)	16.86 (3.53)	.2350 (1.451)	17.54 (1.808)	.1930 (.5122)	13.18 (5.334)	-1.04 (1.514)
Visual semantic memory (PPT)	50.92 (.862)	.2798 (.114)	48.820 (7.515)	-.2976 (1.952)	49.84 (1.724)	.1365 (.2294)	47.66 (4.334)	-.1535 .5767
Non-verbal reasoning (Brixton)	38.23 (7.212)	-.1647 (.039)	68.56 (184.106)	.2057 (.5596)	30.46 (5.980)	-.2069 (.0324)	21.18 (12.66)	.1676 (1.473)
WASI-IQ Score	106 (8.118)	-7.692 (1.00)	104.50 (9.802)	8.333 (1.00)	90.53 (18.20)	.000 (1.00)	77.75 (14.46)	-.3054 (1.117)
Factive condition	.973 (.0388)	7.692 (1.00)	.979 (.0257)	.000 (1.00)	.953 (.0593)	-1.538 (0.99)	.8417 (.1549)	-.8333 (0.95)
Non-factive condition	.964 (.0644)	2.307 (1.00)	.938 (.1153)	.000 (1.00)	.8512 (.1869)	-7.692 (0.99)	.5556 (.3508)	.000 (0.99)

Supplementary Table 5a.

Comparisons of effects size for means between clinical groups and controls

	NC vs. SZ-FTD	NC vs. SZ+FTD	NC vs. FDR	FDR vs. SZ-FTD	FDR vs. SZ+FTD
	Cohen [d]	Cohen [d]	Cohen [d]	Cohen [d]	Cohen [d]
TROG	0.64	1.38	0.64	0.24	-0.813
PPT	0.790	1.04	0.39	0.18	-0.189
Brixton	1.172	1.654	-0.232	-0.29	-0.363
WASI-IQ	1.097	2.40	0.16	-0.95	-2.16
Factive	0.39	1.16	-0.18	-0.56	-1.23
Non-factive	0.80	1.61	0.27	-0.55	-1.46

Supplementary Table 5b.

Comparisons of effects size for z-score means between clinical groups and controls

	NC vs. SZ-FTD	NC vs. SZ+FTD	NC vs. FDR	FDR vs. SZ-FTD	FDR vs. SZ+FTD
	Cohen [d]	Cohen [d]	Cohen [d]	Cohen [d]	Cohen [d]
TROG	-0.63	1.37	0.24	-0.03	-0.85
PPT	0.79	1.04	0.41	0.31	0.10
Brixton	1.117	-0.31	-0.93	-1.04	-0.03
WASI-IQ	-7.692	-6.967	-0.64	-8.333	-8.638
Factive	9.23	8.74	7.692	6.96	7.89
Non-factive	10.049	2.31	2.337	-10.24	-2.51