Effects of Alzheimer’s Disease on the Recognition of Novel Versus Familiar Words: Neuropsychological and Clinico-Metabolic Data

Lekeu Françoise(1)(2)(3) Van der Linden Martial(3)(4) Degueldre Christian(1), Lemaire Christian(1), Franck Georges(2)Luxen André(3) Moonen Gustave(1)(2) Salmon Eric(1)(2)

(1)Cyclotron Research Centre, University of Liège, Liège, Belgium
(2)Department of Neurology, University Hospital of Liège, Liège, Belgium
(3)Neuropsychology Unit, University of Liège, Liège, Belgium
(4)Cognitive Psychopathology Unit, University of Geneva, Geneva, Switzerland

This study explored recognition memory performance for novel versus familiar words in Alzheimer’s disease (AD) patients and normal controls (NCs), using an adaptation of E. Tulving and N. Kroll’s (1995) procedure. Results showed that both groups exhibited more hits and more false alarms for familiar than for novel words. The groups did not differ in the recognition of familiar words, reflecting preserved familiarity processes in AD. However, AD patients made more false alarms than NCs in the recognition of novel words, reflecting impairment of recollection processes in AD. A positron emission tomography analysis of clinico-metabolic correlations in AD patients showed a correlation between recognition of novel words and right hippocampal activity, whereas recognition of familiar words was more related to metabolic activity in the left posterior orbitofrontal cortex.

It is well established that Alzheimer’s disease (AD) is associated with important deficits in episodic memory (Grober & Kawas, 1997), which are generally present very early in the course of the disease (Morris, 1996; Spinnler & Della Sala, 1988). Episodic memory refers to the memory system that makes it possible to consciously recollect both personal happenings and events from one’s personal past and to make mental projections of anticipated events into one’s subjective future (Wheeler, Stuss, & Tulving, 1997).

Memory impairment in AD affects free recall (Grober, Lipton, Hall, & Crystal, 2000; Helkala, Lauluma, Soin-inen, & Riekkinen, 1988), cued recall (Buschke, Sliwinski, Kuslansky, & Lipton, 1995, 1997; Ergis, Van der Linden, & Deweer, 1994; Petersen, Smith, Ivnik, Kokmen, & Tangalos, 1994; Tounsi et al., 1999), and recognition (Braconier-nier, Cole, Spera, & De Vitt, 1982; Wilson, Fox, Kramer, & Kasznia, 1983). These deficits have been attributed to poor encoding (Corkin, 1982; Karlsson et al., 1989; Kasznia, Poon, & Riege, 1986; Weingartner et al., 1981) and retrieval difficulties (Bird & Luszcz, 1991; Cushman, Conon, Booth, & Caine, 1988; Ergis et al., 1994; Tuokko & Crockett, 1989; Tuokko, Vernon-Wilkinson, Weir, & Beattie, 1991). Studies exploring storage in AD have led to opposing results. Some studies suggest that the forgetting rate is normal in AD (Becker, Boller, Saxton, & McGonigle-Gibson, 1987; Christensen, Kopelman, Stanhope, Lorentz, & Owen, 1998; Corkin et al., 1984; Kopelman, 1985; Money, Kirk, & McNaughton, 1992), whereas other studies show an accelerated forgetting rate in AD patients, compared with normal controls (Tröster, Jacobs, Butters, Cullum, & Salmon, 1989; Welsh, Butters, Hughes, Mohs, & Heyman, 1991).

In terms of recognition memory, the dual component hypothesis distinguishes between the involvement of two different processes called recollection and familiarity (Mand-ler, 1980). Recollection refers to the conscious retrieval of events and associated contextual features, which is voluntary and effortful, whereas familiarity corresponds to a more automatic and effortless process. In studies exploring these recognition processes in AD, there is general agreement that recollection processes are impaired, whereas familiarity processes are relatively preserved (Balota, Burgess, Cortese, & Adams, 2002; Balota et al., 1999; Budson, Daffner, Desikan, & Schacter, 2000; Dalla Barba, 1997; Watson, Balota, & Sergent-Marshall, 2001). For example, Dalla Barba (1997) contrasted the “know” and “remember” responses of AD patients and normal controls (NCs) in free recognition and forced-choice recognition memory for words and faces. The results showed that correctly recognized items were less often associated with remember responses (reflecting conscious recollection) in AD patients than in NCs in both recognition conditions, independently of the material used. However, there was no difference
between both groups in terms of know responses (reflecting familiarity-based recognition) in both recognition conditions, apart from the fact that AD patients gave more know responses to target faces in forced-choice recognition than controls. This dissociation in recognition memory performance in AD patients was in keeping with studies demonstrating that AD affects the controlled, but not the automatic, retrieval processes (Fabrigoule et al., 1998; Jorm, 1986; Knight, 1998; Koivisto, Portin, Seinela, & Rinne, 1998).

