The neural bases of discourse semantic and pragmatic deficits in patients with frontotemporal dementia and Alzheimer's disease

*Title page

The neural bases of discourse semantic and pragmatic deficits in patients with frontotemporal dementia and Alzheimer's disease

S. Luzzi [1]*, S. Baldinelli[1], V. Ranaldi[1], C.Fiori[1], A. Plutino[1], F.M. Fringuelli[2], M. Silvestrini[1], G. Baggio[3^] and C. Reverberi[4,5^]

[1]Department of Experimental and Clinical Medicine, Polytechnic University of Marche, Italy

[2]Nuclear Medicine Department, Ospedali Riuniti di Ancona, Italy

[3]Language Acquisition and Language Processing Lab, Department of Language and Literature, Norwegian University of Science and Technology, Trondheim, Norway

[4]Department of Psychology, Università Milano – Bicocca, Italy

[5]NeuroMI, Milan Centre for Neuroscience, Italy

*Correspondence to:

Simona Luzzi, Neurologic Clinic, Department of Experimental and Clinical Medicine, Polytechnic University of Marche, Via Conca, 1, Torrette di Ancona, Italy; Tel. +39 071 596 4637, Fax. +39 071 887 262, e-mail: s.luzzi@univpm.it; simonaluzzi@yahoo.it

^GB and CR equally contributed to this work

Contributions: SL designed the research, SL SB VR CF, AP, FMF MS performed the research, SL GB CR analyzed the data and wrote the paper.

Declaration of interest: none

This research did not receiveany specific grant from funding agenciesin the public, commercial, or not-for-profit sectors.

Abstract. Neuropsychological research on language has largely focused on how the brain processes single words and sentences whose meaning does not depend on the context or on the intentions of the speaker. Fewer studies have investigated the neurobiological bases of discourse semantics and pragmatics in patients and healthy individuals. We studied discourse semantic and pragmatic skills in patients with behavioral variant frontotemporal dementia (bvFTD) or Alzheimer's disease (AD) in comparison to healthy controls. Our goal was to assess whether and how the two patient groups differ in their cognitive and behavioral profiles, and whether these differences may be traced back to disease-specific patterns of neuronal hypometabolism. We combined PET imaging with standard neuropsychological assessment tools and a dedicated test battery designed to evaluate discourse semantics and pragmatics in patients with brain lesions or neurological disorders. We found that AD and bvFTD patients were both impaired compared to controls in discourse comprehension, but largely spared in single word comprehension. Importantly, we also found evidence for behavioral impairments specific to each disease, associated with different brain damage patterns. Compared to AD and controls, bvFTD patients had, behaviorally, more difficulty in evaluating whether certain inferences follow from discourse and in identifying humorous completions of stories; neurally, they had greater damage to medial and lateral regions of PFC. AD patients showed a different pattern of errors in a humor comprehension task than bvFTD patients and controls, and they showed greater posterior temporal and parietal cortical depletion. Both groups had comparable difficulties with understanding idioms and indirect requests. Finally, bvFTD-specific errors were correlated with the severity of hypometabolism in bvFTD. We discuss these results in light of previous research on the dementias as well as consequences for models of semantics and pragmatics in the brain.

Running title: Discourse semantics and pragmatics in bvFTD and AD patients.

Keywords: Neuropsychology; Alzheimer's disease; Behavioral variant frontotemporal dementia; Language; Semantics; Discourse; Pragmatics; Idioms; Inference; Humor

1. Introduction

Understanding everyday language requires more than understanding single words and the way words are arranged in phrases and sentences (Culicover & Jackendoff 2006; Pylkkänen & McElree 2006; Hagoort & van Berkum 2007; Baggio et al. 2012, 2016). Two further factors are normally involved. First, the discourse context. For example, the sentences 'Max fell. John pushed him.' highlight a causal link between John's push and Max's fall, which is not conveyed by these sentences individually. Second, the intentions of the speaker. For example, 'Peter is turbocharged' could mean that Peter is lazy, if the sentence is uttered sarcastically. These are only two examples of a broader range of phenomena analyzed by discourse semantics and pragmatics. Previous research in neuropsychology and neurolinguistics has mostly focused on how the brain processes single words and sentences whose meaning does not depend on the context or the intentions of the speaker. Fewer studies have investigated the neurocognitive bases of discourse semantics and pragmatics in healthy individuals *vs* neurological patients (Hagoort & Indefrey 2014; Hagoort & Levinson 2014). The present study contributes to filling this gap. We conducted a systematic investigation of these abilities in an Indo-European language (Italian) in patients with Alzheimer's disease (AD) or frontotemporal dementia (FTD).

The human ability to recover the meaning of linguistic expressions in context may be damaged or lost as a result of brain injury or disorder (Shallice & Cooper 2011; Faust 2012). The cognitive effects of various neurodegenerative diseases are being increasingly studied, and language and semantics are among the prominent topics of research. Neurodegeneration is often accompanied by impairments of language or communication (Bonner et al. 2010; Mesulam et al. 2014; Grossman 2018). The consequences of brain injury or neurodegenerative processes for semantics at the single word or lexical concept level are relatively well understood (Patterson et al. 2007; Chen et al. 2017), but less is known about the effects of lesions or atrophy on language and semantics beyond single words. Neurodegenerative diseases can impair higher-level discourse semantics and pragmatics (e.g., Bambini et al. 2016 and Baggio et al. 2016 on ALS patients; Roberts et al. 2017; Stemmer 2017). Here, we focus on patients with behavioral variant frontotemporal dementia (bvFTD) or Alzheimer's disease (AD). A diagnosis of possible bvFTD (Rascovsky et al. 2011) requires the presence of 3 out of 6 discriminating features (disinhibition, apathy or inertia, loss of sympathy or empathy, perseverative or compulsive behaviours, hyperorality, and dysexecutive neuropsychological profile), while clinical criteria for probable bvFTD include also functional disability and characteristic cortical degeneration, as revealed by neuroimaging (details below). AD progresses through three main stages (Dubois et al. 2007, 2010, 2014): a preclinical stage with few or no symptoms; a stage with mild cognitive impairment (MCI); and a final stage with classic dementia symptoms (e.g., memory loss and word-finding difficulties).

We assess the ability of AD and bvFTD patients to extract meaning beyond single words. We explore the following four pervasive discourse semantic and pragmatic phenomena: (1) comprehension of idiomatic expressions, (2) discourse inferences, (3) comprehension of indirect requests and (4) humor comprehension. The overall goal is to assess whether AD and bvFTD impact these phenomena differently. This has clinical *and* theoretical implications. From a clinical point of view, the study of

AD- and bvFTD-specific impairment profiles could assist patient management and improve differential diagnosis. From a theoretical stance, the study of discourse semantics and pragmatics in AD and bvFTD could shed new light on how the brain processes meaning. Differences between AD and bvFTD patients may elucidate the functional roles in discourse semantics and pragmatics of posterior brain regions (e.g., temporo-parietal cortices, which are often more damaged in AD than bvFTD) and anterior regions (e.g., the lateral and medial PFC, which present, instead, more atrophy in bvFTD than AD; for network-level and connectivity analysis, see Pievani et al. 2014; for a longitudinal study of (sub)cortical degeneration in AD and bvFTD, see Landin-Romero 2017). These brain regions have been implicated in figurative and pragmatic language processing by neuroimaging experiments (Vrticka 2013; Ferstl et al. 2008; Bambini 2010; Bohrn et al. 2012; Rapp et al. 2012; Hagoort & Indefrey 2014; Hagoort & Levinson 2014). Additionally, bvFTD is an appropriate lesion model to study discourse semantics and pragmatics, also because of known isolated deficits in social cognition and executive function in bvFTD that may affect performance, against largely spared single word and sentence comprehension. AD, on the other hand, may allow researchers to examine the contributions of episodic memory or word finding difficulties to comprehension.

1.1. Discourse semantics and pragmatics in the dementias

Idioms (e.g., ‘To kick the bucket’), much like conventional metaphors (e.g., ‘Time is money’), are expressions whose meaning cannot be derived from the meanings of the words involved, plus the phrasal or sentential syntax, nor can meaning always be inferred from context. Idiomatic meaning is instead recovered or retrieved from memory, like the meanings of single words and of conventional metaphors. Several proposals have linked specific brain structures or processes to the comprehension of figurative (e.g., metaphoric, idiomatic etc.) meaning. Right-hemisphere theories (inspired by Giora 1997 and Jung-Beeman 2005) have found some initial support in brain imaging studies (e.g., Mashal & Faust 2008; Mashal et al. 2005, 2007), in particular for novel metaphors (see Baggio 2018 and Vulchanova et al. 2019 for a discussion of relevant dimensions, including conventionality and decomposability, in figurative language processing). Recent meta-analyses of PET and fMRI studies (Rapp et al. 2012; Bohrn et al. 2012) indeed show RH involvement in processing novel metaphors. However, no RH advantage was found for metaphors and idioms, in general. Figurative language appears to engage primarily the same LH networks as literal (or compositional) language, only to a greater extent, in particular the left inferior frontal gyrus (LIFG), left middle frontal gyrus (MFG), left inferior parietal lobule (IPL), medial prefrontal cortex (mPFC), and large parts of the temporal lobe. RH regions, although present, account for a minority of all activation foci. Idioms, in particular, engage the left IFG (BA44/45) and left MTG (BA21) (Rapp et al. 2012; Bohrn et al. 2012, Table 3 and section 3.2).

The IFG operates in concert with temporal regions (posterior middle and superior temporal gyri, pMSTG) in the service of controlled retrieval of lexical meaning and unification (Hagoort et al. 2009; Baggio 2018). Damage to the temporal cortex or IFG or both should result in impaired processing of idioms. These regions are often damaged in AD and bvFTD. Therefore, both patient groups are expected to exhibit impaired performance in tasks that require comprehension of idiomatic meaning.

Earlier studies on idioms and metaphors in AD have shown that comprehension of non-literal language can be relatively preserved in the early stages of the disease (Papagno 2001; Amanzio et al. 2008; but see Kempler et al. 1988). Newer research suggests that idiomatic meaning may still be available to AD patients. Thus, poorer performance on comprehension of idiomatic meaning, at least in some tasks (e.g., sentence-to-picture matching), should be attributed to the patients' difficulty with suppressing the idiom's literal interpretation (Papagno et al. 2003; Rassiga et al. 2009). Additionally, impaired proverb interpretation—which requires retrieval of 'frozen meanings' from memory, especially for common or familiar proverbs—is associated in bvFTD and AD with loss of cortical volume in left anterior temporal areas (Rapp & Wild 2011; Kaiser et al. 2013).

The second area of interest here is discourse inference. Also in this case, previous work points to impairments in the dementias. For example, Chapman et al. (1998) showed that AD patients have difficulty drawing inferences that link the content of discourse to world knowledge. Patients with AD have specific difficulty integrating relations during reasoning (Waltz et al. 2004). Spotorno et al. (2015) investigated quantifiers and scalar implicatures (e.g., for 'some', meaning 'some but not all') in bvFTD, and reported that the patients' tendency to restrict their interpretations to logical meaning was correlated with atrophy in ventromedial PFC (mPFC). These patients have difficulties in some other tasks requiring inferences, such as Theory of Mind (ToM), again linked to degeneration of regions of the mPFC (Adenzato et al. 2010). Studies have also documented reasoning deficits in bvFTD patients, for example with transitive inferences (A>B; B>C; A>C), when the premises were scrambled, but not when they were chained (Waltz et al. 1999), or when materials involved familiar notions or information about which patients have definite beliefs (Vartanian et al. 2009). Similar patterns were observed for analogical reasoning, where patients with frontotemporal degeneration show poorer performance than controls when the correct relational answer is not supported by the structure of a problem's layout (e.g., perceptual cues) (Morrison et al. 2004). Research on other discourse processes, such as bridging inferences, has largely focused on patients with right-hemisphere damage and not, to our knowledge, on AD or bvFTD.

Observational and clinical studies have reported deficits in pragmatic processes in AD and bvFTD, including understanding and producing indirect requests or replies (Roberts et al. 2017; Guendouzi et al. 2017). Previous research on indirectness has only involved RH damaged patients (e.g., see Foldi 1987), but recent brain imaging experiments leads us to expect deficits in both AD and bvFTD patients in tasks that involve interpretation of indirect replies or requests. For example, Bašnáková et al. (2013) reported that the mPFC and right temporo-parietal junction (TPJ) are more active when an utterance (e.g., 'It is hard to give a good presentation') follows a context in which it serves as an indirect reply ('Did you like my presentation?') *vs* a context relative to which it serves as a direct reply ('How hard is it to give a good presentation?'; for discussion, also of neurocognitive models focusing on the mPFC and TPJ, see Hagoort & Indefrey 2014; see also Hagoort & Levinson 2014, Catani & Bambini 2014, Baggio 2018). These cortical areas involved in pragmatics are often damaged in the dementias (in particular, mPFC and neighboring frontal cortex in bvFTD, TPJ and adjacent cortex in AD), so one might expect impaired processing of requests or replies in AD and bvFTD.

