Verbal fluency discrepancies as a marker of the prehippocampal stages of Alzheimer's disease

Laura M. Wright[1], Matteo De Marco[2] and Annalena Venneri[2]

[1]Translational and Clinical Research Institute, Newcastle University, Newcastle upon Tyne, UK

[2]Department of Life Sciences, Brunel University London, London, UK

Author Note

Laura Wright https://orcid.org/0000-0002-0558-7811

We have no conflicts of interest to disclose. We gratefully acknowledge funding from the Sheffield NIHR Biomedical Research Centre (Translational Neuroscience) (LMW), Neurocare (MDM and AV) and of an Alzheimer Research UK Network small grant scheme (MDM and AV).

Correspondence concerning this article should be addressed to: Annalena Venneri, Department of Life Sciences, Brunel University London, London, UK, Email: annalena.venneri@brunel.ac.uk

Author Contributions:

Laura Wright: Data Curation, Formal Analysis, Funding Acquisition, Investigation, Methodology, Visualisation, Writing – Original Draft

Matteo De Marco: Data Curation, Supervision, Writing - Review and Editing

Annalena Venneri: Conceptualisation, Data Curation, Funding Acquisition, Investigation, Methodology, Project Administration, Resources, Supervision, Writing - Review and Editing

Abstract

Objective: Prior to evidence of episodic memory decline, a lengthy preclinical phase of AD exists characterised by the build-up of tau pathology within extrahippocampal structures. Semantic memory, also impaired in AD, has been linked to degradation within these earliest affected areas. This study aims to assess the utility of performance discrepancies between letter and category verbal fluency tasks to detect neuronal loss in brain regions affected very early by AD. **Methods:** Whole-brain voxel-based morphometry was used to assess the neural correlates of semantic processing in three patient groups: two groups of mild cognitive impairment (MCI) patients split into mildly (n = 58) and moderately (n = 53) affected and a mild AD dementia group (n = 71). Discrepancies between the level of impairment on the semantic category fluency test and non-semantic letter fluency test were calculated for each participant and included in regression models measuring the relationship between semantic memory and whole-brain grey matter volume. **Results:** Patients at all disease stages demonstrated a loss of the normal semantic advantage in fluency tests, showing significantly greater impairments in category relative to letter fluency. Discrepancy scores in mild MCI correlated strongly with the structural integrity of the anterior medial temporal lobes. Correlations in more severely affected groups were weaker and more widespread. **Conclusions:** Semantic memory appears a useful indicator of even the earliest stages of medial temporal damage in AD. With advancing disease severity, the discrepancy index loses its focal anatomical association, reinforcing its value as an early marker of incipient decline.

Keywords: Semantic memory; Alzheimer's disease; preclinical; structural MRI; verbal fluency;

Key Points:

- Question: Can discrepancies in verbal fluency declines identify structural alterations in regions associated with the earliest stages of Alzheimer's disease?
- Findings: In early MCI stages, a category/letter fluency discrepancy score shows a strong correlation with grey matter within areas known to be affected by the initial stages of tau deposition in AD.
- Importance: Changes in semantic memory, identified using a simple verbal fluency test may provide an early and reliable marker for abnormal, AD related cognitive decline.
- Next steps: In the future verbal fluency decline discrepancies should be explored in patients within biomarker confirmation of differing aetiologies to determine the utility of this marker as a tool for early disease differentiation.

Dysfunction of semantic processing has been well established within the literature as representing an extremely early indicator of cognitive decline in Alzheimer's Disease (AD) (Amieva et al., 2008; Vonk et al., 2020; Joubert et al., 2021). Concurrently, many investigations have demonstrated a significant relationship between semantic memory function and structural and functional integrity of anterior medial temporal (aMTL) areas known to be implicated in the initial stages of AD pathological degradation (Barbeau et al., 2012; Kivisaari et al., 2012; Hirni et al., 2013; Gardini et al., 2013; Hirni et al., 2016; Sánchez et al., 2017; Pineault et al., 2018; Venneri et al., 2019; Vonk et al., 2020).

To date, numerous studies have utilised category verbal fluency tasks as a measure of semantic memory when investigating declines in this function among AD patients. It is understood that semantic memory performance relies heavily on the presence of two cognitive processes: storage and access of semantic information and control of semantic retrieval (Troyer et al., 1998; Thompson-Schill et al., 1997; Wagner et al., 2001; Henry & Crawford, 2004a, 2004b; Patterson et al., 2007). During both types of verbal fluency test, phonemic (letter) and semantic (category), engagement of executively mediated controlled retrieval processes, thought to be reliant on frontal lobe structures, is required for successful performance. During category fluency tests however, this retrieval process relies further on the integrity of semantic associations contained within the semantic memory store, thought to be sustained by the temporal lobes (Henry & Crawford, 2004a, 2004b; Vonk et al., 2019). It is now well documented that, although both category and letter fluency are susceptible to AD-type neurodegeneration (Mueller et al., 2015), the semantically-driven category fluency test appears to be consistently impaired to a greater extent (Monsch et al., 1997; Henry, Crawford & Phillips, 2004). Patients, even in prodromal disease stages, have repeatedly shown a loss of the normally observed semantic advantage in verbal fluency performance in healthy individuals (Vaughan et al., 2016; Murphy, Rich & Troyer, 2006; Chasles et al., 2020). Furthermore, longitudinal evaluation of fluency performance has evidenced significantly greater rates of decline in semantic fluency measures compared with phonemic measures in individuals at-risk for AD dementia, in both prodromal and preclinical stages,

with decline discrepancies already appreciable at baseline being significantly related to subsequent progression to dementia (Clark et al., 2009; Papp et al., 2016; Vonk et al., 2020; Marra et al., 2021).

Until recently, there has been limited research in this area investigating the underlying neural correlates that may contribute to such verbal fluency discrepancies in AD. In their longitudinal study, Vonk and colleagues (2020) were, however, able to demonstrate a significant relationship between baseline semantic fluency and semantic decline over time and a number of markers of AD neurodegeneration, including smaller hippocampal volumes, increased white-matter hyperintensities and overall cortical thinning, as well as reduced metabolic functioning within a number of AD-related areas, including the entorhinal cortex, inferior parietal lobule and posterior cingulate gyrus/precuneus. Conversely, no such relationships were found between baseline letter fluency scores and cortical signatures of AD and, although overall indices of neurodegeneration were correlated with the rate of letter fluency decline, this showed no specificity for AD-type alterations.

The present study aimed to assess the relationship between verbal fluency decline discrepancies and brain structural changes across different clinical stages from mild levels of mild cognitive impairment to mild probable AD dementia. Behaviourally, it was expected that all groups, even those at a very mild disease stage, would likely present with a reduction of the semantic advantage in verbal fluency tasks, showing a linear relationship with disease severity. Discrepancy scores were expected to correlate most strongly with aMTL regions in the earliest stages of cognitive impairment, reflecting the involvement of these areas in semantic processing (Didic et al., 2011; Barbeau et al., 2012; Kivisaari et al., 2012; Hirni et al., 2013), but lose this anatomical specificity in patients with more advanced disease stages due to the significant degradation of these areas and the greater involvement of wider temporal neocortices.

Materials and Methods

Participants

A total of 182 cognitively impaired patients were identified retrospectively from a large, multi-centre dataset coordinated by the University of Sheffield, Department of Neuroscience (UK). Patients were recruited through a memory clinic, following referral from their GP, and received a consensus clinical diagnosis following multidisciplinary team review and following published clinical guidelines as highlighted below. All participants underwent extensive clinical assessment, including completion of comprehensive neuropsychological assessment, as well as structural brain imaging, prior to a diagnosis. Seventy-one participants received a diagnosis of probable AD dementia according to the NINCDS-ADRDA clinical criteria (McKhann et al., 2011) and in 111 patients, mild cognitive impairment due to AD, at intermediate level of certainty, was identified following the criteria outlined in Albert et al. (2011). Exclusion criteria in the present study encompassed any diagnostic entity (other than those of interest), medical profile, significant psychiatric condition, or significant pharmacological treatment involving psychotropic medicines that could explain or affect the study outcome. Only those patients who had been followed up longitudinally in clinic and for whom there was support for their initial diagnoses were selected for this retrospective study. All procedures were carried out according to the Declaration of Helsinki. This study received ethical approval from the West of Scotland Regional Ethics Committee 5, Ref No: 19/WS/0177. Written informed consent was obtained from all participants.

MCI patients were further stratified into mild and moderate disease severity groups based on their score on the Mini Mental State Examination (MMSE) (Folstein, Folstein, & McHugh, 1975). This is a clinical instrument that is considered the ‘gold standard’ to quantify general levels of deficit severity (Lladó et al., 2021). In accordance with studies that have found a cut-off of 27 on the MMSE to be a sensitive marker of cognitive decline (O’Bryant et al., 2009; Creavin et al., 2016), participants were categorised as mild (*n* = 58) if

they achieved a score of 27 or more on this cognitive screening measure and as moderate (n = 53) if they achieved a score ≥ 24 but less than 27. The number of patients in each group according to MCI subtype is outlined in **Fig. 1**.