Recently, Tulving and Kroll (1995) developed a memory paradigm that follows a three-phase design, designed to contrast the recognition of novel words with recognition of words made familiar through repeated presentation of items. The first phase of the procedure was devoted to expose subjects to words, during which subjects were asked to make living–nonliving judgments. The subjects were shown each word of this set six times. The goal of this repetition was to familiarize subjects with the words. In the second phase, subjects made the same living–nonliving judgments on a study list, in which half of the words had been presented before (familiar words), and the other half consisted of novel words (words not previously presented in the experiment). During the third phase, subjects were given a recognition test and were asked to respond “yes” only for words that had appeared in the study list.

This paradigm was initially created to test the novelty-encoding hypothesis, which postulates that the efficacy of encoding in long-term memory depends on the novelty of information within an experiment. It is hypothesized that information, which is largely preexposed within the experiment before its inclusion into a study list, will be less well-remembered, compared with information that appears in the study list for the first time within the experiment. In agreement with this hypothesis, Tulving and Kroll (1995) demonstrated that recognition accuracy (hits minus false alarms) for novel words was higher than recognition for familiar words, suggesting that novel words were better remembered. This pattern of results has also been found in other studies using this type of design with larger samples of young subjects (Dobbins, Kroll, Yonelinas, & Liu, 1998; Greene, 1999). Concerning the specific effects of familiarity manipulation, Tulving and Kroll (1995) reported higher hits and lower false-alarm rates for novel words, compared with familiar words. This pattern was referred to as a mirror effect (Glanzer & Adams, 1985). However, this pattern was not found in subsequent studies (Dobbins et al., 1998; Greene, 1999; Maddox & Estes, 1997). Indeed, these later studies showed that manipulation of familiarity affected hits and false-alarm rates in the same direction (higher hits and higher false alarms for familiar words than for novel words), referred to as a concordant pattern. The discrepancy between Tulving and Kroll’s (1995) results and those of the subsequent studies was suggested to be due to the small number of subjects in the first sample (four cases), compared with subsequent studies (Greene, 1999).

Moreover, some authors (Dobbins et al., 1998; Greene, 1999; Maddox & Estes, 1997) have questioned the novelty hypothesis as formulated in Tulving and Kroll’s (1995) paradigm. It has been suggested that the superior recognition performance for novel items could be related to the self-paced presentation of the study-list items, which gives subjects the possibility to rehearse novel stimuli more than the familiar stimuli, resulting in better recognition (Maddox & Estes, 1997). However, this hypothesis has been refuted in an experiment demonstrating better recognition for novel items than for the familiar items while the presentation rate was controlled (Greene, 1999).

Dobbins et al. (1998) conducted a detailed analysis of the recognition processes involved in the Tulving and Kroll paradigm (1995), leading them to reject the novelty-encoding hypothesis. They demonstrated that the superior recognition accuracy (hits minus false alarms) for novel, compared with familiar items, was not related to the level of hits but rather reflected the very high level of false alarms for familiar distractor items, compared with novel distractor items. This may be because the distractor and target familiar items were not mutually exclusive, because half of the items in the study list consisted of items that had already been exposed before study. Consequently, consciously recollecting that an item was present in the preexposition phase would not reduce the uncertainty concerning the potential occurrence of this item in the study list. Moreover, because the same semantic judgment task was performed in both cases (living–nonliving judgments for the processing of both preexposed stimuli and study-list items), the recollection of this semantic information could not be used as a contextual cue to distinguish the source of the list. On the basis of these considerations, Dobbins, Kroll, Yonelinas, and Liu (1998) suggested that recognition of the preexposed material in the Tulving and Kroll paradigm would mainly be based on familiarity processes and little, if at all, on recollection. This suggestion was well developed in an experiment conducted by the same authors (Dobbins et al., 1998). They showed a better discrimination of familiar words not included in the study list (correct rejections of nonstudied familiar words) when the nature of the judgment task varied across the preexposition
phase and the study-list condition. The authors concluded that this manipulation allowed the recollection processes to operate more efficiently.

Although criticized, Tulving and Kroll’s paradigm (1995) is nonetheless useful when exploring familiarity and recollection processes in very contrasted recognition conditions. In fact, the recognition of familiar words will mainly imply familiarity processes (Dobbins et al., 1998), while recollection processes will be more easily used in the recognition of novel words. Such a paradigm is particularly interesting to use in a population of AD patients, given that a dissociation of processes has already been observed in their memory performance, that is, a relative preservation of familiarity-automatic processes, contrasting with an impairment in the use of recollection processes (Balota et al., 2002; Balota et al., 1999; Budson, Daffner, Desikan, & Schacter, 2000; Dalla Barba, 1997; Watson et al., 2001).