Comprehension of humor is the fourth area of interest. Behavioral-variant bvFTD patients may be impaired in tasks that require taking the perspective of others or coordinating with them (Grossman 2018). Healey et al. (2015) showed that bvFTD patients cannot easily convey an object's descriptions to a conversational partner and that impaired performance is related to gray matter atrophy in medial frontal (mPFC) and orbitofrontal cortex. Studies found that bvFTD patients have difficulty establishing focal points in coordination games (McMillan et al. 2011) and are less sensitive to ambiguities in quantified sentences (McMillan et al. 2013). In addition, studies have shown that bvFTD patients are often impaired in their interpretation of humorous language, cartoons, story vignettes, and other material (Snowden et al. 2003; Clark et al. 2015; for case-study data, see Rahman et al. 1999). There are no systematic investigations of humor, irony, and sarcasm in AD, and the imaging literature gives us reasons to believe that only patients with predominantly frontal lobar degeneration will be impaired in tasks that either involve or require humor comprehension. The meta-analysis by Bohrn et al. (2012) shows that humor, irony, and sarcasm processing engage the mPFC and anterior cingulate cortex, together with frontal and temporal areas of the right hemisphere. One fMRI experiment has implicated the left TPJ and the superior frontal gyrus (SFG) in humor processing (Vrticka et al. 2013; Campbell et al. 2015), but the left TPJ appears to be primarily involved in detecting humour in visual stimuli (see Vrticka et al. 2013, Table 1).

1.2. The present study

We report a systematic investigation of discourse semantic and pragmatic deficits in bvFTD and AD, using a combination of neuropsychological testing, PET imaging, and a standardized battery for assessing language comprehension skills (idioms, inferences, indirect requests, humor) in neurological patients (Rinaldi et al. 2006). AD and bvFTD patients might show quantitatively similar patterns of impairment across tasks, but they may differ in the types of errors made, and in the frequency of different error types. Error patterns may also be associated with atrophy in the affected cortical regions in bvFTD and AD, possibly leading to poorer performance in bvFTD patients in tasks that engage more frontal areas, and poorer performance in AD patients in tasks that engage more posterior areas (details below). Our aim here is to assess, through a structured investigation, different levels of impairment between AD and bvFTD. Based on previous research and theory (for reviews, see Seeley 2008; Seeley et al. 2009), we expect the following patterns:

(1) Idiom comprehension, involving the retrieval of stored form-meaning pairings, may be impaired to a similar degree in AD and bvFTD, as both conditions involve damage to temporal lobe regions known to support semantic memory access and retrieval (Davies et al. 2005) and, to a lesser extent, to inferior frontal areas known to contribute to such processes through forms of maintenance and control.

(2) Inference based on linguistic materials relies primarily on regions of the lateral and medial frontal cortex (Monti et al. 2009; Prado et al. 2010; 2011; Reverberi et al. 2007; 2009; 2010; 2012). Although discourse inferences, as studied here, may differ from inference types examined in earlier research, we hypothesize that they would engage areas of the PFC, to a greater extent than the temporal cortex. bvFTD

patients often show greater damage to the PFC (medially and laterally) compared to AD patients (Pievani et al. 2014; Landin-Romero 2017). They should, therefore, have more difficulty in discourse inference tasks than AD patients and controls.

(3) Indirect requests or replies engage frontal (medial PFC) and temporo-parietal (TPJ) cortices; these areas are often damaged in bvFTD patients (frontal cortex, in particular) and in AD patients (temporo-parietal regions, specifically); if these two nodes are equally important functionally, the performance of both patient groups should be impaired relative to controls in tasks requiring pragmatic interpretation of indirect requests or replies.

(4) In healthy individuals, understanding humorous stimuli activates primarily the frontal cortex bilaterally, including the mPFC and adjacent cortex; bvFTD patients should, therefore, have the most difficulty with these materials.

These predictions are largely based on imaging studies in healthy individuals, and to a lesser extent on studies in neurological patients. Our aim here is to contribute to filling a gap by testing and refining these hypotheses using data from bvFTD and AD patients. In brief, we expect to find different levels of impairment in AD and in bvFTD patients, in tasks that involve brain regions that are more severely damaged in one pathological condition than the other. Specifically, we expect that inference and humor processing are more impaired in bvFTD, due to extended PFC damage in these patients. But otherwise, we expect similar impairment in tasks that either involve brain regions disrupted to similar degrees in AD and bvFTD (i.e., in idiom comprehension, which relies largely on temporal areas), or that recruit distributed networks, such that at least one node is damaged in each condition (processing of indirect requests or replies, which recruits frontal and temporo-parietal areas).

2. Methods

We report how we determined our sample size, all data exclusion, all inclusion/exclusion criteria, whether inclusion/exclusion criteria were established prior to data analysis, all manipulations, and all measures in the study.

2.1. Participants

Three groups of participants were enrolled in this study: Alzheimer's disease (AD) patients; patients with behavioral variant frontotemporal dementia (bvFTD); and healthy controls. Inclusion or exclusion criteria were set prior to data analysis. All AD and bvFTD patients included in the study were in the mild stage of the disease, according to their general cognitive condition as measured using the Mini Mental State Examination test (i.e., MMSE score ≥ 18/30; Folstein et al. 1975; Tombaugh & McIntyre 1992; Reisberg et al. 2011). All the patients had a history of fewer than two years of reported cognitive or behavioral difficulties. We excluded patients reporting difficulties in everyday language comprehension and patients with diagnosed deficits pertaining to phonology, syntax, or semantics. We included patients with the limbic variant of AD (McKhann et al. 2011), excluding the visual and language variants of AD (sections 2.2, 3.1; Table 1).

A total of 81 participants were enrolled: 24 patients with bvFTD, 25 AD patients, and 32 healthy individuals. The sample size was determined to ensure a statistical power of ≥0.8 for a two-sample *t*-test, assuming that cortical damage would cause large deficits (effect size d≥0.8). This goal would be achieved by groups equal to or larger than 21 subjects. Post-hoc, the achieved sensitivity (effect size for achieving statistical power =0.8) was d=0.72 for the comparisons between FTD and AD, and d=0.68 for the comparisons between pathological groups and controls. Diagnosis of bvFTD or AD (for criteria, see Neary et al. 1998; Dubois et al. 2007; McKhann et al. 2011; Rascovsky et al. 2011) was performed by a neurologist in our team with expertise in dementia (author SL) and was based on the patient's recent clinical history, on neurological and neuropsychological assessment, and on PET imaging (details below). Healthy controls were recruited among the patients' relatives. The bvFTD patients (15 males and 9 females), AD patients (15 males and 10 females), and controls (18 males and 14 females) had similar ages and levels of education. Mean age was 64.62 years (SD=8.72) in the bvFTD group; 69.36 (4.02) for AD; and 66.37 (6.35) in the control group. Mean education for bfFTD was 9.08 years (4.08); 9.76 (4.36) for AD; 9.21 (4.65) for controls. Kruskal-Wallis and χ^2 tests revealed no differences in these variables between groups. The study procedures and analyses were not pre-registered, and participant confidentiality precludes public archiving of the data. The data may be accessed upon request to the Scientific Committee of the Neurology Clinic at Marche Polytechnic University (m.silvestrini@univpm.it). Interested readers will be required to fill a "Collaboration Statement".

Participants gave their written informed consent to take part in the study, which was conducted in full compliance with the Declaration of Helsinki.

2.2. General neuropsychological assessment

The neuropsychological assessment (Table 1) of bvFTD and AD patients included the tests: MMSE (Folstein et al. 1975); digit span (Caltagirone et al. 1979); Corsi block-tapping test (Spinnler & Tognoni 1998); copy and immediate recall of the Rey-Osterrieth figure B (Luzzi et al. 2011); apraxia tests battery (De Renzi et al. 1980; De Renzi & Lucchelli 1988); phonological and semantic fluency (Caltagirone et al. 1979); Luria's motor sequences (Luzzi et al. 2010); Stroop test (Caltagirone et al. 1979); easy naming, word reading, word-to-picture matching tests (Snowden et al. 2004); verbal and non-verbal versions of the Pyramid and Palm Trees Test (PPTT; Howard & Patterson, 1992); Poppelreuter-Ghent's overlapping figures test (Lezak et al. 2012); visual object and space perception test (VOSP, screening task) (Warrington and James, 1991). We excluded all patients with diagnosed language deficits or with self-reported difficulties in language use. As a partial assessment of receptive grammar, we presented patients with a reduced version (15/60 items) of the sentence comprehension subtest of the BADA aphasia battery (Batteria per l'Analisi dei Deficit Afasici; Miceli et al. 1994), and we included in the final sample only patients who made at most 2 errors in this task.

2.3. Discourse semantics and pragmatics battery

To assess the patients' and controls' discourse semantic and pragmatic abilities we used the battery BLED in Italian (Rinaldi et al. 2006), standardized using a sample of 50 patients with right hemisphere (RH) lesions and 39 controls (age range 29-86; mean education 9 years; see Rinaldi et al. 2004, 2006 for details). BLED was intended as a battery for testing language skills in RH-lesioned patients, in the sense that (a) it was not designed to test residual language skills in aphasic patients (e.g., patients with LH lesions; the present bvFTD and AD samples did not include patients with a diagnosis of aphasia) and (b) it was designed to test discourse semantic and pragmatic skills that, until the late 1990s, were assumed to be subserved by RH regions (see 1.1 above for a more nuanced picture). Some alternatives to BLED are the standardized batteries ABaCO (Sacco et al. 2008) and APACS (Arcara & Bambini 2016). These include tests of pragmatics and figurative language, as well as other verbal or non-verbal communication skills, but they lack a task exclusively devoted to discourse level inference, which is available in BLED. This justifies our preference for BLED, given our aim to investigate both discourse semantic and pragmatic comprehension in bvFTD and AD. We used 4 BLED tests, assessing: comprehension of idioms, discourse inference, request comprehension, and humor comprehension. The stimuli have comparable length and grammatical complexity across subtests and items (10 per subtest). BLED was administered in one hospital session, within at most one month from the PET session. The battery was administered by trained staff (authors SB and VR). There was no time limit for participants to provide a response, and the responses were not timed. Participants could view each test item in its entirety (story and responses, in randomized order across items or trials) at all times during the test.

2.3.1. Idiom comprehension

BLED includes two subtests of figurative language comprehension. In the first test, ('metafore figurate'), each item includes one figurative expression and four images showing the correct figurative interpretation, the literal interpretation, and two alternative (figurative and literal) interpretations. We did not use this first subtest because we were mainly interested in the comprehension of verbal stimuli. In the second subtest ('metafore scritte'), each item includes a single sentence containing an embedded idiomatic phrase, followed by three sentences expressing the literal interpretation (LI) of the first sentence, the correct idiomatic interpretation (II), and a neutral interpretation (NI). Rinaldi et al. (2006) refer to these test stimuli as 'metaphors'. However, they do recognize that "this kind of metaphor is more akin to an idiomatic expression" (Rinaldi et al. 2004), due to the conventional nature of the meaning-form pairing.

The subtest was administered as follows. Ten sentences ('Item n') containing an idiomatic phrase are shown in written form to each participant one by one. At the same time, the participant is also presented with three sentences (LI, II, NI), which are supposed to explain the meaning of the first item. The participant is asked to read aloud all four sentences (Item n, LI, II, and NI) and to indicate the one, among LI, II, and NI, providing a "correct interpretation" of the first item. The participant is told that only one of the three probe sentences is correct. If they point to more than one answer, they are asked to choose only one. One point is scored for every correct answer (II: range 0-10), and errors are classified as either literal (LI; range 0-10) or neutral answers (NI: range 0-10).

2.3.2. Discourse inferences

The discourse inferencing task (Brownell et al. 1986) is based on a set of ten short stories of two sentences each (e.g., 'The postman walked towards the mailbox. For several weeks he had waited in vain for that letter.') Each participant is required to read both sentences aloud. Next, the participant is presented with three sentences, classified as (a) a true conclusion (TC; a sentence expressing the intended meaning of the short story; e.g., 'The postman checked whether the letter he was waiting for had arrived'), or (b) a false conclusion (FC; a sentence expressing an unwarranted inference, e.g., 'The postman delivered the mail', which should be inhibited by the participant: the postman usually delivers the mail, but in this case he is waiting for the letter), or (c) factual information (FI; a sentence that expresses a true or false factual proposition from the vignette, e.g., 'The postman approached the mailbox'). Participants are shown these sentences one by one and are then requested to say if each sentence is true or false. Participants' answers are scored separately for each response type—TC (range 0-10), FC (0-10), and FI (0-10)—as incorrect (0 score) if the participant responds that a TC is false, or that an FC is true, or if an FI sentence is incorrectly judged true or false; responses are scored as correct (1) otherwise.

2.3.3. Request comprehension

In the indirect requests subtest, the participant is asked to read aloud each of 10 short stories; the stories are complete vignettes, but lack the final sentence (e.g., 'A boy meets a girl at a party and falls in love with her. When he has to leave and say goodbye he asks the girl: "Could you give me your phone number, please?" And the girl answers:'). Next, three sentences are shown simultaneously to the participant, who is asked to read them aloud and point to the correct reply. One sentence is the reply that would be given by a hearer who (correctly) understands the question in the vignette as an indirect request (IR: 'My phone number is 1587932'). The other two sentences are a reply to the same question interpreted as a direct request (DR: 'I know it by heart') and an unrelated reply (UR: 'I could give you a lift'). One point is scored for each correct answer (IR: range 0-10). We considered errors the cases in which participants choose a DR answer or a UR answer (range 0-10).