Please insert Fig. 1 about here

The demographic data and MMSE scores of all participants, including a group of 82 healthy older controls with available MRI scans, can be seen in **Table 1**. A *Pearson Chi Square* test revealed no significant differences in the proportion of female and male participants between the groups, χ^2 (32, N = 264) = 5.66, p = .13. A *Kruskal-Wallis H* test revealed no significant differences between groups in terms of age at scan, H(3) = 1.00, p = .80 or years of education, H(3) = 2.71, p = .44.

Please insert Table 1 about here

Neuropsychological Assessment

All participants completed a neuropsychological test battery, including a range of tests measuring semantic memory, episodic memory, speed of processing and executive functions, among others. An exhaustive list of cognitive tests can be seen in **Supplementary Table 1**. Normality of the cognitive data was assessed using Shapiro-Wilk tests. Differences between groups on continuous neuropsychological variables were assessed using a *Kruskall Wallis H* test and *independent-samples t-tests* as appropriate based on normality of the data.

The verbal fluency tests for this study consisted of six one-minute trials. Three category fluency trials (‘Animals’, ‘Fruits’ and either ‘Cities’ or ‘Car Brands’) and three letter fluency trials (‘F’, ‘P’ and ‘L’) were completed by all participants, who were administered the

two parts in a random order. Participants' scores for each test amounted to the sum of the total number of unique words produced within each of the three trials.

To ascertain the relative deficits in letter and category fluency in each of the patient groups, fluency scores were taken from 113 control participants, selected to match with patients for age (*mdn*: 72.00, *IQR*: 11.00), education (*mdn*: 11.00, *IQR*: 5.50) and gender (m/f: 47/66). The data from these healthy older adults were used to create sample-based normative data for both fluency scores. To characterise impairment, the means and standard deviations taken from controls for each of the verbal fluency measures (Total Letter Fluency: *m* = 39.81, *sd* = 14.10; Total Category Fluency: *m* = 47.37, *sd* = 14.55) were used to obtain standardised *z*-scores for the patient data. To control for the effects of shared cognitive processes associated with both verbal fluency measures, the category fluency *z*-scores were subtracted from the letter fluency *z* scores to obtain a discrepancy score for each participant reflecting the relative difference in the amount of impairment on each verbal fluency test (Marra et al., 2021). Normality of the fluency data was assessed using Shapiro-Wilk tests. Between group differences on fluency measures were calculated using independent sample *t-tests* or *Mann-Whitney U* tests and within-group differences between verbal fluency *z*-scores were assessed using paired *t-tests* or *Wilcoxon Signed Rank* tests, as appropriate, based on the normality of the data.

MRI Protocol

Three-dimensional T1-weighted scans were collected from 142 participants using a 1.5 T Philips Achieva scanner with parameters as follows: Turbo Field Echo 3D sequence, voxel dimension 1.1 × 1.1 × 0.6 mm, field of view 250 mm, matrix size 256 × 256 × 124, repetition time: 7.4 ms, echo delay time: 3.4 ms and flip angle: 8°. The remaining 122 participants were scanned using a Philips Ingenia 3.0 T scanner with parameters as follows: voxel dimension .94 × .94 × 1.0 mm, field of view 256 mm, matrix size 256 × 256 × 124, repetition time 8.2 ms, echo time: 3.8 ms, and flip angle 8°.

Pre-processing Procedures

Using SPM12 software run in a Matlab environment (version R2011b; Mathworks Inc., UK), the latest standard Voxel Based Morphometry (VBM) preprocessing procedures (Ashburner & Friston, 2000) were applied to 182 anatomical scans taken from patients, as well as 82 scans that were available for control participants. Images were reoriented, segmented into the three major tissue classes, grey matter, white matter, and cerebrospinal fluid, registered to the Montreal Neurological Institute (MNI) template, modulated, and smoothed with a 10-mm full-width at half-maximum Gaussian filter. Total intracranial volumes were calculated as the sum of the volumes quantified for each of the tissue class maps in the native space using the "get_totals" Matlab function (http://www0.cs.ucl.ac.uk/staff/g.ridgway/vbm/get_totals.m).

To account for group variability in global signal intensity, a 'contrast-to-noise ratio' (CNR) index was calculated for each preprocessed image. Spherical regions of 2 mm radius were defined: 1) in the 'trunk region' of the primary motor cortex (M1), following the coordinates of the Human Brainnettome Atlas (Fan et al., 2016): [x = -13, y = -20, z = 73] and [x = 15, y = -22, z = 71]; 2) in ventricular seeds of 2 mm radius containing no neural tissue, centred at [x = -3, y = 8, z = 12] and [x = 4, y = 8, z = 12]. The MarsBaR toolbox (https://marsbar-toolbox.github.io/) was used for this purpose. MarsBaR was then also used to extract the average signal from the combined M1 seeds, the average signal from the combined ventricular seeds, and signal variance from the ventricular seeds. CNR was calculated as the difference between M1 and ventricular signal, divided by the standard deviation of ventricular signal.

Whole Brain VBM Analytical Procedures

The relationship between grey matter volumes and verbal fluency discrepancy scores was assessed separately for each group using whole-brain regression analyses carried out in SPM12. Analyses were run using the discrepancy scores as the independent variable.

The dependent variable in all cases was the grey matter volume determined by VBM and all multiple regression models included age at scan, years of education, MMSE scores, total intracranial volume (ml) and CNR as covariates. Although MMSE scores were at the basis of group stratification, they were added to the models to control for the possibility that the results found were due to variances in overall cognitive impairment. The threshold of significance chosen in this study was an uncorrected set-level *p* value equal to .001. Clusters surviving a cluster-level Family-Wise Error-corrected (FWE) $p < .05$ were the only observed clusters considered for interpretation. Peak coordinates of clusters surviving the FWE ($p < .05$) were converted into Talairach space using a non-linear transform (http://imaging.mrc-cbu.cam.ac.uk/downloads/MNI2tal/mni2tal.m) and were interpreted using the Talairach Daemon client (http://www.talairach.org/client.html), selecting the "Nearest Grey Matter" search option (Lancaster et al., 2000). Findings were interpreted only at a cluster level to capture significant effects at group diagnostic level and avoid peak-level bias due to inter-individual variability in atrophy patterns.

Transparency and openness

Journal Article Reporting Standards have been followed and how all data exclusions (if any), all manipulations, and all measures used in the study were determined is reported. Data will be made available on request. The study was not pre-registered.

Results

Neuropsychological Comparisons

Significant differences between groups on the neuropsychological test battery are reported in **Supplementary Table 1**. Missing data, due to patients being unable to complete testing, are reflected in the number of participants listed along with each test. Statistical comparisons between patients and the 82 control participants who were included in the imaging analysis revealed a pattern of impairment in each of the patient groups that is

consistent with an AD diagnosis. A *Kruskall-Wallis H* test, with post-hoc *Dunn* and *independent-sample t-tests*, revealed that all patients scored significantly below controls on a number of episodic and semantic memory tests. No differences were found on episodic memory performance between the mild and moderate MCI groups. Only performance on the Token Test (corrected $p = .03$) and number of errors on the Stroop Task (corrected $p = .02$) demonstrated significant differences between these groups when applying a Bonferroni correction for multiple comparisons (medians and inter-quartile ranges are reported in **Supplementary Figure 1**). Differences between these groups were also found in letter fluency performance ($p = .01$), digit cancellation ($p = .02$), copying of the Rey-Osterrieth complex figure ($p = .009$) and digit span forward ($p = .04$) although none survived correction for multiple comparisons. Post-hoc *independent-samples t-tests* also showed significant differences between these groups on the similarities subset of the Wechsler Adult Intelligence Scale (mild MCI: $m = 18.59$, $sd = 4.36$, moderate MCI: $m = 16.26$, $sd = 4.33$, $t(109) = 2.81$, $p = .006$) and the Raven's progressive matrices (mild MCI: $m = 26.78$, $sd = 4.94$, moderate MCI: $m = 24.02$, $sd = 5.28$, $t(109) = 2.85$, $p = .005$) although these also did not survive correction for multiple comparisons.

Neuroimaging Comparison with Controls

When compared with a control group, the two groups of MCI patients had significantly reduced grey-matter volumes in the mediotemporal lobe, bilaterally, while the group of patients with a diagnosis of dementia had extensive atrophy across the majority of cortical regions. These neuroanatomical profiles were consistent with clinical diagnoses and provided group-level, neuroimaging-informed confirmation of the suspected aetiology. The methodology at the basis of these subsidiary descriptive analyses of the sample and an illustration of these atrophy profiles are included in the **Supplementary Material**.

Verbal Fluency z-Scores

Significant differences between patient groups were observed in both letter and category fluency *z*-scores. Medians and inter-quartile ranges are reported in **Fig. 2** and **Supplementary Table 1**. Means and standard deviations are reported in **Table 2**. The dementia group tended to perform significantly further below the control mean on the category fluency test than both the mild MCI group (U = 401, p <.001) and the moderate MCI group (t(122) = 5.38, p <.001). They also performed significantly further below the control mean on the letter fluency test compared with the mild MCI group (U = 860, p <.001). A *Mann-Whitney U* test also found a modest difference between the moderate MCI and dementia groups on letter fluency *z* scores (U = 1455, p = .03), although this did not survive correction for multiple comparisons. The moderate MCI group also demonstrated impairments in category fluency that were significantly further from the control mean than the mild MCI group (U = 885, p <.001). Likewise, these groups demonstrated similar differences on letter fluency *z* scores (U = 925, p < .001). A summary of these findings in outlined in **Fig. 2** and **Table 2**.