From a cerebral metabolic viewpoint, some studies have explored the cerebral regions implied in processing novel versus familiar items. Using positron emission tomography (PET), Tulving, Markowitsch, Kapur, Habib, and Houle (1994) explored novelty encoding networks. Before the scan, the subjects were presented with 80 complex, colored geographic pictures that appeared twice during the encoding session. After a delay of 24 hr, subjects underwent the scan trials when viewing either “old” or “new” pictures. For the novelty detection, the results showed greater activation for new than for old pictures in regions located in the right “expanded” limbic system including the hippocampal formation, the parahippocampal gyrus, retrosplenial cortex, thalamus, subcallosal area, anterior and inferior cingulate cortex, putamen, and medial prefrontal cortex. Activation for familiarity processing was located in the right dorsolateral prefrontal cortex, the left and right cingulate region (extending to the supplementary motor cortex), and the left medial retrosplenial-parietal region. A subsequent PET study by Tulving, Markowitsch, Craik, Habib, and Houle (1996) replicated these results with colored pictures of people, scenes, and landscapes, which led the authors to suggest the existence of brain regions specialized in transmodal novelty assessment. In a similar manner, a recent functional magnetic resonance imaging (fMRI) study contrasted regions differentially implied in the listening of either recently learned words (familiar) or completely new words (Saykin et al., 1999). Familiar words activated the right prefrontal cortex, left parahippocampal gyrus, left medial parietal cortex, and right superior temporal gyrus. Also, in agreement with previous studies, novel words activated the hippocampal region. The involvement of the hip-pocampal region in novelty detection has been demonstrated in other neuroimaging studies using PET (Martin, Wiggs, & Weisberg, 1997) as well as fMRI (Kanwisher, Chun, McDermott, & Ledden, 1996; Stern et al., 1996; Strange, Fletcher, Henson, Friston, & Dolan, 1999) and has been suggested in studies using event-related potentials (Grunwald, Lehnertz, Heinze, Helmstaedter, & Elger, 1998; Knight, 1996; Knight & Nakada, 1998). These results have led Eichenbaum (1999) to refer to “the shock of the new” for hippocampal involvement.

In the case of AD, there have been very few neuroimaging studies exploring familiarity and recollection memory processes. Only one study recorded event-related potentials (ERPs) in AD patients and NCs while executing a test of verbal recognition memory accompanied by contextual information retrieval (color associated to items; Tendolkar et al., 1999). The results showed that the item recognition rate was above chance in Alzheimer patients, owing to preserved familiarity processes. In this case, hits were accompanied by maximum ERP amplitude at frontal sites. However, contrary to NCs, recollection of the presentation color was below chance in AD patients, and this disability of recollection processes was shown to be related to hippocampal atrophy measured by MRI volumetry.

On the basis of these considerations, our study used an adaptation of Tulving and Kroll’s paradigm (1995) in AD patients and NCs. The first goal of the present study was to explore recollection and familiarity processes in each group of subjects. Previous studies have postulated an impairment of recollection processes and a relative preservation of familiarity processes in AD patients. Our hypothesis was that, if the recognition of familiar words is essentially explained in terms of familiarity processes, both groups should not differ in their recognition of familiar items. However, whereas NCs use recollection processes in novel word recognition, AD patients should manifest an impaired recognition performance for novel word recognition.

Given the very few neuroimaging studies exploring familiarity-recollection processes in AD, a second objective of this study was to use PET clinico-metabolic correlations in AD patients. The aim was to emphasize specific regions related to performance in the recognition of novel versus familiar words. On the basis of the suggestion that recollection processes would depend on hippocampal functioning, we anticipated a greater involvement of the hippocampus in the recognition performance for novel, compared with familiar words.
Method

Subjects

Sixteen patients (7 men and 9 women) who met National Institute of Neurological and Communications Disorders and Stroke and Alzheimer’s Disease and Related Disorders Association (NINCDS-ADRDA) criteria for probable AD underwent both a three-phase design memory task and resting (18F) fluorodeoxy-glucose PET (18FDG-PET). They all suffered from dementia with progressive memory impairment (McKhann et al., 1984). They were all recruited from the Cyclotron Research Centre in Liege, Belgium. The diagnosis was based on a general medical and neurological examination, complete neuropsychological testing, and clinical follow-up. At most, computerized tomography scans showed mild atrophy. Mean age was 70.1 (SD = 6.2) years, and mean (estimated) disease duration was 25.8 (SD = 15.1) months. Mean score on the Mini-Mental State Examination (MMSE; Folstein, Folstein, & McHugh, 1975) was 22.2 (SD = 4.2). Neuropsychological results were compared with results obtained in a group of 16 healthy, elderly NCs (12 men and 4 women). NCs were recruited by word of mouth and were tested either at home or at the Cyclotron Research Centre. Screening of NCs included a neuropsychiatric interview and medical examination. Exclusion criteria included actual or past history of depression, neurological disorder, alcohol or drug abuse, recent hospitalization, memory complaints, or medication that could influence memory performance (e.g., anticholinergic medication or benzodiazepine). Informed consent was obtained from all subjects. Mean age of normal controls was 65.4 (SD = 7.1) years. Age showed a tendency to be statistically different between AD patients and NCs, t(30) = 2.01, p = .053, so that age was taken as a confounding covariate for analyses.