2.3.4. Humor comprehension

In the humor comprehension subtest, participants are presented with a short joke without the final sentence and punchline (e.g., 'The doctor says to a senior patient: "Dear Madam, the pain in your right arm is not a serious problem. It is simply due to old age." Then the woman says:'). The participant is asked to read aloud the joke and three sentences: a humorous conclusion and punchline (HC: 'This is nonsense, doctor. My left hand has the same age but it is healthy.'); a coherent distractor, i.e., a non-humorous sentence that fits with the context (CD: 'There is no cure for old age ailments!'); and a surprising distractor without a logical connection with the story (SD: 'The government fell today'). The participant is asked which sentence makes the vignette a joke. One point is provided for each correct answer (SC: range 0-10). Errors are classified as coherent distractor answers (CD: range 0-10) or surprising distractor answers (SD: range 0-10).

2.4. PET image acquisition

PET (positron emission tomography) images were acquired using a GE Discovery PET/CT 690 VCT scanner, with spatial resolutions of 1.17 mm, 1.17 mm, and 3.27 mm at full-width-half-maximum (FWHM) in radial, tangential, and axial directions, respectively. The participant's head was fixed to a head holder and a plastic spacer to minimize head movements. The participants received an average dose of [^{18}F] 2-deoxy-2-fluoro-D-glucose (FDG) tracer of 185 ± 10MBq (5±0.27mCi). PET image acquisition started about 30 to 40 minutes after FDG injection. Three-dimensional scanning was performed for 15 minutes (in list mode). Post-injection transmission scanning was conducted using CT for tissue attenuation correction (helical full 0.6 sec., 3.75 mm, 47 slices, 120kV, 250mA). The PET images were corrected for tissue attenuation and reconstructed into 256×256 matrices based on an Iterative Reconstruction Algorithm: VPFX-S (DFOV 30cm, VUE Point FX, 3.2mm, 32/8).

2.5. Statistical analysis

2.5.1. Behavioral data analysis

Neuropsychological and behavioral data were analyzed using non-parametric tests due to non-normal data distributions. Kruskal-Wallis and Mann-Whitney-Wilcoxon tests were employed to test the null hypothesis that the samples are from identical populations.

2.5.2. PET data analysis

Imaging analysis was performed using SPM12. Pre-processing included rigid-body transformation (realignment) to correct for head movement. The PET images were then normalized to the MNI space by means of non-linear warping as implemented in SPM12, smoothed with a Gaussian filter of 12 mm full-width at half maximum (FWHM) to increase the signal-to-noise ratio and to facilitate group-level analyses (similar results were obtained using 8mm FWHM). Individual global counts were finally normalized via proportional scaling to a mean value of 50 mg/100 mL/min. The resulting linear contrasts were then used to test for region-specific differences between groups, producing t-statistic maps in MNI space. We explored which brain regions show a reduced FDG uptake in each patient group as compared to healthy subjects, and the brain regions specifically damaged in bvFTD or AD when the two patient groups are compared against each other. The direct comparison between pathological groups is particularly important, because it allows us one identify the brain regions that are more likely to cause the behavioral impairments specific to each group. We considered significant effects at $p<0.05$ (alpha) FWE corrected for multiple comparisons at the cluster level. Clusters were identified by means of a voxel-level threshold at $p<0.001$, not corrected for multiple comparisons.

We examined correlations in patients between PET signal loss and performance in the BLED tasks. We used a region-of-interest (ROI) approach to reduce the severity of multiple comparison corrections of whole-brain approaches and to increase the power of the analysis. We considered two ROIs: one ROI including all brain regions showing significant signal loss in bvFTD patients, and one ROI including all brain regions showing significant signal loss in AD patients (Fig. 2). Correlations were examined in each patient group using the relevant ROI, i.e., we used the bvFTD-ROI for bvFTD patients and AD-ROI for AD patients. We further tested for correlations considering the entire patient group, to increase the sample size and therefore the robustness of the correlation estimation. Collapsing patient groups is appropriate when it may be safely assumed that the only variable affecting behavior is location of brain damage, not the pathology that caused it. We were able to analyze a large group for a relatively rare focal dementia. Yet, we are aware that, from a statistical point of view, the robustness of the correlation estimation would benefit from the availability of a larger group (Schönbrodt & Perugini, 2013). In several BLED tests, we noticed a reduction in the coverage of the full range of the behavioral variables, due to either ceiling or floor effects. This might reduce the reliability of correlation analyses. Therefore, we excepted from the correlation analyses any variables with severely reduced coverage of the measurement range: we excluded those variables where 90% of the patients covered less than 50% of the range; these variables are literal interpretation and neutral interpretation in the idiom comprehension task, direct request reply and unrelated reply in the request comprehension task, and surprising distractor in the humor comprehension task (Fig. 1). The scripts used in the neuroimaging analyses are available as Supplementary Material.

3. Results

3.1. Neuropsychological results

The standard neuropsychological tests (Table 1) show that AD and bvFTD patients are matched for general cognitive abilities, perceptual-spatial skills, ideomotor and constructional praxis and, crucially, language skills. There are only two domains in which the two groups differ (Table 1): AD patients present poorer memory skills than bvFTD patients, and bvFTD patients show poorer executive functions than AD patients in the Stroop test, the Brixton test, and Raven's progressive matrices. These findings are largely consistent with current knowledge of bvFTD and AD. We did not systematically and specifically assess the patients' grammar skills here, but we excluded bvFTD or AD patients with diagnoses of language disorders (see 2.1). Moreover, all BLED stimuli were comparable across tasks in terms of grammatical or syntactic complexity (see 2.3). Therefore, any differences in performance across

BLED tasks between the two patient groups cannot be attributed to (undetected) grammar deficits in one patient group *vs* the other.

3.2. Behavioral results

3.2.1. Analysis of correct responses

Kruskal-Wallis rank-sum tests on correct responses in each task revealed effects of group (3 samples: AD, bvFTD, controls) in idiom comprehension ($\chi 2=39.4$; df=2; $p<0.001$), discourse inferences ($\chi 2=20.74$; df=2; $p<0.001$), request comprehension ($\chi 2=27.97$; df=2; $p<0.001$), and humor comprehension ($\chi 2=40.55$; df=2; $p<0.001$).

Pairwise comparisons between groups by means of Mann-Whitney-Wilcoxon tests on *correct responses* produced the following results. In the idiom comprehension task, bvFTD patients gave fewer correct responses (idiomatic interpretations, II) than controls (W=101.5; $p<0.001$), and so did AD patients (W=44.5; $p<0.001$). Yet, there was no difference between AD and bvFTD patients (W=245; p=0.27; Fig. 1a). In the discourse inference task, bvFTD patients produced fewer correct responses (true conclusions, TC) than the controls (W=136; $p<0.001$), and so did AD patients (W=197; p=0.002), but there was no difference between AD and bvFTD (W=204.5; p=0.08; Fig. 1d). In the request comprehension task, bvFTD patients gave fewer correct responses (indirect request replies, IR) than controls (W=103; $p<0.001$), and the same holds for AD patients (W=130; $p<0.001$). There were no differences between AD and bvFTD (W=226.5; p=0.55; Fig. 1g). In the humor comprehension task, bvFTD patients gave fewer correct responses (humorous conclusions, HR) than controls (W=53; p<0.001), and so did AD patients (W=83.5 p<0.001; bvFTD vs AD, W=175 p=0.03; Fig. 1j).

In conclusion, when considering correct responses, we found the same pattern for all four subtests of the BLED battery (Figs. 1a, 1d, 1g, 1j): bvFTD and AD patients perform worse than controls (they produce fewer correct responses) and there are no differences between the two patient groups. Clear differences between bvFTD and AD patients emerge only when one analyzes the errors in the BLED subtests.

3.2.2. Analysis of errors

Kruskal-Wallis tests revealed an effect of group in all types of errors we analyzed: in idiom comprehension (LI: $\chi 2=27.5$; df=2; p<0.001; NI: $\chi 2=25.6$; df=2; p<0.001), discourse inferences (FC: $\chi 2=23.81$; df=2; p<0.001; FI: $\chi 2=26.19$; df=2; p<0.001), requests (DR: $\chi 2=20.39$; df=2; p<0.001; UR: $\chi 2=22.26$; df=2; p<0.001), and humor (CD: $\chi 2=21.54$; df=2; p<0.001; SD: ($\chi 2=43.83$; df=2; p<0.001).

We further performed pairwise comparisons on errors between groups by means of Mann-Whitney-Wilcoxon tests. In the idiom comprehension test, bvFTD patients produced more literal interpretation errors (LI) than controls (W=609; p<0.001), and the same applies to AD patients (W=691.5; p<0.001). There was no difference between AD and bvFTD patients (W=344; p=0.37; Fig. 1b). The same pattern was found for neutral interpretation errors (NI): bvFTD patients made more NI errors than controls (W=587.5; p<0.001), and so did AD patients (W=689; p<0.001), with no difference between groups (W=275; p=0.61; Fig. 1c). For idiom comprehension, neither the analysis of correct responses nor the analysis of errors discriminates between AD and bvFTD. In this task, AD and bvFTD patients perform similarly.

In the discourse inferencing task, bvFTD patients had worse scores than controls. They made more errors in evaluating false conclusions as false (FC, Fig. 1e; that is, they considered them to be *true*) than controls ($W=78.5$; $p<0.001$) and AD patients ($W=141$; $p=0.002$). There were no differences between AD patients and controls in the frequency of FC errors ($W=350$; $p=0.57$; Fig. 1e). Therefore, error patterns for false conclusions discriminate between AD and bvFTD. In factual errors (FI), both bvFTD and AD patients had worse scores than controls (FTD: $W=184$; $p=0.0005$; AD: $W=104.5$; $p<0.001$; bvFTD *vs* AD, $W=188.5$; $p=0.038$; Fig. 1f).

In the request comprehension task, bvFTD patients produced more direct request (DR) errors than controls ($W=586$; $p<0.001$), and so did AD patients ($W=495.5$; $p<0.001$; bvFTD *vs* AD, $W=327$; $p=0.08$; Fig. 1h). Patients with bvFTD produced more unrelated reply (UR) errors that control participants ($W=479$; $p=0.001$). The same holds for AD patients ($W=606$; $p<0.001$). No difference was found between the bvFTD and AD groups ($W=299$; $p=0.27$; Fig. 1i).

In the humor comprehension task, bvFTD patients made more coherent distractor (CD) errors than controls ($W=588$; $p<0.001$) and AD patients ($W=438$; $p=0.0004$). There were no differences between AD and controls ($W=455.5$; $p=0.3299$; Fig. 1k). Though bvFTD patients made more surprising distractor (SD) errors than controls ($W=585$; $p<0.001$), AD made more such errors than bvFTD ($W=153.5$; $p=0.0081$) and controls ($W=770.5$; $p<0.001$; Fig. 1l). Error patterns in humor comprehension, for both coherent and surprising distractors, discriminate between AD and bvFTD.

3.2.3. Summary of behavioral results

Our goal here is to identify the tasks and error types that discriminate between AD and bvFTD patients, and that may contribute to a clarification of some differences between the cognitive impairment profiles of the two disorders. The two groups of patients showed different response patterns in only two tasks in this study.

In the inference task, bvFTD patients, but not AD patients, tend to evaluate a false conclusion to be true more often than did both AD and controls (e.g., 'The postman delivered the mail', given the premises: 'The postman walked towards the mailbox. For several weeks he had waited in vain for that letter'; Fig. 1e). This pattern of results is only observed in bvFTD patients, while the performance of AD patients is closer to that of healthy controls.

In the humor comprehension task, bvFTD patients, but not AD patients, indicate a coherent distractor ('There is no cure for old age ailments!') as a humorous ending of a joke ('The doctor says to a senior patient: "Dear Madam, the pain in your right arm is not a serious problem. It is simply due to old age." Then the woman says:'), more often than both AD and controls. This pattern is seen only in bvFTD patients, whereas AD patients are closer to controls. AD patients tend to choose a surprising distractor ('The government fell today') as a humorous ending for the joke, more often than both bvFTD patients and controls.

3.3. Neuroimaging results

We first examined where a decrease in FDG uptake could be detected in the brain of bvFTD and AD patients, relative to a group of healthy participants of similar age.

When compared to healthy controls, bvFTD patients showed a severe decrease in FDG uptake, primarily in the frontal and anterior temporal cortices, with a milder involvement of the parietal lobe (Fig. 2). In contrast, AD patients showed a more posterior pattern of cortical thinning, involving principally the temporal, parietal, and occipital lobes bilaterally, with a milder involvement of frontal cortex. We then examined whether and where AD and bvFTD patients showed a different pattern of cortical thinning. Compared to AD, bvFTD patients had a more severe reduction of FDG uptake affecting to a greater extent the inferior frontal gyrus (IFG), medial prefrontal cortex, and the anterior cingulate bilaterally. AD patients, compared to bvFTD, presented more severe reduction of FDG uptake in the posterior temporal lobe, lateral parietal lobe, lateral occipital lobe, cuneus, and precuneus. The results are summarized in Table 2.