Within-group analyses, using either paired *t-tests* or *Wilcoxon Signed Rank* tests depending on data normality, were carried out to assess the difference in the relative impairment on each type of verbal fluency at each stage of disease progression. All groups demonstrated significantly lower category fluency *z*-scores than letter fluency z scores (mild MCI: z = -4.17, p < .001, moderate MCI: t(52) = 4.41, p < .001, dementia: z = -6.66, p < .001). A summary of these findings is outlined in **Fig. 2**.

Please insert Table 2 about here

Please insert Fig. 2 about here

No differences in discrepancy scores were found between the mild and moderate MCI groups $t(109) = .25$, $p = .80$ (means and standard deviations reported in **Table 2**). The dementia group showed higher discrepancies in verbal fluency impairment when compared with both the mild MCI ($U = 1570$, $p = .021$) and moderate MCI groups ($U = 1491$, $p = .049$), although these did not survive correction for multiple comparisons (medians and inter-quartile range reported in **Fig. 2**).

Whole Brain Imaging Results

Mild MCI

In the mild MCI group, verbal fluency discrepancy scores correlated negatively with regions of the right anterior temporal lobes, including the right uncus, anterior parahippocampal gyrus/perirhinal cortex (BA 36, 35), hippocampus and some lateral areas of temporal neocortex (BA 20, 21) (**Table 3** and **Fig. 3**).

Please insert Fig. 3 about here

Please insert Table 3 about here

Moderate MCI

No significant clusters were found in the moderate MCI group when controlling for FWE. However, a trend of negative correlation between discrepancy scores and grey matter volume was revealed at the uncorrected level (thresholded $p < .05$) in areas of the temporal lobe including aMTL areas such as perirhinal regions of anterior parahippocampal gyrus (BA 36) and temporal pole (BA 38) as well as more lateral areas such as the inferior and middle temporal and fusiform gyri (BA 20 and 21). These areas are highlighted in **Fig. 3**.

Dementia

In the dementia patient group, a variety of areas was found to be negatively correlated with the verbal fluency discrepancy scores (**Table 3** and **Fig. 3**). These included

bilateral regions of the occipital cortex, including lingual and middle occipital gyri, as well as the cuneus (BA 18), but also extended further to include posterior temporal regions such as areas of parahippocampal (BA 19) and the right superior temporal gyrus (BA 22). Further correlations with parietal areas in the proximity of the central sulcus were also observed bilaterally (BA 7).

Discussion

The findings of this study indicated that there is a significant loss of the semantic advantage in verbal fluency tasks across the AD spectrum, even at the earliest prodromal stages. Furthermore, this loss shows a specific and focused negative correlation with grey matter volumes in aMTLs at the mild MCI stage that is progressively lost with increasing disease severity.

At the behavioural level, the present study successfully replicated findings of previous work that has demonstrated that declines in verbal fluency in MCI and AD cohorts are often characterised by significantly greater impairments in category fluency relative to letter fluency (Henry, Crawford & Phillips, 2004; Murphy, Rich & Troyer, 2006; Clark et al., 2009; Chasles et al., 2020; Vonk et al., 2020; Marra et al., 2021). Not only was this found to be the case at all disease stages, but in both the moderate MCI and dementia groups, category fluency *z* scores reached below clinical cut off (i.e., 1.5 standard deviations from the control mean) while letter fluency z-scores did not, suggesting that patients, even in the later stages of disease, are impaired to a lesser extent on this non-semantic verbal fluency test when compared with tests requiring additional semantic retrieval (see De Marco et al., 2017, for a detailed discussion on this pattern of findings). In contrast with previous studies however, which have shown a reduction of verbal fluency discrepancies with increasing disease severity (Marra et al., 2021), the present study demonstrated equal, if not slightly greater, discrepancies in fluency decline among dementia patients when compared with MCI groups. Unlike the study by Marra and colleagues, in which dementia patients were split into mild and moderate groups, with MMSE scores spanning 18-23 and 10-17 respectively, most

patients included in the present study were in the very early stages of dementia with interquartile range of 4. It is likely, therefore, that dementia patients included in the present study were more closely aligned, in terms of disease severity, with the group of aMCI converters described by Marra et al., hence showing an exacerbation, rather than reduction, of discrepancy scores compared with MCI. Furthermore, the discrepancy scores described in the present study do not equate to the semantic–phonological delta (SPD) described by Marra and colleagues. In Marra et al's study, the SPD reflected the difference in raw fluency scores, corrected for the number of items in each task. Therefore, verbal fluency discrepancies were not, as they are in the present study, operationalised in terms of the level of decline relative to healthy controls. It is expected that, as disease progresses, patients will likely perform similarly on each of the fluency tests due to greater impairments in executive functioning and overall effortful retrieval (Henry, Crawford & Phillips, 2004). However, due to the semantic advantage demonstrated by healthy adults (Vaughan et al., 2016), even in cases in which raw performance on both tests is similar, category fluency performance will likely, as is demonstrated in the present study, be substantially further from the level expected for a given age and education level than performance on letter fluency.

Interestingly, no statistically significant differences were found between the mild and moderate MCI groups on any of the standard episodic memory tests. Such findings are in line with longitudinal studies of MCI patients that have found verbal fluency discrepancies to be a better predictor of progression to dementia in these individuals than episodic memory tests such as the Rey Auditory Verbal Learning Task (Marra et al., 2021). This supports the theory put forward by Didic et al., (2011) that suggests that, in early AD, semantic memory decline likely occurs as a result of the initial progressive degradation of subhippocampal cortical structures, whereas episodic memory decline may remain relatively stable, only becoming increasingly exacerbated at a stage in which degeneration of the hippocampus intensifies. In the earliest stages of disease, currently relied upon tests of episodic memory, it may, therefore, not be sufficient to separate early from late-stage cognitive impairment in a prodromal group. Tests involving language and semantic processing however, such as the

similarities and Token tests, which did show differences between the stages of MCI in this study, as well as declines on fluency tests, could be more informative in identifying the subtle progressions in underlying pathology in early disease stages, and represent good prognostic indicators for further cognitive decline (Clark et al., 2009; Vonk et al., 2020; Marra et al., 2021).

VBM Findings

In line with the initial hypothesis, discrepancy scores in the mild MCI group correlated predominately with areas of the aMTLs, specifically, with areas of the right perirhinal regions, as well as the amygdala and uncus. Discrepancy scores in this cohort, therefore, appear to capture accurately degradation in areas related to semantic memory processing, highlighting inferior and medial regions of the anterior temporal lobes, thought to be integral to this function (Patterson et al., 2007; Venneri et al., 2008; Visser et al., 2010; Kivisaari et al., 2012). The right lateralisation of correlations in this group is interesting, given the left lateralised nature of language functions (Vigneau et al., 2006). However, the verbal fluency discrepancy score serves to isolate semantic retrieval processes by controlling for any variance explained by differences in language production ability that may affect both fluency tests. Semantic processing has been shown to relate to brain structure and function bilaterally, particularly within anterior temporal regions (Barbeau et al., 2012; Binder & Desai, 2011). The right sided lateralisation of these results may, therefore, be related to the pattern of atrophic change seen in the earliest stages of AD, which has been found by previous studies to follow an asymmetric pattern in the very early stages, with the left hemisphere showing an initial acceleration in degradation before atrophy evens out bilaterally (Thompson et al., 2003; Shi et al., 2009). Right lateralised correlations with discrepancy scores may therefore represent the relationship between semantic functioning and brain volume in more well preserved medial temporal regions.

In accordance with Didic's model of memory impairment in AD, the present study demonstrates how the semantic component of a verbal fluency test, isolated by decline

discrepancies, can accurately illustrate degradation of discrete areas of rhinal cortices associated with the very earliest moments of AD pathology (Braak & Braak, 1991). Perirhinal involvement in semantic forms of declarative memory has been repeatedly demonstrated by animal lesion studies (Meunier et al., 1993; Mumby & Pinel, 1994; Brown & Aggleton, 2001; Barker et al., 2007). Such research has demonstrated that ablations to the rhinal cortices, without involvement of hippocampal formation, can cause significant impairments on recognition memory tasks. Furthermore, in an early study, Meunier et al. (1993) found that even when such lesions were limited to the perirhinal cortex, impairments of the same severity were observed, a finding that was not replicated when ablations were restricted to the entorhinal cortex. This observation, therefore, suggests that the perirhinal cortex in particular, plays a significant role in context-free memory processing. Studies of human participants with frontotemporal lobar degeneration have similarly demonstrated that the extent of degradation within the perirhinal cortex directly correlates with the severity of a semantic memory deficit (Davies et al., 2004). Similarly, studies examining amnestic patients with discrete damage to the hippocampus, but relative sparing of the perirhinal cortex, have demonstrated that patients of this description retain the ability to acquire new semantic knowledge despite significant damage to the hippocampal formation (Corkin et al., 1997; Mishkin et al., 1998; Vargha-Khadem et al., 1997), signifying the cognitive and neural dissociation of semantic and episodic memory processes. Many previous neuroimaging investigations have also shown a significant relationship between semantic memory function and structural involvement of the perirhinal cortex among AD dementia and MCI cohorts (Kivisaari et al., 2012; Barbeau et al., 2012; Hirni et al., 2013; Hirni et al., 2016). In particular, attributes of category fluency performance in AD patients have been found to be useful in detecting degeneration within rhinal cortices and the wider aMTLs (Venneri et al., 2008). In their recent study, Vonk and colleagues (2020) were successful in showing that both baseline category fluency performance, as well as declines in category fluency over time, correlated with neuroimaging indices of AD related change, finding no such associations with letter fluency. The present study was able to replicate and extend these