Finally, we focused on the relation between the severity of brain damage produced by AD or bvFTD and behavior in the BLED tasks. We tested for anatomo-functional correlations using a regions-of-interest (ROI) approach. To generate the ROIs, we considered the contrast reported in Fig. 2 (top and middle lower panel), i.e., those including the brain regions damaged in bvFTD and AD, compared to controls. First, we performed a correlation analysis in each patient group to assess group-specific effects in the corresponding ROIs. Second, we analyzed correlations by considering the entire patient group to increase the sample size and to generate more robust correlation estimates. Collapsing patient groups is admissible, if the only variable affecting behavior is the location of brain damage as opposed to the pathology that caused the damage: this can be assumed in our case.

We found significant anatomo-functional correlations with 3 behavioral variables: false conclusions (FC) in the discourse inference task, humorous conclusions (HC) and coherent distractors (CD) in the humor comprehension task. These variables correlated with signal strength in the designated bvFTD-ROI when bvFTD patients are considered (inferencing: FC, r=0.42, p=0.019; humor: HC: r=0.39, p=0.037, CD: r=-0.43, p=0.023). The variable coherent distractor correlated with signal strength also when all patients (bvFTD and AD) are considered (r=-0.27, p=0.036). None of the three behavioral variables correlated with PET signal strength in the AD-ROI, either when only AD or when all patients are considered (all p values > 0.1).

4. Discussion

Language comprehension involves more than just processing individual words and combining them into phrases or sentences. The broader discourse context and the intentions of the speaker, or the source of the written message, must be taken into account when constructing interpretations of the input. These discourse semantic and pragmatic processes are crucial for everyday verbal communication and other forms of social interaction. Patients with focal lesions in cortical regions associated with language or communication have been extensively studied in recent decades. Much less is known about the effects of neurodegenerative disorders on discourse semantics and pragmatics. We conducted a study of these skills in Italian-speaking patients with Alzheimer's disease (AD) or with behavioral variant frontotemporal dementia (bvFTD), using PET imaging, neuropsychological testing in key cognitive domains, and a standardized battery (BLED; Rinaldi et al. 2006) including tests for discourse inferences and comprehension of idioms, requests, and humor.

In general, we found that both AD and bvFTD patients were impaired compared to controls. Both patient groups had general difficulties in discourse comprehension, despite the fact that both were largely spared in single word comprehension and grammar (the lack of a more thorough assessment of language capacities in these patients is a limitation of the present study). Importantly, however, the two groups showed also specific patterns of impairment, due to the different distributions of cortical depletion in AD and bvFTD. Their performance was comparably impaired in the idiom comprehension and the request comprehension tasks: neither overall error frequencies nor the frequencies of specific error types could differentiate the two patient groups in these two tasks. By contrast, bvFTD patients showed, when compared to AD, specific difficulties in the *discourse inference* task (i.e., they were more likely to consider false conclusions as true) and in the *humor comprehension* task (they tended to indicate a coherent distractor as the joke's punchline). These specific impairments are associated with more severe damage in bvFTD than AD to the bilateral inferior frontal gyrus (IFG), medial prefrontal cortex, and anterior cingulate. Instead, AD patients made different errors in the *humor comprehension* task (i.e., occasionally they indicated a surprising distractor as a joke's punchline). These more subtle differences between bvFTD and AD patients are also supported by the results of anatomo-functional correlations.

4.1. Specific discourse semantic and pragmatic deficits in AD and bvFTD

Performance differences in the inferencing and in the humor comprehension tasks may be explained with reference to neurobehavioral impairments specific to either the AD or the bvFTD group. bvFTD patients made more errors than AD patients in only two cases: in the discourse inference task, FC (false conclusion) errors; and in the humor comprehension task, CD (coherent distractor) errors. Patients with AD, in these two cases, performed as controls. Therefore, these errors point to deficits that are not shared between bvFTD and AD patients.

To explain these findings, it is useful to highlight what these stimuli and tasks have in common. In the inference task, participants are presented with a brief discourse context (e.g., 'The postman walked towards the mailbox. For several weeks he had waited in vain for that letter') and then assess the truth of three target sentences: a true conclusion (TC: 'The postman checked whether the letter he was waiting for had arrived'), a false conclusion (FC: 'The postman delivered the mail'), and factual information (FI: 'The postman approached the mailbox'). bvFTD patients make more FC errors than both AD patients and controls. Avoiding these errors requires integration of information provided by the context and suppression of inferences that would follow from a subset of the premises plus background knowledge (e.g., the postman delivered the mail). In the humor comprehension task, patients are given a context (e.g., 'The doctor says to a senior patient: "Dear Madam, the pain in your right arm is not a serious problem. It is simply due to old age." Then the woman says:'), and they are asked to indicate the sentence that makes the vignette a joke, choosing between a humorous conclusion (HC: 'This is nonsense, doctor. My left hand has the same age but is healthy.'), a non-humorous coherent distractor (CD: 'There is no cure for old age ailments!'), and a surprising distractor with no logical connection with the rest of the discourse (SD: 'The government fell today'). bvFTD patients make more CD errors than both AD patients and controls. Avoiding these errors requires the ability to integrate of information given by the context (an incomplete joke), appreciation of the aims of the task as instructed (choosing a humorous conclusion), correct attribution of mental states to participants (e.g., the woman's intention to say something humorous), and suppression of the responses that normally, in the absence of special communicative aims, such as humor, might be appropriate in the given situation (e.g., a non-humorous coherent reply would suit the context of a medical examination). On the basis of the present results, it is difficult to identify precisely the processes that are impaired in bvFTD, but not in AD patients, and that could therefore explain the observed differences. However, a few considerations may help rule out some possibilities.

FTD patients here do not appear to have special difficulties with the integration of discourse or contextual information, at least not to a degree that would explain the observed differences with AD patients. For example, the frequency of CD errors in the humor comprehension task, if anything, would be suggestive of an 'excess of coherence' in their representations of discourse-level meanings. In general, correct integration of the elements that constitute a discourse (a) is implicated in all tasks of BLED and (b) is logically required, as a preliminary step, by further task-specific processes: if bvFTD patients had a specific impairment in discourse integration, not shared by AD patients, one would in general expect (i) more frequent errors by bvFTD patients than both AD patients and controls *across BLED subtests*, and (ii) a more random distribution of errors across error types in each subtest. None of this was found in our data. This line of argument suggests that the relative prevalence of FC errors in bvFTD patients could be a consequence of their impaired ability to suppress generally valid inferences that are however disallowed in a given context (Burgess & Shallice 1996; Hornberger et al. 2008; Robinson et al. 2015). This may link poorer performance by these patients in the inferencing task with executive dysfunction, evidence of which was indeed found in our own neuropsychological assessment results. From a neurological stance, it seems most parsimonious, and also consistent with earlier results, to connect this particular deficit to atrophy in regions of the PFC, which present greater damage in bvFTD than in AD patients (Viskontas et al. 2007; Huey et al. 2009; Harciarek & Cosentino 2013; Irish et al. 2012; Spotorno et al. 2015). It should be noted that previous research suggests that bvFTD patients tend to be overly rigid in their interpretations of language. For example, Spotorno et al. (2015) have shown that bvFTD patients have difficulties generating pragmatic readings of expressions that trigger scalar implicatures (e.g., 'some' implicating some, but not all). However, given a choice between logical and pragmatic readings, they prefer the latter, as controls also do. These findings may be reconciled with our results, if one assumes that, in both cases, bvFTD patients have greater difficulty generating the contextually preferred or valid inference: in one case, they fall back on logical readings (e.g., of a scalar term); in the other, they fall back on world knowledge (e.g., of what postmen typically do).

As regards the explanation of CD errors in the humor comprehension task, again the discourse integration deficit account seems ruled out by the argument outlined in the previous paragraph. Likewise, bvFTD patients' impaired ability to suppress contextually inappropriate responses (e.g., a non-humorous distractor), related to executive dysfunction in bvFTD, could explain the relative frequency of CD errors.

This hypothesis may provide an overarching account of deficit differences between bvFTD and AD patients. Yet, one cannot rule out a contribution of a mentalizing deficit in bvFTD patients for explaining more frequent CD errors, given greater damage to mPFC in bvFTD compared to AD in our sample. On that assumption, one account could go as follows. bvFTD patients have difficulty assigning appropriate mental states (e.g., the intention to say something humorous) to the agent in the vignette, and so they pick less often the correct humorous conclusion (HC): greater damage to mPFC in bvFTD patients may explain this specific aspect of the deficit. Moreover, as in the inference task, bvFTD patients may have difficulty suppressing a generally appropriate response, so they tend to pick a coherent distractor (CD) more often than controls and AD patients: damage to areas of PFC associated with executive function and cognitive control (i.e., rostral and lateral PFC) may explain this aspect of the deficit.

The pattern of errors by AD patients in the humor comprehension subtest may be understood as follows. If TPJ supports spontaneous mental state attribution, then AD patients should be able to attribute mental states in this task, in which explicit instructions are given (see Methods). In contrast to bvFTD patients, mPFC in AD patients is relatively spared. Here, the question is how to explain the fact that AD patients make more errors (like bvFTD patients they choose less often the correct HR response); however, the errors they make more frequently are SD errors, and not CD errors (unlike bvFTD patients). One hypothesis is that AD patients, guided by task instructions, are able to assign an appropriate (i.e., a humorous) intention to the speaker in the story, but they cannot then select a response that is effectively humorous. AD patients might, in contrast to bvFTD patients, (i) be able to suppress the inappropriate coherent (CD) response, but (ii) they may be unable to generate or select a suitable humorous response, which conventionally requires more than a surprising contribution to discourse: (i) might be explained by relatively spared PFC in AD, supporting executive function and cognitive control (e.g., suppression of generally appropriate or default responses), and (ii) may be explained by damage to posterior temporal or temporo-parietal areas, encoding conventional mappings of form to meaning types or conditions on those mappings, given a communicative intent or purpose (e.g., what is required to produce a humorous effect).

In brief, the observed differences between bvFTD and AD patients in the discourse inferencing and humor comprehension tasks may be explained within a coherent framework, attributing impaired mentalizing and default response suppression to bvFTD patients (linked to frontal lesions, in bvFTD), and impaired generation or selection of contextually appropriate responses in AD patients (linked to temporal or temporo-parietal lesions). Previous research with bvFTD patients (Spotorno et al. 2015) suggests that these patients may have difficulty generating alternative interpretations of an utterance, whereas they may still be able to draw inferences. Our work indicates instead that bvFTD patients may be impaired in their ability to suppress inferences that are generally valid, but are contextually blocked. As for humor comprehension, our findings are generally in agreement with earlier work (Clark et al. 2015, 2016), showing difficulties in bvFTD patients with processing humor, irony, or sarcasm (Shany-Ur et al., 2012). The comparable response profiles of bfFTD and AD in the idioms task is not entirely consistent with earlier work on other fixed form-meaning constructions (e.g., on proverbs; Kaiser et al., 2013), but these discrepancies may be due to differences between the stimuli or task: idioms and proverbs are different on many levels; e.g., proverbs are full sentences or even short stories, whereas idioms are syntactic phrases, and moreover only proverbs require comprehenders to extract a general truth or moral from the material.

4.2. From common deficit patterns to models of language in the brain

The performance of AD and bvFTD patients is comparable in two BLED subtests: i.e., comprehension of idioms and indirect requests. AD and bvFTD patients show similar response patterns, evidenced by a lack of a between-groups difference in the analysis of correct responses (Figs. 1a, 1g) and of errors (Figs. 1b-c, 1h-i). Yet, these findings would require different explanations, due to the different nature of the stimuli and tasks for idioms and requests.

In the idiom comprehension subtest, in each trial the participant is presented with a sentence in which an idiom is embedded (e.g., 'The father read between the lines a request of help') and is then asked to choose, among three additional sentences, the one giving the "correct interpretation" of the first item. Both AD and bvFTD patients choose less often than controls the idiomatic interpretation (II) (i.e., 'The father understood that his son needed help'; Fig. 1a), and both groups choose more often a literal sentence (LI; 'Between one line and the next the father found written a request of help by his son'; Fig. 1b) or a neutral sentence (NI: 'The father had to ask for help to read his son's letter'; Fig. 1c). This task requires that participants recover the meaning of the idiom (e.g., for 'read between the lines': to understand an implicit suggestion or message), which often cannot be derived compositionally.

Debates in psycholinguistics have focused largely on how idiomatic meanings may be accessed and retrieved—whether indirectly (e.g., via pragmatic inferences) or directly. Most current models of language in the brain are virtually unanimous in assigning to the temporal cortex bilaterally, in particular the posterior middle and superior temporal gyri (pMTG/STG), a critical role in accessing and retrieving the meanings of lexical items: morphemes, words, idioms, conventional metaphors etc. (see Hagoort 2005; Hickok & Poeppel 2007; Lau et al. 2008; Hagoort et al. 2009; Friederici 2017; Baggio 2018). These regions are damaged in both AD and bvFTD patients. A possible explanation of the observed deficits in idiom comprehension, in bvFTD and AD patients, is damage to a bilateral perisylvian network subserving access and retrieval of idiomatic meaning, implicating in particular pMTG/STG.

The shared deficit in the request comprehension task requires a different analysis. Meta-analyses, individual imaging experiments, and reviews point to a distributed medial frontal and temporo-parietal network (including TPJ) subserving inference to the communicative intentions of the speaker and pragmatic reasoning (Ferstl et al. 2008; Bambini 2010; Bohrn et al. 2012; Rapp et al. 2012; Bašnáková et al. 2013; Hagoort & Indefrey 2014; Hagoort & Levinson 2014; Catani & Bambini 2014). Three types of models are relevant in this context:

(A) Models that emphasize the role of a "temporo-parietal network for pragmatic integration", involving AG/TPJ ('Geschwind's region') and superior temporal cortex ('Wernicke's region'), i.e., for computing "speakers' meaning at the highest level of communication" (Catani & Bambini 2014). This network is moreover connected to dorsolateral and ventromedial PFC (vmPFC), and it "supports complex integration and inferential mechanisms well beyond the simple recognition of communicative intentions" (Catani & Bambini 2014).

(B) Models that emphasize functional interactions between medial frontal (mPFC) and temporo-parietal (e.g., TPJ) areas in high-level pragmatic processing, including comprehension of requests (Hagoort & Indefrey 2014; Hagoort & Levinson 2014).

(C) Models where pragmatic processing arises from the interplay of (left) inferior parietal regions, representing salient aspects of the reference structure (or model) in which discourse is interpreted (e.g., numerosities and spatio-temporal locations of entities and events), and medial frontal regions (e.g., the anterior-rostral mPFC; Amodio & Frith 2006; Frith & Frith 2006), encoding the speaker's communicative intentions about relevant entities and events in discourse (Baggio 2018).

Our results challenge these models in different ways. Firstly, A-type accounts, such as the SCALED model by Catani & Bambini (2014), correctly predict that bvFTD and AD patients have difficulty in tasks that require "pragmatic integration"—here, the request comprehension task and the humor comprehension task, at least—due to damage to temporo-parietal cortex. Indeed, this region presents atrophy in both our bvFTD and AD samples. Nonetheless, AD patients present greater damage than bvFTD patients in these regions (Fig. 2): the SCALED model, and related A-type accounts, predict a greater impairment for AD patients across all tasks requiring "pragmatic integration": this was not found in the present data set. Conversely, B- and C-type models attribute to the mPFC a prominent role in high-level pragmatic processes, and predict a greater deficit in bvFTD than AD patients, as the former show greater deterioration of medial frontal cortex than the latter. Again, this was not found in the present study. Further research, with test batteries more sensitive than BLED, are needed to gather additional evidence bearing on these issues.

There are several independent results which challenge models assigning a primary function to the TPJ specifically, that is, both A-type and B-type models. Firstly, TPJ, in contrast to the inferior parietal lobule (IPL), for example, “is not associated with any objective landmarks” anatomically, and can be fractionated into sub-areas with different functions and connectivities (Ingelstrom and Graziano 2017). Second, in the SCALED model, the TPJ is part of a left perisylvian network of areas connected via the arcuate fasciculus (Catani et al. 2005). Thus, SCALED assigns a role to the *left* TPJ in pragmatic processing, although contributions from the right TPJ are not ruled out by it, and are even supported by recent results (Carotenuto et al. 2018). Studies indicate that the right TPJ is a key temporo-parietal region for pragmatics (see Bašnáková et al. 2013; Hagoort & Indefrey 2014; Hagoort & Levinson 2014). Recent results of imaging experiments and studies of brain-damaged patients are mixed. Overall, they do not offer support for A-type or B-type models. For example, lesions to the left posterior TPJ can impair the patient’s spontaneous ability to take into account others’ beliefs in a false-belief task without any explicit mentalizing instructions. However, these patients show no deficits when attention is explicitly directed to what that the other person believes or to what they might do (Biervoye et al. 2016). Pragmatic processing requires precisely *spontaneous* consideration of others’ beliefs. Therefore, this result is consistent with A-type models like SCALED. But this finding also suggests that representation of others’ mental states involves other brain regions in addition to the left TPJ: the mPFC is one candidate. Likewise, the right TPJ has been shown to be non-selective for mentalizing (Mitchell 2007), and to respond to a functionally broader range of stimuli and tasks (e.g., biological motion processing, mentalizing, attention re-orienting), or to common underlying cognitive or neural processes (Lee & McCarthy 2016).

Taken together, these data indicate that a functional specialization for spontaneous pragmatic inference, as required by processing indirect requests, is not a feature of single brain regions, and most likely not a feature of either the left or right TPJ, but emerges instead at the level of fronto-parietal interactions. Theories that stress the functional role of mPFC in pragmatics predict greater impairment in bvFTD, and theories that stress the role of left or right TPJ predict more marked deficits in AD: our findings in the request comprehension task—a paradigmatic test of pragmatic skills—do not support these predictions. We tentatively conclude that (i) medial frontal and parietal or temporo-parietal regions are *jointly necessary* for pragmatic processing and (ii) damage to either of these regions suffices to disrupt pragmatic processes involved in understanding indirect requests. Fully in line with previous work, we account for deficit associations in AD and bvFTD by arguing that lesions to perisylvian networks are sufficient to impair retrieval of idiom meaning. Lesions to fronto-parietal networks suffice to impair comprehension of indirect requests.

5. Conclusion

We combined PET scans with standard neuropsychological assessment tools and a standardized battery in Italian (BLED) to study pragmatic and discourse semantic deficits in bvFTD and AD patients. PET scans revealed greater damage to posterior temporal and parietal cortex bilaterally in AD patients, and to frontal and temporal regions bilaterally in bvFTD patients. We found a similar deficit in bvFTD and AD patients in comprehension of idioms and indirect requests, and we attributed this to damage to fronto-temporal networks subserving the retrieval of form-meaning pairings (idioms) or to fronto-parietal regions subserving the attribution of mental states to others (requests). We also reported differences between patient groups in evaluating inferences from discourse and selecting humorous completions of short stories. Here, bvFTD patients had greater difficulty suppressing contextually inappropriate responses, which we attributed to frontal atrophy, while AD patients had more difficulty generating or selecting a response that is contextually and conventionally apt, also given the pragmatic and communicative intentions of the speaker, as is required by humor. We attributed this deficit to damage to temporal and parietal regions in AD, leading to a specific impairment in the selection and generation of semantically or pragmatically suitable responses.

References

Adenzato, M., Cavallo, M., & Enrici, I. (2010). Theory of mind ability in the behavioural variant of frontotemporal dementia: an analysis of the neural, cognitive, and social levels. Neuropsychologia, 48(1), 2-12.

Amanzio, M., Geminiani, G., Leotta, D., & Cappa, S. (2008). Metaphor comprehension in Alzheimer's disease: Novelty matters. Brain and Language, 107(1), 1-10.

Amodio, D. M., & Frith, C. D. (2006). Meeting of minds: the medial frontal cortex and social cognition. Nature Reviews Neuroscience, 7(4), 268.

Arcara, G., & Bambini, V. (2016). A test for the assessment of pragmatic abilities and cognitive substrates (APACS): Normative data and psychometric properties. Frontiers in psychology, 7, 70.

Baggio, G. (2018). Meaning in the Brain. MIT Press.

Baggio, G., Van Lambalgen, M., & Hagoort, P. (2012). The processing consequences of compositionality. In The Oxford handbook of compositionality (pp. 655-672). Oxford University Press.

Baggio, G., Stenning, K., & van Lambalgen, M. (2016). Semantics and cognition. The Cambridge Handbook of Formal Semantics, 756-774.

Baggio, G., Granello, G., Verriello, L., & Eleopra, R. (2016). Formal Semantics in the Neurology Clinic: Atypical Understanding of Aspectual Coercion in ALS Patients. Frontiers in psychology, 7, 1733.

Bambini, V. (2010). Neuropragmatics: A foreword. Rivista di Linguistica, 22(1).

Bambini, V., Arcara, G., Martinelli, I., Bernini, S., Alvisi, E., Moro, A., Cappa, S., & Ceroni, M. (2016). Communication and pragmatic breakdowns in amyotrophic lateral sclerosis patients. Brain and Language, 153, 1-12.

Bašnáková, J., Weber, K., Petersson, K. M., van Berkum, J., & Hagoort, P. (2013). Beyond the language given: the neural correlates of inferring speaker meaning. Cerebral Cortex, 24(10), 2572-2578.

Biervoye, A., Dricot, L., Ivanoiu, A., & Samson, D. (2016). Impaired spontaneous belief inference following acquired damage to the left posterior temporoparietal junction. Social cognitive and affective neuroscience, 11(10), 1513-1520.

Bohrn, I. C., Altmann, U., & Jacobs, A. M. (2012). Looking at the brains behind figurative language—A quantitative meta-analysis of neuroimaging studies on metaphor, idiom, and irony processing. Neuropsychologia, 50(11), 2669-2683.

Bonner, M. F., Ash, S., & Grossman, M. (2010). The new classification of primary progressive aphasia into semantic, logopenic, or nonfluent/agrammatic variants. Current neurology and neuroscience reports, 10(6), 484-490.

Brownell, H. H., Potter, H. H., Bihrle, A. M., & Gardner, H. (1986). Inference deficits in right brain-damaged patients. Brain and language, 27(2), 310-321.

Burgess, P. W., & Shallice, T. (1996). Response suppression, initiation and strategy use following frontal lobe lesions. Neuropsychologia, 34(4), 263–272.

Caltagirone, C., Gainotti, G., Masullo, C., & Miceli, G. (1979). Validity of some neuropsychological tests in the assessment of mental deterioration. Acta Psychiatrica Scandinavica, 60(1), 50-56.

Campbell, D. W., Wallace, M. G., Modirrousta, M., Polimeni, J. O., McKeen, N. A., & Reiss, J. P. (2015). The neural basis of humour comprehension and humour appreciation: The roles of the temporoparietal junction and superior frontal gyrus. Neuropsychologia, 79, 10-20.

Carotenuto, A., Cocozza, S., Quarantelli, M., Arcara, G., Lanzillo, R., Morra, V.B., Cerillo, I., Tedeschi, E., Orefice, G., Bambini, V. and Brunetti, A., (2018). Pragmatic abilities in multiple sclerosis: The contribution of the temporo-parietal junction. Brain and Language, 185, 47-53.

Catani, M., & Jones, D. K. (2005). Perisylvian language networks of the human brain. Annals of neurology, 57(1), 8-16.

Catani, M., & Bambini, V. (2014). A model for social communication and language evolution and development (SCALED). Current Opinion in Neurobiology, 28, 165-171.

Chapman, S. B., Highley, A. P., & Thompson, J. L. (1998). Discourse in fluent aphasia and Alzheimer's disease: Linguistic and pragmatic considerations. Journal of Neurolinguistics, 11(1-2), 55-78.

Chen, L., Ralph, M. A. L., & Rogers, T. T. (2017). A unified model of human semantic knowledge and its disorders. Nature human behaviour, 1(3), 0039.

Clark, C. N., Nicholas, J. M., Henley, S. M., Downey, L. E., Woollacott, I. O., Golden, H. L., Fletcher, P. D., Mummery, C. J., Schott, J. M., Rohrer, J. D., Crutch, S. J., & Warren, J. D. (2015). Humour processing in frontotemporal lobar degeneration: a behavioural and neuroanatomical analysis. Cortex, 69, 47-59.

Clark, C.N., Nicholas, J.M., Gordon, E., Golden, H.L., Cohen, M.H., Woodward, F.J., Macpherson, K., Slattery, C.F., Mummery, C.J., Schott, J.M. & Rohrer, J.D. (2016). Altered sense of humor in dementia. Journal of Alzheimer's Disease, 49(1), 111-119.

Culicover, P. W., & Jackendoff, R. (2006). The simpler syntax hypothesis. Trends in cognitive sciences, 10(9), 413-418.

Davies, R. R., Hodges, J. R., Kril, J. J., Patterson, K., Halliday, G. M., & Xuereb, J. H. (2005). The pathological basis of semantic dementia. Brain, 128(9), 1984-1995.

De Renzi, E., Motti, F., & Nichelli, P. (1980). Imitating gestures: a quantitative approach to ideomotor apraxia. Archives of neurology, 37(1), 6-10.

De Renzi, E., & Lucchelli, F. (1988). Ideational apraxia. Brain, 111(5), 1173-1185.

Dubois, B., Feldman, H.H., Jacova, C., DeKosky, S.T., Barberger-Gateau, P., Cummings, J., Delacourte, A., Galasko, D., Gauthier, S., Jicha, G. & Meguro, K., (2007). Research criteria for the diagnosis of Alzheimer's disease: Revising the NINCDS–ADRDA criteria. The Lancet Neurology, 6(8), pp.734-746.