findings by demonstrating the utility of discrepancy scores, which inherently control for the non-semantic aspects of verbal fluency measures, to detect directly variances in structural loss within areas of the aMTL known to be affected by the earliest known stages of tau deposition in AD, even when implementing a whole-brain approach. These results, therefore, confirm and extend previous hypotheses relating fluency discrepancies in AD to damage within the temporal lobes (Henry & Crawford, 2004a, 2004b; Henry, Crawford & Philips, 2004) and, in accordance with previous studies (Venneri et al., 2008; Barbeau et al., 2012; Kivisaari et al., 2012; Hirni et al., 2013), contribute to the assertion that semantic and episodic memory functions are sufficiently distinct, in terms of both their cognitive processes and their neural bases, in order that semantic tasks may provide a means to identify AD related pathological change within rhinal cortices prior to hippocampal involvement.

The non significant trends seen among the moderate MCI group within areas of the MTLs and the more widespread posterior temporal and occipital correlations found in the dementia patients suggest that the specificity of verbal fluency discrepancy scores to detect structural damage within aMTLs is most optimal in the earliest stages of disease. Previous imaging studies (Binder et al., 2009; Visser et al., 2010), models of the semantic system (Patterson et al., 2007; Binder & Desai, 2011) and models describing visual object recognition (Saksida & Bussey, 2010; Kivisaari et al., 2012) together point to a hierarchical organisation of semantic processing within the temporal lobes, in which anterior structures represent the most specified regions. This shift in grey matter correlates of semantic memory towards posterior temporal structures, in both the moderate MCI and dementia groups, could therefore be thought of as representative of increased reliance on lower-level posterolateral temporal structures to facilitate semantic retrieval in the presence of significant medial and anterior temporal atrophy. Such an observation is, therefore, supportive of the notion that loss of the semantic advantage in AD may represent a specific marker for focal degeneration associated only with the earliest stages of neurodegeneration.

Limitations

A possible limitation of the present study is the reliance on clinical diagnosis to define patients and the lack of biomarker confirmation. However, all patients were followed up extensively and only those who showed no indications of improvement or evidence of clinical entities other than AD were included in the final sample. Furthermore, a proportion of patients (around 25%) who entered clinical trials or further studies have since been found to demonstrate amyloid positivity in cerebrospinal fluid or on positron emission tomography. It is, therefore, unlikely that any diagnostic errors, although still a theoretical possibility, might have contaminated the findings of this study, as the extensive period of longitudinal clinical observation has greatly limited any potential confound due to clinical error.

Conclusions

The current study confirms and extends the results of previous research indicating that a significant decline in semantic memory, apparent on measures of verbal fluency, is present even in the earliest prodromal stages of AD. These findings further support the assertion that a significant discrepancy in the amount of decline on semantic and phonemic fluency tests is present throughout the course of disease (Henry, Crawford & Phillips, 2004; Murphy, Rich & Troyer, 2006; Amieva et al., 2008; Clark et al., 2009; Gardini et al., 2013; Chasles et al., 2020; Vonk et al., 2020; Joubert et al., 2021; Marra et al., 2021). Imaging analyses have corroborated previous findings that suggest the semantic deficit in prodromal AD is likely underpinned by pathological changes within the aMTL, occurring in the initial stages of disease (Barbeau et al., 2012; Hirni et al., 2013), and further demonstrated a pattern of pathological progression, across the disease spectrum, that is traceable via semantic memory measures. These analyses have demonstrated for the first time that the specificity of semantic memory decline for identifying structural variance of the aMTLs is greatest in the earliest stages of disease and is lost with increasing disease severity. Together with longitudinal behavioural analyses demonstrating the power of this marker for

identifying incipient progression to dementia (Marra et al., 2021), the work presented here, therefore, supports the utility of fluency discrepancies as a proxy for prehippocampal neurodegeneration in AD.

References

Albert, M. S., DeKosky, S. T., Dickson, D., Dubois, B., Feldman, H. H., Fox, N. C., Gamst, A., Holtzman, D. M., Jagust, W. J., Petersen, R. C., & Phelps, C. H. (2011). The diagnosis of mild cognitive impairment 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), 270-279. https://doi.org/10.1016/j.ppees.2011.04.003

Amieva, H., Le Goff, M., Millet, X., Orgogozo, J. M., Pérès, K., Barberger-Gateau, P., Jacqmin-Gadda, H., & Dartigues, J. F. (2008). Prodromal Alzheimer's disease: Successive emergence of the clinical symptoms. *Annals of Neurology*, 64(5), 492-498. https://doi.org/10.1002/ana.21509

Ashburner, J., & Friston, K. J. (2000). Voxel-based morphometry—the methods. *Neuroimage*, 11(6), 805-821. https://doi.org/10.1006/nimg.2000.0582

Barbeau, E. J., Didic, M., Joubert, S., Guedj, E., Koric, L., Felician, O., Ranjeva, J. P., Cozzone, P., & Ceccaldi, M. (2012). Extent and neural basis of semantic memory impairment in mild cognitive impairment. *Journal of Alzheimer's Disease*, 28(4), 823-837. https://doi.org/10.3233/JAD-2011-110989

Barker, G. R., Bird, F., Alexander, V., & Warburton, E. C. (2007). Recognition memory for objects, place, and temporal order: A disconnection analysis of the role of the medial prefrontal cortex and perirhinal cortex. *Journal of Neuroscience*, 27(11), 2948-2957. https://doi.org/10.1523/JNEUROSCI.5289-06.2007

Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex*, 19(12), 2767-2796. https://doi.org/10.1093/cercor/bhp055

Binder, J. R., & Desai, R. H. (2011). The neurobiology of semantic memory. *Trends in Cognitive Sciences*, *15*(11), 527-536. https://doi.org/10.1016/j.tics.2011.10.001

Braak, H., & Braak, E. (1991). Neuropathological stageing of Alzheimer-related changes. *Acta Neuropathologica*, 82(4), 239-259. https://doi.org/10.1007/BF00308809

Brown, M. W., & Aggleton, J. P. (2001). Recognition memory: what are the roles of the perirhinal cortex and hippocampus? *Nature Reviews Neuroscience*, *2*(1), 51-61. https://doi.org/10.1038/35049064

Chasles, M. J., Tremblay, A., Escudier, F., Lajeunesse, A., Benoit, S., Langlois, R., Joubert, S., & Rouleau, I. (2020). An examination of semantic impairment in amnestic MCI and AD: What can we learn from verbal fluency? *Archives of Clinical Neuropsychology*, 35(1), 22-30. https://doi.org/10.1093/arclin/acz018

Clark, L. J., Gatz, M., Zheng, L., Chen, Y. L., McCleary, C., & Mack, W. J. (2009). Longitudinal verbal fluency in normal aging, preclinical, and prevalent Alzheimer's disease. *American Journal of Alzheimer's Disease & Other Dementias*, 24(6), 461-468. https://doi.org/10.1177/1533317509345154

Corkin, S., Amaral, D. G., González, R. G., Johnson, K. A., & Hyman, B. T. (1997). HM's medial temporal lobe lesion: findings from magnetic resonance imaging. *Journal of Neuroscience*, 17(10), 3964-3979. https://doi.org/10.1523/JNEUROSCI.17-10-03964.1997

Creavin, S. T., Noel-Storr, A. H., Smailagic, N., Giannakou, A., Ewins, E., Wisniewski, S., & Cullum, S. (2016). Mini-Mental State Examination (MMSE) for the detection of Alzheimer's dementia and other dementias in asymptomatic and previously clinically unevaluated people aged over 65 years in community and primary care populations. *Cochrane Database of Systematic Reviews*. Chichester, UK: John Wiley & Sons, Ltd. https://doi.org/10.1002/14651858.CD011145.pub2

Davies, R. R., Graham, K. S., Xuereb, J. H., Williams, G. B., & Hodges, J. R. (2004). The human perirhinal cortex and semantic memory. *European Journal of Neuroscience*, 20(9), 2441-2446. https://doi.org/10.1111/j.1460-9568.2004.03710.x

De Marco, M., Duzzi, D., Meneghello, F., & Venneri, A. (2017). Cognitive efficiency in Alzheimer's disease is associated with increased occipital connectivity. *Journal of Alzheimer's Disease*, 57(2), 541-556. https://doi.org/10.3233/JAD-161164