Dubois, B., Feldman, H. H., Jacova, C., Cummings, J. L., DeKosky, S. T., Barberger-Gateau, P. et al. (2010). Revising the definition of Alzheimer's disease: A new lexicon. The Lancet Neurology, 9(11), 1118-1127.

Dubois, B., Feldman, H. H., Jacova, C., Hampel, H., Molinuevo, J. L., Blennow, K. et al. (2014). Advancing research diagnostic criteria for Alzheimer's disease: The IWG-2 criteria. The Lancet Neurology, 13(6), 614-629.

Faust, M. (Ed.). (2012). *The Handbook of the Neuropsychology of Language* (Vols. 1-2). John Wiley & Sons.

Ferstl, E. C., Neumann, J., Bogler, C., & Von Cramon, D. Y. (2008). The extended language network: a meta-analysis of neuroimaging studies on text comprehension. Human brain mapping, 29(5), 581-593.

Foldi, N. S. (1987). Appreciation of pragmatic interpretations of indirect commands: Comparison of right and left hemisphere brain-damaged patients. Brain and Language, 31(1), 88-108.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). “Mini-mental state”: a practical method for grading the cognitive state of patients for the clinician. Journal of psychiatric research, 12(3), 189-198.

Friederici, A. D. (2017). Language in our brain: The origins of a uniquely human capacity. MIT Press.

Frith, C. D., & Frith, U. (2006). The neural basis of mentalizing. Neuron, 50(4), 531-534.

Giora, R. (1997). Understanding figurative and literal language: The graded salience hypothesis. Cognitive Linguistics, 8(3), 183-206.

Grossman, M. (2018). Linguistic aspects of primary progressive aphasia. Annual Review of Linguistics, 4, 377-403.

Guendouzi, J., & Savage, M. (2017). Alzheimer's dementia. In Research in clinical pragmatics (pp. 323-346). Springer, Cham.

Hagoort, P. (2005). On Broca, brain, and binding: a new framework. Trends in cognitive sciences, 9(9), 416-423.

Hagoort, P., & van Berkum, J. (2007). Beyond the sentence given. Philosophical Transactions of the Royal Society B: Biological Sciences, 362(1481), 801-811.

Hagoort, P., Baggio, G., & Willems, R. M. (2009). Semantic unification. In The cognitive neurosciences, 4th ed. (pp. 819-836). MIT press.

Hagoort, P., & Indefrey, P. (2014). The neurobiology of language beyond single words. Annual Review of Neuroscience, 37, 347-362.

Hagoort, P., & Levinson, S. C. (2014). Neuropragmatics. In The cognitive neurosciences (pp. 667-674). MIT Press.

Harciarek, M., & Cosentino, S. (2013). Language, executive function and social cognition in the diagnosis of frontotemporal dementia syndromes. International Review of Psychiatry, 25(2), 178-196.

Healey, M. L., McMillan, C. T., Golob, S., Spotorno, N., Rascovsky, K., Irwin, D. J., Clark, R., & Grossman, M. (2015). Getting on the same page: the neural basis for social coordination deficits in behavioral variant frontotemporal degeneration. Neuropsychologia, 69, 56-66.

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. Nature Reviews Neuroscience, 8(5), 393.

Hornberger, M., Piguet, O., Kipps, C., & Hodges, J. R. (2008). Executive function in progressive and nonprogressive behavioral variant frontotemporal dementia. Neurology, 71(19), 1481–1488. doi:71/19/1481 [pii] 10.1212/01.wnl.0000334299.72023.c8

Howard, D., & Patterson, K. (1992). The pyramid and palm trees test: a test of semantic access from words and pictures. Thames Valley Test Company.

Huey, E. D., Goveia, E. N., Paviol, S., Pardini, M., Krueger, F., Zamboni, G., Tierney, M. C, Wassermann, E. M., & Grafman, J. (2009). Executive dysfunction in frontotemporal dementia and corticobasal syndrome. Neurology, 72(5), 453-459.

Irish, M., Piguet, O., & Hodges, J. R. (2012). Self-projection and the default network in frontotemporal dementia. Nature Reviews Neurology, 8(3), 152.

Jung-Beeman, M. (2005). Bilateral brain processes for comprehending natural language. Trends in cognitive sciences, 9(11), 512-518.

Kaiser, N. C., Lee, G. J., Lu, P. H., Mather, M. J., Shapira, J., Jimenez, E., Thompson, P. M., & Mendez, M. F. (2013). What dementia reveals about proverb interpretation and its neuroanatomical correlates. Neuropsychologia, 51(9), 1726-1733.

Kempler, D., Van, D. L., & Read, S. (1988). Proverb and idiom comprehension in Alzheimer disease. Alzheimer disease and associated disorders, 2(1), 38-49.

Landin-Romero, R., Kumfor, F., Leyton, C. E., Irish, M., Hodges, J. R., & Piguet, O. (2017). Disease-specific patterns of cortical and subcortical degeneration in a longitudinal study of Alzheimer's disease and behavioural-variant frontotemporal dementia. Neuroimage, 151, 72-80.

Lau, E. F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics:(de) constructing the N400. Nature Reviews Neuroscience, 9(12), 920.

Lee, S. M., & McCarthy, G. (2016). Functional heterogeneity and convergence in the right temporoparietal junction. Cerebral Cortex, 26(3), 1108-1116.

Lezak, L.M., Howieson, D.B., Bigler, E.D., Tranel, D. (2012). Neuropsychological assessment. Oxford University Press, 5th Edition.

Luzzi, S., Piccirilli, M., Pesallaccia, M., Fabi, K., & Provinciali, L. (2010). Dissociation apraxia secondary to right premotor stroke. Neuropsychologia, 48(1), 68-76.

Luzzi, S., Pesallaccia, M., Fabi, K., Muti, M., Viticchi, G., Provinciali, L., & Piccirilli, M. (2011). Non-verbal memory measured by Rey–Osterrieth Complex Figure B: normative data. Neurological Sciences, 32(6), 1081-1089.

Mashal, N., & M. Faust (2008). Right hemisphere sensitivity to novel metaphoric relations: Application of the signal detection theory. Brain and Language 104: 103–112.

Mashal, N., M. Faust, & T. Hendler (2005). The role of the right hemisphere in processing nonsalient metaphorical meanings: Application of Principal Components Analysis to fMRI data. Neuropsychologia 43: 2084–2100.

Mashal, N., M. Faust, T. Hendler, & M. Jung-Beeman (2007). An fMRI investigation of the neural correlates underlying the processing of novel metaphoric expressions. Brain and Language 100: 115–126.

McKhann, G.M., Knopman, D.S., Chertkow, H., Hyman, B.T., Jack Jr, C.R., Kawas, C.H., Klunk, W.E., Koroshetz, W.J., Manly, J.J., Mayeux, R. & Mohs, R.C. (2011). The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimer's & dementia, 7(3), 263-269.

McMillan, C. T., Rascovsky, K., Khella, M. C., Clark, R., & Grossman, M. (2011). The neural basis for establishing a focal point in pure coordination games. Social cognitive and affective neuroscience, 7(8), 881-887.

McMillan, C. T., Coleman, D., Clark, R., Liang, T. W., Gross, R. G., & Grossman, M. (2013). Converging evidence for the processing costs associated with ambiguous quantifier comprehension. Frontiers in psychology, 4, 153

Mesulam, M.M., Rogalski, E.J., Wieneke, C., Hurley, R.S., Geula, C., Bigio, E.H., Thompson, C.K. and Weintraub, S., (2014). Primary progressive aphasia and the evolving neurology of the language network. Nature Reviews Neurology, 10(10), 554-569.

Miceli, G., Laudanna, A., Burani, C., & Capasso, R. (1994) Batteria per l'Analisi dei Deficit Afasici. BADA (Battery for Analysis of Aphasic Deficits). Università Cattolica del Sacro Cuore, CEPSAG.

Mitchell, J. P. (2007). Activity in right temporo-parietal junction is not selective for theory-of-mind. Cerebral cortex, 18(2), 262-271.

Monti, M. M., Parsons, L. M., & Osherson, D. N. (2009). The boundaries of language and thought in deductive inference. Proceedings of the National Academy of Sciences, 106(30), 12554-12559.

Morrison, R. G., Krawczyk, D. C., Holyoak, K. J., Hummel, J. E., Chow, T. W., Miller, B. L., & Knowlton, B. J. (2004). A neurocomputational model of analogical reasoning and its breakdown in frontotemporal lobar degeneration. Journal of cognitive neuroscience, 16(2), 260-271.

Neary, D., Snowden, J.S., Gustafson, L., Passant, U., Stuss, D., Black, S.A.S.A., Freedman, M., Kertesz, A., Robert, P.H., Albert, M. & Boone, K., (1998) Frontotemporal lobar degeneration A consensus on clinical diagnostic criteria. Neurology, 51(6), pp.1546-1554.

Papagno, C. (2001). Comprehension of metaphors and idioms in patients with Alzheimer's disease: A longitudinal study. Brain, 124(7), 1450-1460.

Papagno, C., Lucchelli, F., Muggia, S., & Rizzo, S. (2003). Idiom comprehension in Alzheimer's disease: The role of the central executive. Brain, 126(11), 2419-2430.

Patterson, K., Nestor, P. J., & Rogers, T. T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. Nature Reviews Neuroscience, 8(12), 976.

Pievani, M., Filippini, N., Van Den Heuvel, M. P., Cappa, S. F., & Frisoni, G. B. (2014). Brain connectivity in neurodegenerative diseases—from phenotype to proteinopathy. Nature Reviews Neurology, 10(11), 620.

Prado, J., Van Der Henst, J. B., & Noveck, I. A. (2010). Recomposing a fragmented literature: How conditional and relational arguments engage different neural systems for deductive reasoning. Neuroimage, 51(3), 1213-1221.

Prado, J., Chadha, A., & Booth, J. R. (2011). The brain network for deductive reasoning: a quantitative meta-analysis of 28 neuroimaging studies. Journal of cognitive neuroscience, 23(11), 3483-3497.

Pylkkänen, L., & McElree, B. (2006). The syntax-semantics interface: On-line composition of sentence meaning. In Handbook of psycholinguistics (pp. 539-579). Academic Press.

Rahman, S., Sahakian, B. J., Hodges, J. R., Rogers, R. D., & Robbins, T. W. (1999). Specific cognitive deficits in mild frontal variant frontotemporal dementia. Brain, 122(8), 1469-1493.

Rapp, A. M., & Wild, B. (2011). Nonliteral language in Alzheimer dementia: a review. Journal of the International Neuropsychological Society, 17(2), 207-218.

Rapp, A. M., Mutschler, D. E., & Erb, M. (2012). Where in the brain is nonliteral language? A coordinate-based meta-analysis of functional magnetic resonance imaging studies. Neuroimage, 63(1), 600-610.

Rascovsky, K., Hodges, J. R., Knopman, D., Mendez, M. F., Kramer, J. H., Neuhaus, J. et al. (2011). Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia. *Brain*, 134(9), 2456-2477.

Rassiga, C., Lucchelli, F., Crippa, F., & Papagno, C. (2009). Ambiguous idiom comprehension in Alzheimer's disease. Journal of clinical and experimental neuropsychology, 31(4), 402-411.

Reisberg, B., Jamil, I. A., Khan, S., Monteiro, I., Torossian, C., Ferris, S., & Wegiel, J. (2011). Principles and practice of geriatric psychiatry.

Reverberi, C., Cherubini, P., Rapisarda, A., Rigamonti, E., Caltagirone, C., Frackowiak, R. S., Macaluso, E. & Paulesu, E. (2007). Neural basis of generation of conclusions in elementary deduction. Neuroimage, 38(4), 752-762.

Reverberi, C., Shallice, T., D'agostini, S., Skrap, M., & Bonatti, L. L. (2009). Cortical bases of elementary deductive reasoning: Inference, memory, and metadeduction. Neuropsychologia, 47(4), 1107-1116.

Reverberi, C., Cherubini, P., Frackowiak, R. S., Caltagirone, C., Paulesu, E., & Macaluso, E. (2010). Conditional and syllogistic deductive tasks dissociate functionally during premise integration. Human brain mapping, 31(9), 1430-1445.

Reverberi, C., Bonatti, L. L., Frackowiak, R. S., Paulesu, E., Cherubini, P., & Macaluso, E. (2012). Large scale brain activations predict reasoning profiles. Neuroimage, 59(2), 1752-1764.

Rinaldi, M. C., Marangolo, P., & Baldassarri, F. (2004). Metaphor comprehension in right brain-damaged patients with visuo-verbal and verbal material: A dissociation (re)considered. Cortex, 40(3), 479-490.

Rinaldi M. C., Marangolo P., Lauriola M. BLED Santa Lucia – Batteria sul Linguaggio dll'Emisfero Destro SantaLucia. Giunti O. S. 2006.

Roberts, A., Savundranayagam, M., & Orange, J. B. (2017). Non-Alzheimer Dementias. In Research in Clinical Pragmatics (pp. 347-377). Springer.

Robinson, G. A., Cipolotti, L., Walker, D. G., Biggs, V., Bozzali, M., & Shallice, T. (2015). Verbal suppression and strategy use: a role for the right lateral prefrontal cortex? Brain, awv003. doi:10.1093/brain/awv003

Sacco, K., Angeleri, R., Bosco, F. M., Colle, L., Mate, D., & Bara, B. G. (2008). Assessment Battery for Communication—ABaCo: A new instrument for the evaluation of pragmatic abilities. Journal of Cognitive Science, 9, 111–157.

Seeley, W. W. (2008). Selective functional, regional, and neuronal vulnerability in frontotemporal dementia. Current opinion in neurology, 21(6), 701.

Seeley, W. W., Crawford, R. K., Zhou, J., Miller, B. L., & Greicius, M. D. (2009). Neurodegenerative diseases target large-scale human brain networks. Neuron, 62(1), 42-52.

Shallice, T. (1988). From neuropsychology to mental structure. Cambridge University Press.

Shallice, T., & Cooper, R. (2011). The organisation of mind. Oxford University Press.

Shany-Ur, T., Poorzand, P., Grossman, S.N., Growdon, M.E., Jang, J.Y., Ketelle, R.S., Miller, B.L. & Rankin, K.P., (2012). Comprehension of insincere communication in neurodegenerative disease: lies, sarcasm, and theory of mind. *Cortex* 48(10), 1329-1341.

Snowden, J. S., Gibbons, Z. C., Blackshaw, A., Doubleday, E., Thompson, J., Craufurd, D., Foster, J., Happé, F., & Neary, D. (2003). Social cognition in frontotemporal dementia and Huntington's disease. Neuropsychologia, 41(6), 688-701.

Schönbrodt, F. D., & Perugini, M. (2013). At what sample size do correlations stabilize? Journal of Research in Personality, 47(5), 609–612. doi:10.1016/j.jrp.2013.05.009

Snowden, J. S., Thompson, J. C., & Neary, D. (2004). Knowledge of famous faces and names in semantic dementia. Brain, 127(4), 860-872.

Spinnler, H. & Tognoni, G. (1987), Standardizzazione e taratura italiana di test neuropsicologici. Italian Journal of Neurological Sciences, 8.

Spotorno, N., McMillan, C. T., Rascovsky, K., Irwin, D. J., Clark, R., & Grossman, M. (2015). Beyond words: Pragmatic inference in behavioral variant of frontotemporal degeneration. Neuropsychologia, 75, 556-564.

Stemmer, B. (2017). Neural aspects of pragmatic disorders. In: Cummings L. (eds) *Research in Clinical Pragmatics. Perspectives in Pragmatics, Philosophy & Psychology*, Vol. 11. Springer.

Tombaugh, T. N., & McIntyre, N. J. (1992). The mini-mental state examination: A comprehensive review. Journal of the American Geriatrics Society, 40(9), 922-935.

Vartanian, O., Goel, V., Tierney, M., Huey, E. D., & Grafman, J. (2009). Frontotemporal dementia selectively impairs transitive reasoning about familiar spatial environments. Neuropsychology, 23(5), 619.

Viskontas, I. V., Possin, K. L., & Miller, B. L. (2007). Symptoms of frontotemporal dementia provide insights into orbitofrontal cortex function and social behavior. Annals of the New York Academy of Sciences, 1121(1), 528-545.

Vrticka, P., Black, J. M., & Reiss, A. L. (2013). The neural basis of humour processing. Nature Reviews Neuroscience, 14(12), 860.

Waltz, J. A., Knowlton, B. J., Holyoak, K. J., Boone, K. B., Mishkin, F. S., de Menezes Santos, M., Thomas, C.R., & Miller, B. L. (1999). A system for relational reasoning in human prefrontal cortex. Psychological science, 10(2), 119-125.

Waltz, J. A., Knowlton, B. J., Holyoak, K. J., Boone, K. B., Back-Madruga, C., McPherson, S., Masterman, D., Chow, T., Cummings, J. L., & Miller, B. L. (2004). Relational integration and executive function in Alzheimer's disease. Neuropsychology, 18(2), 296.

Warrington E.K., James M. VOSP: The visual object and space perception battery. Suffolk, UK: Thames Valley Company, 1991.

Tables

Table 1. Summary of results of the background neuropsychological assessment. Mean (SD) scores in each test are shown, along with the results of Mann-Whitney U tests for the comparisons between patient groups (AD *vs* bvFTD).

Test (maximum score)	FTD	AD	Mann-Whitney U (*p*)
General abilities			
MMSE	23.25 (2.63)	21.76 (6.32)	207 (n.s.)
Memory			
Digit span	4.83 (0.91)	3.96 (0.93)	157 (0.003)
Corsi blocks	4.25 (0.73)	3.76 (0.87)	200 (0.031)
Rey AVLT test long term (15)	4.68 (2.51)	1.52 (2.10)	104 (<0.001)
Rey figure B: delayed recall (31)	12.82 (7.12)	4.68 (5)	107 (<0.001)
Executive function			
Letter fluency (F, A, S)	13.61 (12.33)	16.28 (10.71)	214 (n.s.)
Luria's motor sequences (50)	36.61 (14.14)	36.15 (12.85)	226 (n.s.)
Stroop test (errors)	4.7 (7.43)	17.16 (3.49)	100 (0.001)
Brixton test (55)	24.83 (8.23)	34.1(5.1)	98.5 (<0.001)
Raven's progressive matrices	20.34 (6.17)	24.4 (5.00)	190 (0.02)
Perceptuo-spatial skills			
Overlapping figures	7.58 (1.01)	7.64 (1.22)	277.5 (n.s.)
Screening test from the VOSP battery	19.41(0.71)	19.47 (0.69)	218.5 (n.s.)
Ideomotor praxis			
Right upper limb (20)	19.87 (0.44)	19.72 (0.68)	276 (n.s.)
Left upper limb (20)	19.83 (0.48)	19.64(0.7)	264 (n.s.)
Constructional praxis			
Rey-Osterrieth Figure B copy (31)	27.73 (3.53)	26.61 (4.64)	184 (n.s.)
Language			
Picture naming (40)	34 (5.09)	32.68 (3.32)	215.5 (n.s.)
Word reading (40)	39.58 (1.24)	39.64 (0.57)	268.5 (n.s.)
Word-picture matching (40)	39.91 (0.28)	39.8 (0.41)	265 (n.s.)
Verbal fluency (3 categories 1 min. each)	25.04 (8.03)	25.85 (7.74)	192 (n.s.)

Table 2 Statistical results from comparisons between FDG-uptake in frontotemporal dementia (bvFTD), Alzheimer's disease (AD), and healthy control groups. Coordinates [x, y, z] are reported in the space of the Montreal Neurological Institute (MNI) template.

	Cluster size	MNI coordinates	*t* score
AD < CTL	21664	-3 -55 28	10.35
		-51 -61 16	10.27
		0 -37 34	9.77
	1683	30 -1 70	7.82
		45 11 58	6.67
		24 41 49	6.38
	1606	-39 11 58	7.09
		-21 41 46	6.08
		-27 2 67	6.03
	509	-18 -46 -65	6.37
		39 -52 -65	5.65
		15 -43 -65	5.55
FTD < CTL	27023	-42 8 55	8.86
		45 14 52	8.74
		60 23 19	8.74
AD < CTL	5393	-33 -64 22	6.56
		-18 -46 76	6.01
		39 -61 22	5.53
	923	18 -82 -50	4.97
		6 -70 -50	4.81
		-24 -52 -62	4.74
FTD < AD	3450	60 20 16	5.97
		18 74 4	5.51
		9 20 34	5.15
	221	9 -25 -2	5.51

1100	-51 11 13	5.4
	-30 65 -11	5.13
	-48 29 10	4.73

Figure legends

Figure 1. Behavioral results from the BLED tasks. Violin plots show the distribution and the probability density of data from different subtests, in different groups (i.e., control, AD, bvFTD). In each plot, the white dot represents the median, the thick black bar represents the interquartile range, thin lines are 95% confidence intervals, and shaded areas indicate kernel density estimates. Significant effects are marked with asterisks: for each group */X, where X is the group of comparison in Wilcoxon tests, and CTR is the control group.

Figure 2. Comparisons between [^{18}F] 2-deoxy-2-fluoro-D-glucose (FDG) tracer uptake in frontotemporal dementia (bvFTD), Alzheimer's disease (AD), and healthy control groups. In the top two rows, we show the brain regions where AD and bvFTD patients respectively showed a significant decrease in FDG-uptake. In the bottom row, we present a comparison of the [^{18}F] 2-deoxy-2-fluoro-D-glucose (FDG) uptake in AD and bvFTD. We show in green the regions where AD patients showed a significant reduction of FDG uptake as compared to bvFTD (masked for the presence of a significant decrease of FDG uptake in the baseline contrast between AD and controls) and in yellow the areas where bvFTD patients showed a significant reduction compared to AD (masked for bvFTD<controls).

Figure 3. Relation between the average [^{18}F] 2-deoxy-2-fluoro-D-glucose (FDG) uptake in hypometabolic brain regions in bvFTD and three behavioral indices: false conclusions in the discourse inferencing task, humorous conclusions and coherent distractors in humor processing. These behavioral indices are those displaying a significant anatomo-functional correlation (see 2.5.2). In scatter plots, each point represents a bvFTD or AD patient, and

lines represent the linear trend for each patient group. On the top and right of each scatter plot, we report boxplots showing the distribution of the variables in the patient groups.

Tables

Table 1. Summary of results of the background neuropsychological assessment. Mean (SD) scores in each test are shown, along with the results of Mann-Whitney U tests for the comparisons between patient groups (AD *vs* bvFTD).

Test (maximum score)	FTD	AD	Mann-Whitney U (*p*)
General abilities			
MMSE	23.25 (2.63)	21.76 (6.32)	207 (n.s.)
Memory			
Digit span	4.83 (0.91)	3.96 (0.93)	157 (0.003)
Corsi blocks	4.25 (0.73)	3.76 (0.87)	200 (0.031)
Rey AVLT test long term (15)	4.68 (2.51)	1.52 (2.10)	104 (<0.001)
Rey figure B: delayed recall (31)	12.82 (7.12)	4.68 (5)	107 (<0.001)
Executive function			
Letter fluency (F, A, S)	13.61 (12.33)	16.28 (10.71)	214 (n.s.)
Luria's motor sequences (50)	36.61 (14.14)	36.15 (12.85)	226 (n.s.)
Stroop test (errors)	4.7 (7.43)	17.16 (3.49)	100 (0.001)
Brixton test (55)	24.83 (8.23)	34.1(5.1)	98.5 (<0.001)
Raven's progressive matrices	20.34 (6.17)	24.4 (5.00)	190 (0.02)
Perceptuo-spatial skills			
Overlapping figures	7.58 (1.01)	7.64 (1.22)	277.5 (n.s.)
Screening test from the VOSP battery	19.41(0.71)	19.47 (0.69)	218.5 (n.s.)
Ideomotor praxis			
Right upper limb (20)	19.87 (0.44)	19.72 (0.68)	276 (n.s.)
Left upper limb (20)	19.83 (0.48)	19.64(0.7)	264 (n.s.)
Constructional praxis			
Rey-Osterrieth Figure B copy (31)	27.73 (3.53)	26.61 (4.64)	184 (n.s.)
Language			
Picture naming (40)	34 (5.09)	32.68 (3.32)	215.5 (n.s.)
Word reading (40)	39.58 (1.24)	39.64 (0.57)	268.5 (n.s.)
Word-picture matching (40)	39.91 (0.28)	39.8 (0.41)	265 (n.s.)
Verbal fluency (3 categories 1 min. each)	25.04 (8.03)	25.85 (7.74)	192 (n.s.)

Table 2 Statistical results from comparisons between FDG-uptake in frontotemporal dementia (bvFTD), Alzheimer's disease (AD), and healthy control groups. Coordinates [x, y, z] are reported in the space of the Montreal Neurological Institute (MNI) template.

	Cluster size	MNI coordinates	*t* score
AD < CTL	21664	-3 -55 28	10.35
		-51 -61 16	10.27
		0 -37 34	9.77
	1683	30 -1 70	7.82
		45 11 58	6.67
		24 41 49	6.38
	1606	-39 11 58	7.09
		-21 41 46	6.08
		-27 2 67	6.03
	509	-18 -46 -65	6.37
		39 -52 -65	5.65
		15 -43 -65	5.55
FTD < CTL	27023	-42 8 55	8.86
		45 14 52	8.74
		60 23 19	8.74
AD < CTL	5393	-33 -64 22	6.56
		-18 -46 76	6.01
		39 -61 22	5.53
	923	18 -82 -50	4.97
		6 -70 -50	4.81
		-24 -52 -62	4.74
FTD < AD	3450	60 20 16	5.97
		18 74 4	5.51
		9 20 34	5.15
	221	9 -25 -2	5.51

1100	-51 11 13	5.4
	-30 65 -11	5.13
	-48 29 10	4.73

Figure legends

Figure 1. Behavioral results from the BLED tasks. Violin plots show the distribution and the probability density of data from different subtests, in different groups (i.e., control, AD, bvFTD). In each plot, the white dot represents the median, the thick black bar represents the interquartile range, thin lines are 95% confidence intervals, and shaded areas indicate kernel density estimates. Significant effects are marked with asterisks: for each group */X, where X is the group of comparison in Wilcoxon tests, and CTR is the control group.