Didic, M., Barbeau, E. J., Felician, O., Tramoni, E., Guedj, E., Poncet, M., & Ceccaldi, M., 2011. Which memory system is impaired first in Alzheimer's disease? *Journal of Alzheimer's Disease*, 27(1), 11-22. https://doi.org/10.3233/JAD-2011-110557

Fan, L., Li, H., Zhuo, J., Zhang, Y., Wang, J., Chen, L., Yang, Z., Chu, C., Xie, S., Laird, A. R., Fox, P. T., Eickhoff, S. B., Yu, C., & Jiang, T., 2016. The Human Brainnetome Atlas: A new brain atlas based on connectional architecture. *Cerebral Cortex,* 26(8), 3508-3526. https://doi.org/10.1093/cercor/bhw157

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. https://doi.org/10.1016/0022-3956(75)90026-6

Gardini, S., Cuetos, F., Fasano, F., Ferrari Pellegrini, F., Marchi, M., Venneri, A., & Caffarra, P. (2013). Brain structural substrates of semantic memory decline in mild cognitive impairment. *Current Alzheimer Research*, 10(4), 373-389. https://doi.org/10.2174/1567205011310040004

Henry, J. D., & Crawford, J. R. (2004a). A meta-analytic review of verbal fluency performance in patients with traumatic brain injury. *Neuropsychology*, 18(4), 621-628. https://doi.org/10.1037/0894-4105.18.4.621

Henry, J. D., & Crawford, J. R. (2004b). A meta-analytic review of verbal fluency performance following focal cortical lesions. *Neuropsychology*, 18(2), 284-295. https://doi.org/10.1037/0894-4105.18.2.284

Henry, J. D., Crawford, J. R., & Phillips, L. H. (2004). Verbal fluency performance in dementia of the Alzheimer's type: A meta-analysis. *Neuropsychologia*, 42(9), 1212-1222. https://doi.org/10.1016/j.neuropsychologia.2004.02.001

Hirni, D. I., Kivisaari, S. L., Krumm, S., Monsch, A. U., Berres, M., Oeksuez, F., Reinhardt, J., Ulmer, S., Kressig, R. W., Stippich, C. & Taylor, K. I. (2016). Neuropsychological

markers of medial perirhinal and entorhinal cortex functioning are impaired twelve years preceding diagnosis of Alzheimer's dementia. *Journal of Alzheimer's Disease*, 52(2), 573-580. https://doi.org/10.3233/JAD-150158

Hirni, D. I., Kivisaari, S. L., Monsch, A. U., & Taylor, K. I. (2013). Distinct neuroanatomical bases of episodic and semantic memory performance in Alzheimer's disease. *Neuropsychologia*, 51(5), 930-937. https://doi.org/10.1016/j.neuropsychologia.2013.01.013

Joubert, S., Gardy, L., Didic, M., Rouleau, I., & Barbeau, E. J. (2021). A meta-analysis of semantic memory in mild cognitive impairment. *Neuropsychology Review*, 31(2), 221-232. https://doi.org/10.1007/s11065-020-09453-5

Kivisaari, S. L., Tyler, L. K., Monsch, A. U., & Taylor, K. I. (2012). Medial perirhinal cortex disambiguates confusable objects. *Brain*, 135(12), 3757-3769. https://doi.org/10.1093/brain/aws277

Lancaster, J. L., Woldorff, M. G., Parsons, L. M., Liotti, M., Freitas, C. S., Rainey, L., Kochunov, P. V., Nickerson, D., Mikiten, S.A., & Fox, P. T. (2000). Automated Talairach atlas labels for functional brain mapping. *Human Brain Mapping*, 10(3), 120-131. https://doi.org/10.1002/1097-0193(200007)10:3<120::aid-hbm30>3.0.co;2-8

Lladó, A., Froelich, L., Khandkher, R. K., Roset, M., Black, C. M., Lara, N., Chekani, F., & Ambegaonkar, B. M. (2021). Assessing the progression of Alzheimer's disease in real-world settings in three European countries, *Journal of Alzheimer's Disease*, 80(2), 749-759. https://doi.org/10.3233/JAD-201172

Marra, C., Piccininni, C., Iacobucci, G. M., Caprara, A., Gainotti, G., Maria Costantini, E., Callea, A., Venneri, A., & Quaranta, D. (2021). Semantic memory as an early cognitive marker of Alzheimer's disease: Role of category and phonological verbal fluency tasks. *Journal of Alzheimer's Disease*, 81(2), 619-627. https://doi.org/10.3233/JAD-201452

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., & Phelps, C. H. (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. https://doi.org/10.1016/j.jalz.2011.03.005

Meunier, M., Bachevalier, J., Mishkin, M., & Murray, E. A. (1993). Effects on visual recognition of combined and separate ablations of the entorhinal and perirhinal cortex in rhesus monkeys. *Journal of Neuroscience*, 13(12), 5418-5432. https://doi.org/10.1523/JNEUROSCI.13-12-05418.

Mishkin, M., Vargha-Khadem, F., & Gadian, D. G. (1998). Amnesia and the organization of the hippocampal system. *Hippocampus*, 8(3), 212-216. https://doi.org/10.1002/(SICI)1098-1063(1998)8:3<212::AID-HIPO4>3.0.CO;2-L

Monsch, A. U., Seifritz, E., Taylor, K. I., Ermini-Fünfschilling, D., Staehelin, H. B., & Spiegel, R. (1997). Category fluency is also predominantly affected in Swiss Alzheimer's disease patients. *Acta Neurologica Scandinavica*, 95(2), 81-84. https://doi.org/10.1111/j.1600-0404.1997.tb00073.x

Mueller, K. D., Koscik, R. L., LaRue, A., Clark, L. R., Hermann, B., Johnson, S. C., & Sager, M. A. (2015). Verbal fluency and early memory decline: results from the Wisconsin registry for Alzheimer's prevention. *Archives of Clinical Neuropsychology*, 30(5), 448-457. https://doi.org/10.1093/arclin/acv030

Mumby, D. G., & Pinel, J. P. (1994). Rhinal cortex lesions and object recognition in rats. *Behavioral Neuroscience*, 108(1), 11-18. https://doi.org/10.1037/0735-7044.108.1.11

Murphy, K. J., Rich, J. B., & Troyer, A. K. (2006). Verbal fluency patterns in amnestic mild cognitive impairment are characteristic of Alzheimer's type dementia. *Journal of the International Neuropsychological Society*, 12(4), 570-574. https://doi.org/10.1017/S1355617706060590

O'Bryant, S. E., Humphreys, J. D., Smith, G. E., Ivnik, R. J., Graff-Radford, N. R., Petersen, R. C., & Lucas, J. A. (2008). Detecting dementia with the Mini-Mental State

Examination. Archives of Neurology, 65(7), 963-967. https://doi.org/10.1001/archneur.65.7.963

Papp, K. V., Mormino, E. C., Amariglio, R. E., Munro, C., Dagley, A., Schultz, A. P., Johnson, K. A., Sperling, R. A., & Rentz, D. M. (2016). Biomarker validation of a decline in semantic processing in preclinical Alzheimer's disease. *Neuropsychology*, 30(5), 624–630. https://doi.org/10.1037/neu0000246

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), 976-987. https://doi.org/10.1038/nrn2277

Pineault, J., Jolicoeur, P., Grimault, S., Bermudez, P., Brambati, S. M., Lacombe, J., Villalpando, J. M., Kergoat, M. J. & Joubert, S. (2018). Functional changes in the cortical semantic network in amnestic mild cognitive impairment. *Neuropsychology*, 32(4), 417-435. https://doi.org/10.1037/neu0000466

Sánchez, S. M., Abulafia, C., Duarte-Abritta, B., de Guevara, M., Castro, M. N., Drucaroff, L., Sevlever, G., Nemeroff, C. B., Vigo, D. E., Loewenstein, D. A., & Guinjoan, S. M. (2017). Failure to recover from proactive semantic interference and abnormal limbic connectivity in asymptomatic, middle-aged offspring of patients with late-onset Alzheimer's disease. *Journal of Alzheimer's Disease*, 60(3), 1183-1193. https://doi.org/10.3233/JAD-170491

Saksida, L. M., & Bussey, T. J. (2010). The representational–hierarchical view of amnesia: Translation from animal to human. *Neuropsychologia*, 48(8), 2370-2384. https://doi.org/10.1016/j.neuropsychologia.2010.02.026

Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. K., & Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: A reevaluation. *Proceedings of the National Academy of Sciences*, 94(26), 14792-14797. https://doi.org/10.1073/pnas.94.26.14792

Troyer, A. K., Moscovitch, M., Winocur, G., Alexander, M. P., & Stuss, D. O. N. (1998). Clustering and switching on verbal fluency: The effects of focal frontal-and temporal-

lobe lesions. *Neuropsychologia*, 36(6), 499-504. https://doi.org/10.1016/S0028-3932(97)00152-8

Vargha-Khadem, F., Gadian, D. G., Watkins, K. E., Connelly, A., Van Paesschen, W., & Mishkin, M. (1997). Differential effects of early hippocampal pathology on episodic and semantic memory. *Science*, 277(5324), 376-380. https://doi.org/10.1126/science.277.5324.376

Vaughan, R. M., Coen, R. F., Kenny, R., & Lawlor, B. A. (2016). Preservation of the semantic verbal fluency advantage in a large population-based sample: Normative data from the TILDA study. *Journal of the International Neuropsychological Society*, 22(5), 570-576. https://doi.org/10.1017/S1355617716000291

Venneri, A., McGeown, W. J., Hietanen, H. M., Guerrini, C., Ellis, A. W., & Shanks, M. F. (2008). The anatomical bases of semantic retrieval deficits in early Alzheimer's disease. *Neuropsychologia*, 46(2), 497-510. https://doi.org/10.1016/j.neuropsychologia.2007.08.026

Venneri, A., Mitolo, M., Beltrachini, L., Varma, S., Della Pietà, C., Jahn-Carta, C., Frangi, A. F., & De Marco, M. (2019). Beyond episodic memory: Semantic processing as independent predictor of hippocampal/perirhinal Vol. in aging and mild cognitive impairment due to Alzheimer's disease. *Neuropsychology*, 33(4), 523–533. https://doi.org/10.1037/neu0000534

Visser, M., Jefferies, E., & Lambon Ralph, M. A. (2010). Semantic processing in the anterior temporal lobes: A meta-analysis of the functional neuroimaging literature. *Journal of Cognitive Neuroscience*, 22(6), 1083-1094. https://doi.org/10.1162/jocn.2009.21309

Vonk, J. M., Bouteloup, V., Mangin, J. F., Dubois, B., Blanc, F., Gabelle, A., Ceccaldi, M., Annweiler, C., Krolak-Salmon, P., Belin, C. & Rivasseau-Jonveaux, T. (2020). Semantic loss marks early Alzheimer's disease-related neurodegeneration in older adults without dementia. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, 12(1), e12066. https://doi.org/10.1002/dad2.12066

Vonk, J. M., Rizvi, B., Lao, P. J., Budge, M., Manly, J. J., Mayeux, R., & Brickman, A. M. (2019). Letter and category fluency performance correlates with distinct patterns of cortical thickness in older adults. *Cerebral Cortex*, 29(6), 2694-2700. https://doi.org/10.1093/cercor/bhy138

Wagner, A. D., Paré-Blagoev, E. J., Clark, J., & Poldrack, R. A. (2001). Recovering meaning: left prefrontal cortex guides controlled semantic retrieval. *Neuron*, 31(2), 329-338. https://doi.org/10.1016/S0896-6273(01)00359-2

Table 1. Demographic Characteristics

	Controls (*n* = 82)	**Mild MCI (*n* = 58)**	**Moderate MCI (*n* = 53)**	**AD Dementia (*n* = 71)**
Age (Years)	75.00 (8.50)	75.00 (9.00)	74.00 (14.00)	74.00 (19.00)
Years of Education	11.00 (5.00)	10.00 (5.00)	10.00 (5.00)	11.00 (8.00)
Gender (M/F)	31/51	21/37	20/33	38/33
MMSE	29.00 (3.00)	28.00 (2.00)	25.00 (1.00)[ab]	21.00 (4.00)[abc]

Gender ratios, medians (and interquartile range) for age, years of education and MMSE scores are presented for all patient groups and a group of matched controls with available MRI scans. Between-group differences assessed with a *Kruskall-Wallis H* test. Gender-ratio differences calculated with a χ^2 test. MMSE: Mini Mental State Examination; M/F: Male/Female. [a] Significantly lower than controls at a $p < 0.001$; [b] Significantly lower than mild MCI $p < 0.001$; [c] Significantly lower than Moderate MCI at a $p < 0.001$.

Table 2. Verbal Fluency Measures

	Mild MCI (*n* = 58)	**Moderate MCI (*n* = 53)**	**AD Dementia (*n* = 71)**
Letter Fluency *z*-Scores	-0.26 (1.17)	-0.95 (1.00)[a]	-1.35 (0.97)[a]
Category Fluency *z*-Scores	-0.90 (0.87)	-1.54 (0.86)[a]	-2.35 (0.81)[ab]
Discrepancy Scores	0.64 (1.25)	0.58 (0.97)	1.01 (0.92)

Group means (standard deviations) of verbal fluency measures. Significant differences (corrected for multiple comparisons) were calculated using independent samples t-tests and Mann-Whitney U tests as appropriate. Significant differences determined by *Mann-Whitney U* tests are in bold. [a] Significantly lower than mild MCI at a $p < 0.001$; [b] Significantly lower than moderate MCI at a $p < 0.001$

Table 3. Correlational Findings

Group	Brain Region (BA)	Side	Cluster-Level *pFWE*	Cluster Extent (voxels)	Peak-level *Z* Score	Talairach Coordinates X	Y	Z
Mild MCI (*n* = 58) CFT *p* < 0.001	Uncus (20)	R	0.019	1174	3.81	30	-11	-26
	Amygdala	R			3.73	30	-5	-20
	Uncus (20)	R			3.58	27	0	-34
	Middle Temporal Gyrus (21)	R			3.39	42	-4	-30
	Inferior Temporal Gyrus (20)	R			3.29	38	-12	-36
	Brain Region (BA)	**Side**	**Cluster-Level *pUNC***	**Cluster Extent (voxels)**	**Peak-level *z*-Score**	**Talairach Coordinates X**	**Y**	**Z**
Moderate MCI (n = 53) CFT *p* < 0.05	Middle Frontal Gyrus (9)	L	0.032	5946	4.05	-30	23	34
	Medial Frontal Gyrus (9)	L			3.03	-15	36	29
	Superior Frontal Gyrus (6)	L			2.96	-15	14	51
	Middle Frontal Gyrus (6)	L			2.77	-30	-12	42
	Angular Gyrus (39)	L			2.77	-39	-62	34
	Cingulate Gyrus (24)	L			2.65	-16	-2	46
	Inferior Parietal Lobule (40)	L			2.64	-48	-37	33
	Medial Frontal Gyrus (6)	L			2.61	-10	-15	58
	Postcentral Gyrus (2)	L			2.56	-33	-23	38
	Inferior Parietal Lobule (40)	L			2.53	-34	-39	37
	Middle Temporal Gyrus (21)	L	0.120	2877	3.46	-44	-3	-18
	Inferior Temporal Gyrus (21)	L			3.12	-56	-16	-14
	Middle Frontal Gyrus (6)	R	0.350	1001	2.71	33	8	44
	Middle Frontal Gyrus (9)	R			2.64	33	19	35
	Temporal Sub-Gyral (20)	R	0.301	1222	2.65	38	-19	-21
	Brain Region (BA)	**Side**	**Cluster-Level *pFWE***	**Cluster Extent (voxels)**	**Peak-level *z*-Score**	**Talairach Coordinates X**	**Y**	**Z**
AD Dementia (n = 71) CFT *p* < 0.001	Cuneus (17)	L	<.001	22718	4.60	-22	-84	12
	Cuneus (18)	R			4.57	27	-77	17
	Lingual Gyrus (18)	L			4.52	-20	-78	-6
	Lingual Gyrus (19)	R			4.46	26	-58	0
	Precuneus (18)	L			4.41	-20	-75	23
	Middle Occipital Gyrus (19)	R			4.39	44	-83	8
	Superior Parietal Lobule (7)	R			4.36	20	-61	56
	Parahippocampal Gyrus (19)	R			4.27	36	-44	-1
	Thalamus-Pulvinar	L			4.21	-18	-33	5
	Middle Occipital Gyrus (19)	L			4.18	-40	-79	2
	Thalamus-Ventral Anterior Nucleus	L			4.17	-15	-8	14
	Lingual Gyrus (18)	R			4.13	15	-73	-1
	Lingual Gyrus (19)	L			4.13	-20	-60	-2
	Superior Parietal Lobule (7)	R			4.12	32	-63	53
	Middle Occipital Gyrus (19)	L			4.11	-40	-84	4
	Thalamus-Pulvinar	R			4.04	21	-33	3

Areas of significant negative correlation between grey matter volume and verbal fluency discrepancy scores in each patient group are reported. Covariates: Age, Education, MMSE,

Total Intracranial Volume and CNR. Unc: Uncorrected; BA: Brodmann's Area. In the moderate MCI group, uncorrected set-level thresholded $p < .05$ results and coordinates with a *z*-score at local maximum > 2.5 are shown.

Figure captions

Figure 1. MCI Classification in Mild and Moderate Subgroups.

Bar chart showing the number of MCI patients assigned to each of the MCI sub-types according to neuropsychological test score cut-offs defined as 1.5 *SD* below the mean score of age, education, and gender matched controls. aMCI-md: amnestic MCI multi-domain; aMCI-sd: amnestic MCI single-domain; naMCI-md: non-amnestic MCI multi-domain; naMCI-sd: non-amnestic MCI single-domain.