Figure 2. Comparisons between [^{18}F] 2-deoxy-2-fluoro-D-glucose (FDG) tracer uptake in frontotemporal dementia (bvFTD), Alzheimer's disease (AD), and healthy control groups. In the top two rows, we show the brain regions where AD and bvFTD patients respectively showed a significant decrease in FDG-uptake. In the bottom row, we present a comparison of the [^{18}F] 2-deoxy-2-fluoro-D-glucose (FDG) uptake in AD and bvFTD. We show in green the regions where AD patients showed a significant reduction of FDG uptake as compared to bvFTD (masked for the presence of a significant decrease of FDG uptake in the baseline contrast between AD and controls) and in yellow the areas where bvFTD patients showed a significant reduction compared to AD (masked for bvFTD<controls).

Figure 3. Relation between the average [^{18}F] 2-deoxy-2-fluoro-D-glucose (FDG) uptake in hypometabolic brain regions in bvFTD and three behavioral indices: false conclusions in the discourse inferencing task, humorous conclusions and coherent distractors in humor processing. These behavioral indices are those displaying a significant anatomo-functional correlation (see 2.5.2). In scatter plots, each point represents a bvFTD or AD patient, and

lines represent the linear trend for each patient group. On the top and right of each scatter plot, we report boxplots showing the distribution of the variables in the patient groups.

Credit author statement

Simona Luzzi : designed the research, selected the patients enrolled in the research, analyzed the neuropsychological data and wrote the paper

Sara Baldinelli: performed the reasearch: she applied the experimental neuropsychological battery

Valentina Ranaldi: performed the reasearch: she applied the experimental neuropsychological battery

Chiara Fiori: performed the reasearch: she applied the background neuropsychological battery

Andrea Plutino: performed the reasearch: he applied background neuropsychologial battery

Fabio Massimo Fringuelli: performed the reasearch: he performed neuroimaging (cerebral PET scan)

Mauro Silvestrini: performed the research: he selected the patients enrolled in the reasearch and contributed to revise the literature

Giosuè Baggio: revised the literarure, analyzed the data and wrote the paper

Carlo Reverberi: analyzed the neurimaging data and wrote the paper

Figure 1

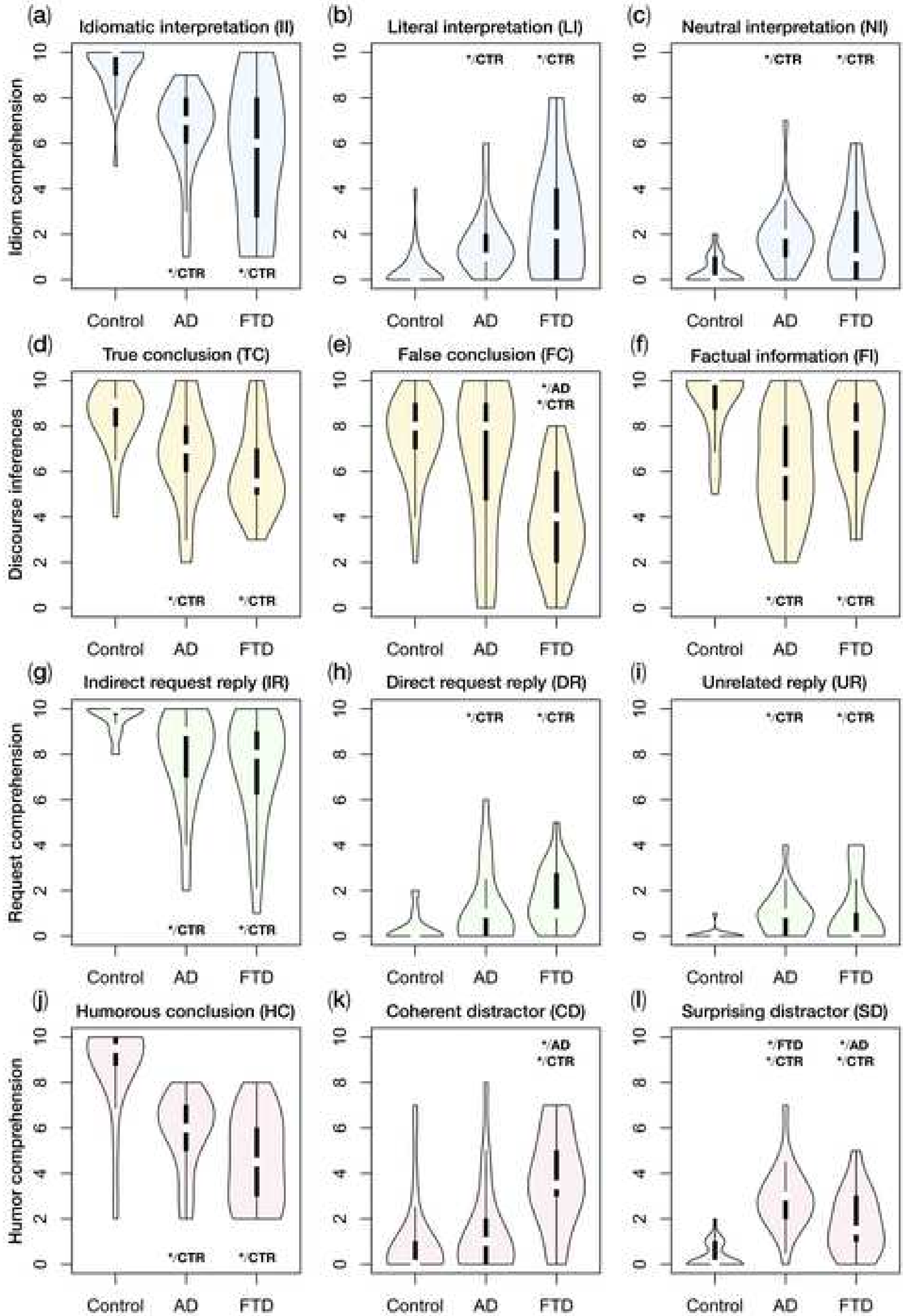

(a) Idiomatic interpretation (II)
(b) Literal interpretation (LI)
(c) Neutral interpretation (NI)
Idiom comprehension
(d) True conclusion (TC)
(e) False conclusion (FC)
(f) Factual information (FI)
Discourse inferences
(g) Indirect request reply (IR)
(h) Direct request reply (DR)
(i) Unrelated reply (UR)
Request comprehension
(j) Humorous conclusion (HC)
(k) Coherent distractor (CD)
(l) Surprising distractor (SD)
Humor comprehension
Control
AD
FTD
*/CTR
*/AD
*/FTD

Figure 2

Click here to access/download;Figure;FIGURE2-FINAL.jpg

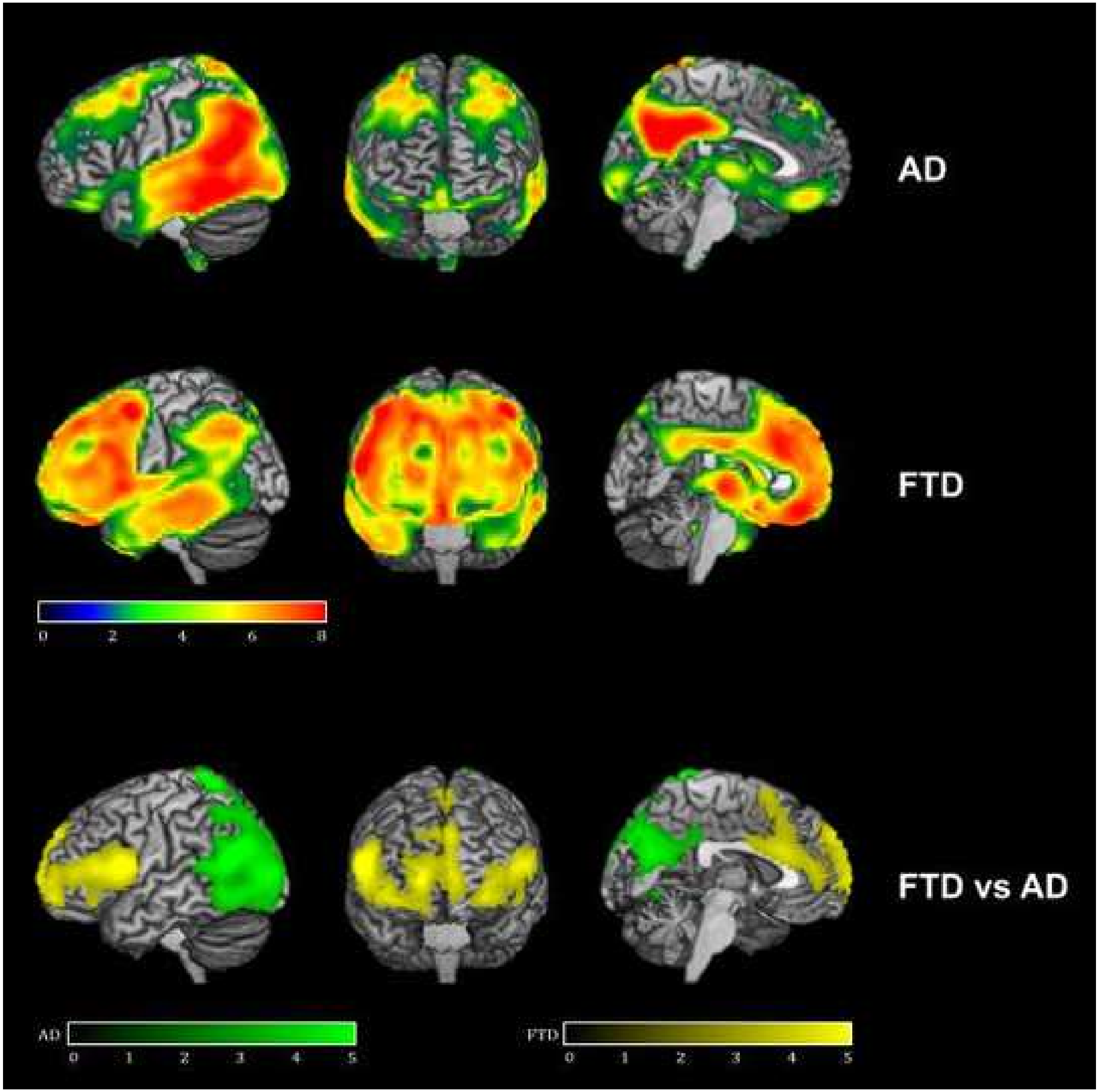
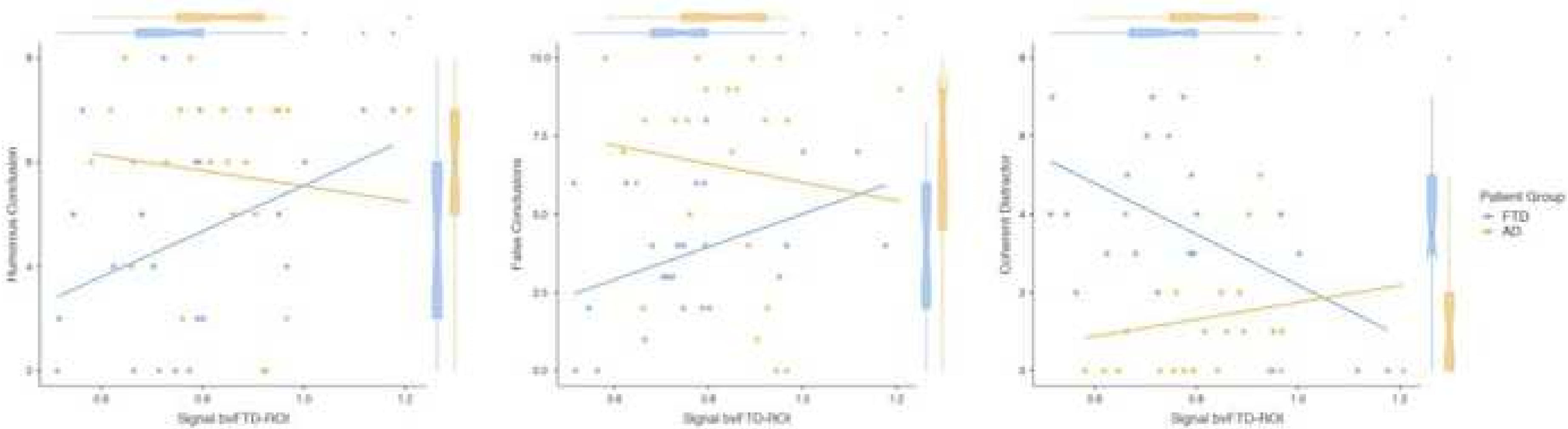


Figure 3


Humorous Conclusion
Signal bvFTD-ROI
False Conclusions
Signal bvFTD-ROI
Coherent Distractor
Signal bvFTD-ROI
Patient Group
FTD
AD