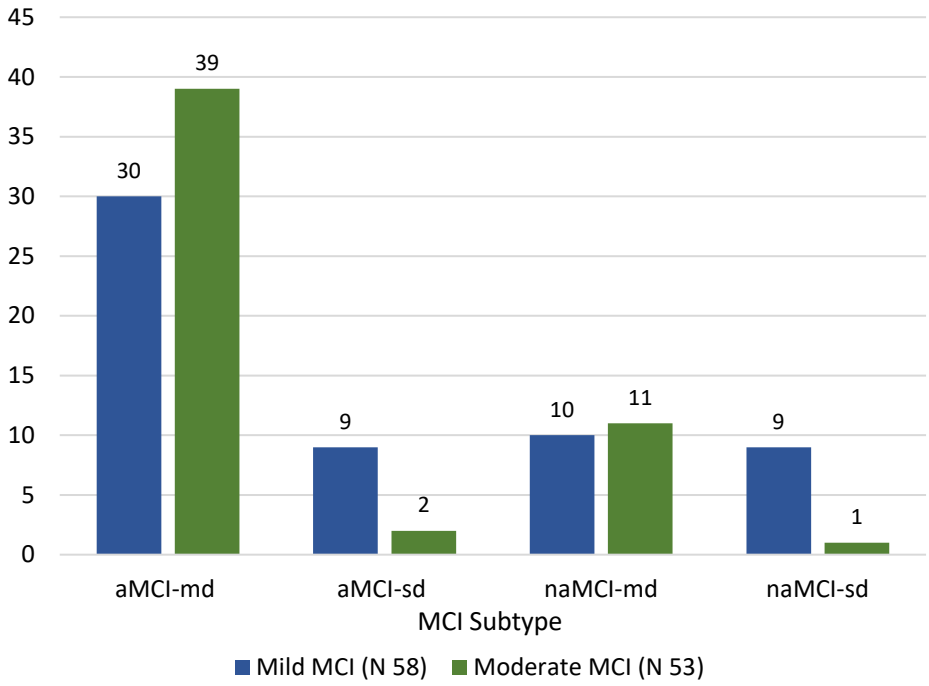
Figure 2. Verbal Fluency z-Scores and Discrepancy Scores

Box plots showing the *median* and range of verbal fluency *z*-scores and discrepancy scores in each patient group. Boxes depict the *median* and *interquartile range* and error bars represent the range. Significant differences are highlighted. [a] Significantly lower than letter fluency; * Significantly lower than mild MCI; ** Significantly lower than mild and moderate MCI; *** Significantly higher than mild and moderate MCI but did not survive correction for multiple comparisons.

Figure 3. Areas of Significant Negative Correlation Between Grey Matter Volumes and Verbal Fluency Discrepancy Scores.

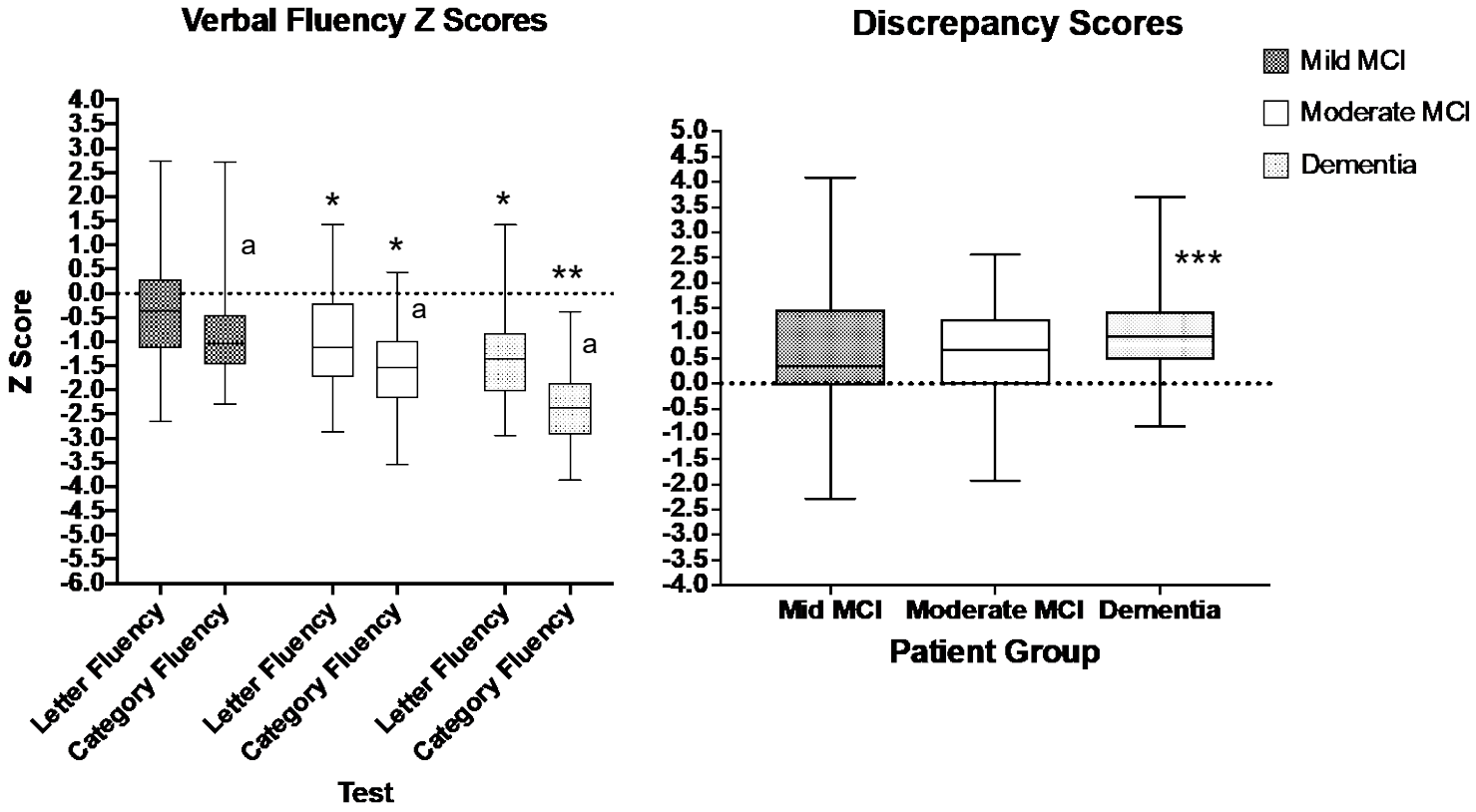
Figure showing the areas of significant correlation between grey matter volumes and verbal fluency discrepancy score in each patient group. Coordinates correspond to MNI space. Results are thresholded at a $p < .001$ in mild MCI and dementia groups and at a $p < .05$ in the moderate MCI group.

Figure 1. MCI Classification in Mild and Moderate Subgroups



Bar chart showing the number of MCI patients assigned to each of the MCI sub-types according to neuropsychological test score cut-offs defined as 1.5 *SD* below the mean score of age, education, and gender matched controls. aMCI-md: amnestic MCI multi-domain; aMCI-sd: amnestic MCI single-domain; naMCI-md: non-amnestic MCI multi-domain; naMCI-sd: non-amnestic MCI single-domain.

Figure 2. Verbal Fluency z-Scores and Discrepancy Scores



Box plots showing the *median* and range of verbal fluency *z*-scores and discrepancy scores in each patient group. Boxes depict the *median* and *interquartile range* and error bars represent the range. Significant differences are highlighted. [a] Significantly lower than letter fluency; * Significantly lower than mild MCI; ** Significantly lower than mild and moderate MCI; *** Significantly higher than mild and moderate MCI but did not survive correction for multiple comparisons.

Figure 3. Areas of Significant Negative Correlation Between Grey Matter Volumes and Verbal Fluency Discrepancy Scores.

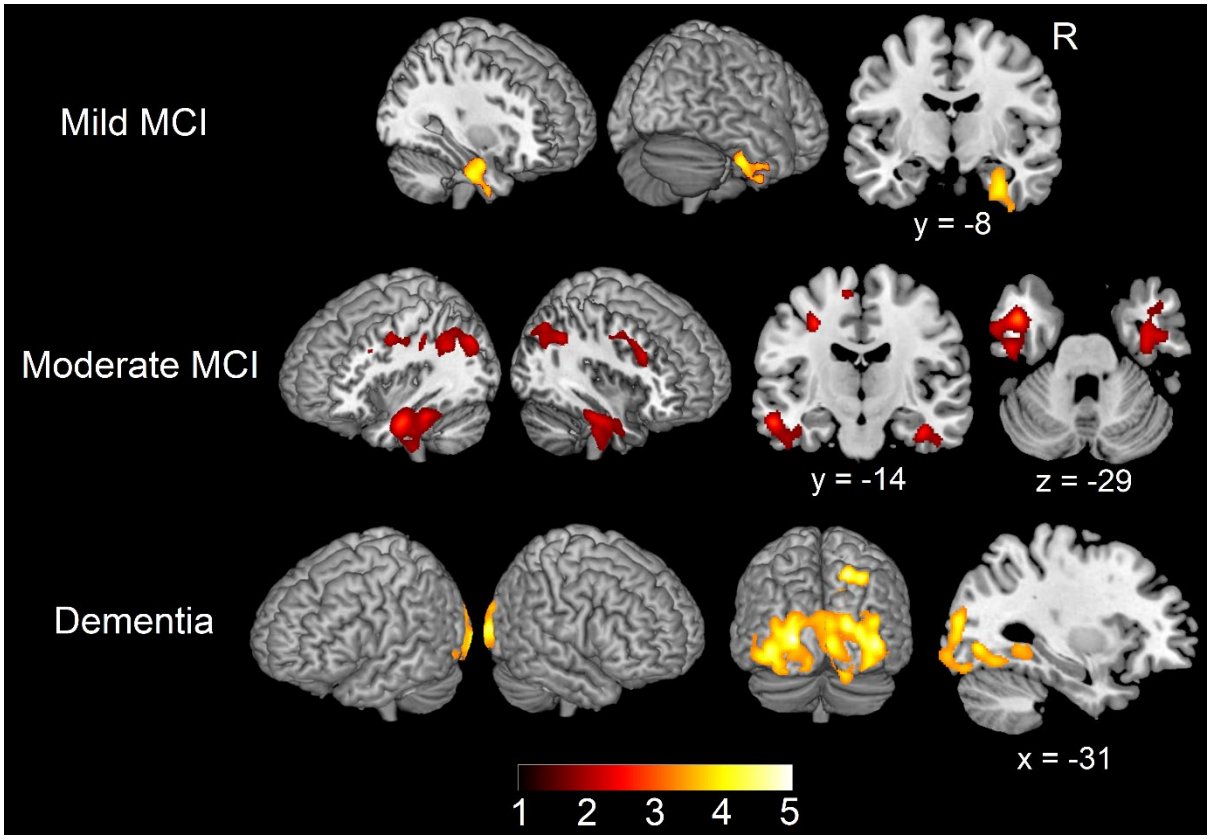


Figure showing the areas of significant correlation between grey matter volumes and verbal fluency discrepancy score in each patient group. Coordinates correspond to MNI space. Results are thresholded at a $p < .001$ in mild MCI and dementia groups and at a $p < .05$ in the moderate MCI group.

Supplementary material

Supplementary Table 1. *Table containing the results of a Kruskal-Wallis H test to determine which neuropsychological tests scores differed significantly between healthy controls, MCI and probable AD dementia groups. Post-hoc Dunn tests with a Bonferroni correction were applied. Significant differences between group pairs are highlighted (corrected p < .05)*

Test		Controls	Mild MCI	Moderate MCI	Dementia	Kruskal-Wallis H	df	p value
Letter Fluency (Raw Scores)	**n**	82	58	53	71			
	Median (IQR)	37.50 (20.00)	31.00 (16.50)	23.00 (14.50)[ab]	20.00 (16.00)[ab]			
	Mean Rank	177.58	143.54	106.63	90.73	57.18	3	<.001
Letter Fluency Z Scores	**n**	82	58	53	71			
	Median (IQR)	0.06 (1.53)	-0.38 (1.45)	-1.13 (1.53)[ab]	-1.35 (1.22)[ab]			
	Mean Rank	173.09	160.45	107.44	81.50	28.25	3	<.001
Category Fluency (Raw Scores)	**n**	82	58	53	71			
	Median (IQR)	41.00 (18.00)	29.50 (10.25)[a]	26.00 (14.50)[a]	19.00 (13.00)[abc]			
	Mean Rank	199.02	136.56	110.72	68.61	116.53	3	<.001
Category Fluency (Z Scores)	**n**	82	58	53	71			
	Median (IQR)	-0.12 (1.37)	-1.03 (1.04)[a]	-1.54 (1.20)[ab]	-2.37 (1.09)[abc]			
	Mean Rank	199.73	152.57	107.29	57.28	66.36	3	<.001
Discrepancy Scores	**n**	82	58	53	71			
	Median (IQR)	0.14 (1.27)	0.35 (1.51)[a]	0.67 (1.31)[a]	0.94 (0.97)[a]			
	Mean Rank	98.30	134.98	138.08	165.80	6.47	3	<.001
Raven	**n**	82	58	53	67			
	Median (IQR)	30.50 (7.25)	27.00 (7.25)[a]	24.00 (8.00)[a]	19.00 (11.00)[a]			
	Mean Rank	180.60	144.33	111.67	72.11	82.28	3	<.001
Digit Cancellation	**n**	82	58	53	71			
	Median (IQR)	53.00 (9.00)	49.00 (12.00)[a]	42.00 (10.00)[ab]	35.00 (19.00)[abc]			
	Mean Rank	186.36	149.97	114.97	69.11	95.67	3	<.001
Similarities	**n**	82	58	53	67			
	Median (IQR)	22.00 (9.00)	19.00 (5.25)[a]	16.00 (6.00)[ab]	12.00 (9.00)[abc]			
	Mean Rank	180.57	143.72	113.22	71.46	82.46	3	<.001
Token Test	**n**	81	57	53	68			
	Median (IQR)	35.00 (3.00)	33.00 (3.00)	32.00 (2.75)[ab]	30.00 (5.00)[abc]			
	Mean Rank	180.71	151.58	112.22	65.37	96.46	3	<.001
Rey-Osterrieth Complex Figure - Copy	**n**	82	58	52	67			
	Median (IQR)	32.50 (4.00)	31.25 (7.00)	28.00 (8.63)[ab]	21.50 (15.50)[abc]			
	Mean Rank	174.29	150.20	113.02	71.49	76.69	3	<.001
Rey-Osterrieth Complex Figure - Recall	**n**	82	58	52	67			
	Median (IQR)	14.75 (7.13)	7.25 (8.38)[a]	6.75 (5.38)[a]	3.00 (6.50)[abc]			
	Mean Rank	193.59	128.30	109.76	69.36	106.96	3	<.001
Stroop Test - Time Interference	**n**	80	58	52	65			
	Median (IQR)	24.00 (16.75)[cd]	31.50 (25.63)	34.50 (32.63)	40.00 (47.75)			
	Mean Rank	99.73	124.29	147.56	150.46	21.59	3	<.001
Stroop Test - Error Interference	**n**	80	58	52	65			
	Median (IQR)	0.00 (1.00)[cd]	0.50 (4.00)[cd]	2.50 (7.50)	7.00 (15.50)			
	Mean Rank	81.79	111.58	152.38	180.02	77.06	3	<.001
Digit Span Forward	**n**	82	58	53	71			
	Median (IQR)	6.00 (2.00)	6.00 (1.00)	5.00 (1.00)[a]	5.00 (1.00)[ab]			
	Mean Rank	160.70	149.65	121.37	94.23	36.10	3	<.001
Digit Span Backward	**n**	82	58	53	71			
	Median (IQR)	4.00 (1.00)	4.00 (2.00)[a]	4.00 (1.00)[a]	3.00 (2.00)[ab]			
	Mean Rank	174.80	140.66	119.46	86.70	57.48	3	<.001
Prose Memory Test - Immediate Recall	**n**	82	58	53	70			
	Median (IQR)	11.00 (6.00)	8.00 (4.25)[a]	7.00 (5.00)[a]	3.00 (5.00)[abc]			
	Mean Rank	191.80	139.12	119.09	65.81	106.20	3	<.001
Prose Memory Test - Delayed Recall	**n**	82	58	53	70			
	Median (IQR)	15.00 (8.25)	8.00 (7.00)[a]	7.00 (8.50)[a]	2.00 (5.00)[abc]			
	Mean Rank	199.76	128.22	121.63	63.60	123.22	3	<.001
Verbal Paired Associates Learning Test	**n**	82	58	52	68			
	Median (IQR)	13.25 (5.50)	9.25 (3.63)[a]	8.25 (4.75)[a]	4.50 (3.00)[abc]			
	Mean Rank	192.93	136.56	121.41	56.99	122.81	3	<.001
Confrontation Naming Test	**n**	82	57	53	67			
	Median (IQR)	19.00 (2.00)	19.00 (3.00)	18.00 (2.00)	16.00 (6.00)[abc]			
	Mean Rank	159.64	142.56	125.58	86.54	38.78	3	<.001

[a] Significantly lower than controls, [b] Significantly lower than mild MCI, [c] Significantly lower than Moderate MCI, [d] Significantly lower than dementia. Differences determined by independent-samples t-tests are in bold.

Supplementary Figure 1

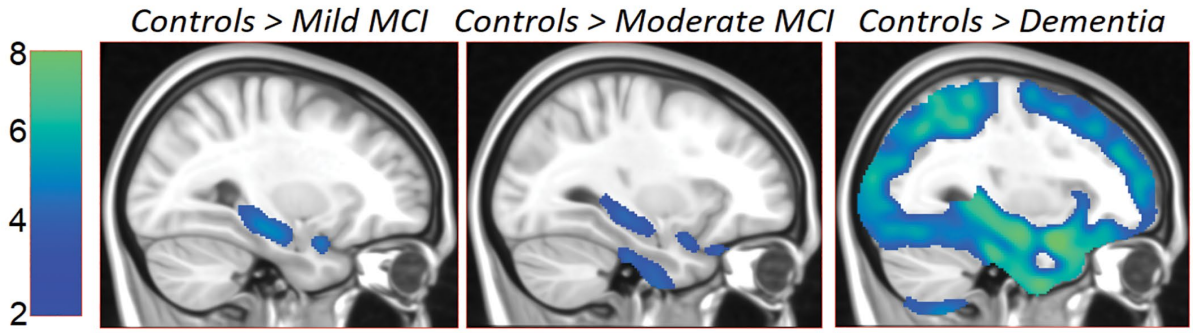


Figure 1. *Areas of reduced grey matter in each patient group compared with healthy controls emerging from voxel-brain morphometry analyses*