Materials and Procedure

An adaptation of the procedure developed by Tulving and Kroll (1995) was used. This procedure permits a distinction between familiarity and novelty effects on recognition memory. The material consisted of a total of 160 concrete French nouns of two or three syllables that constituted the basic pool of experimental words. This pool was divided into four subsets: A, B, C, and D, each containing 40 words (15 living and 25 nonliving). There was no significant difference in degree of imagery and frequency of words between the four subsets (Baudot, 1992), for imagery, F(3, 156) = 1.46, p = .227, and, for frequency, F(3, 156) = .07, p = .978; the frequency ranged from 1 to 42 in each subset of words.

The procedure consisted of an initial familiarization phase (1) during which two sets of items (80 total: Sets A and B for Version 1 and Sets A and D for Version 2) were presented four times. The words appeared on a screen one at a time and at a subject-determined rate. Subjects had to make living–nonliving decisions for the words. When subjects had seen all the items, they were given a yes-no recognition test of the 80 words without any distractors. This test was followed by an additional trial of living-nonliving decisions on the 80 familiar words. Subjects saw these two sets of words a total of six times. Following this familiarization phase, subjects were submitted to a distractor task, which consisted of counting backward for 2 min. In the study phase (2), subjects were presented with 80 words, half of which had appeared in the familiarization phase, and the other half of which were novel (Sets B and C for Version 1 and Sets D and B for Version 2). Subjects were instructed to memorize the words, and they were again required to make living-nonliving judgments for the words. The rhythm of the presentation of words was self-paced, and judgment time was recorded. The learning phase was followed by a distractor task, which consisted of counting backward for 2 min. In the recognition test (3), subjects were shown all 160 words from the pool (Sets A, B, C, D), one at a time, at a self-paced rate, and for each word, they were asked to indicate whether the word came from the study list (yes-no recognition).

To avoid any list effects on memory performance, we used two versions of the test. Seven AD patients received Version 1 of the memory test, and the 9 others received Version 2. For the NCs, half of them received Version 1 of the test, whereas the others received Version 2. There were no significant differences in the ability to make correct living-nonliving judgments between groups (M_NC = .94, SD = .04; M_AD = .94, SD = .02), t(30) = - .60, p = .551.
PET Data

All images of cerebral glucose uptake were obtained in the resting condition on a SIEMENS CTI 951 R 16/31 tomograph (CTI, Knoxville, TN) in two-dimensional mode with collimating septa extended, using the ¹⁸Ffluorodeoxyglucose (¹⁸FDG) technique. The camera had a 10.8-cm field of view in the axial direction. A 20-min transmission scan was acquired for attenuation correction using three ⁶⁸Ge rotating sources prior to the tracer injection. An 8 mCi intravenous bolus injection of ¹⁸FDG was followed 35 min later by 20-min acquisition of the emission data under standard conditions (eyes closed in dimmed ambient light and ears unplugged). This scan was then reconstructed using a Hanning filter at a cutoff of 0.5 cycles per pixel, giving a transaxial resolution of 8.7 mm full width at half-maximum (FWHM) for each of the 31 planes.

All calculations and image transformations were carried out on a SUN SPARC Ultra 60 workstation (Sun Computers Europe Inc., Surrey, United Kingdom). For each PET scan, the 31 transverse planes were interpolated to 44 planes to render the voxels cubic (2.347 x 2.347 x 2.347). All subsequent steps of image processing were made using the statistical parametric mapping software (SPM99, Wellcome Department of Cognitive Neurology, London, England) implemented in Matlab 5.3 (Math Works, Natick, MA). First, PET images were manually reoriented to the interhemispheric fissure and intercommissural plane. Functional imaging data were then normalized into a standard stereotaxic anatomical space (Talairach & Tournoux, 1988), using a bilinear interpolation method to allow for intersubject averaging. The normalizing spatial transformation matches (in a least-squared sense) each scan to a template image that already conforms to the standard anatomical space. The estimation of spatial normalization parameters consists of two parts: affine registration and basis function. All images were checked visually before and after the normalization procedure to ensure that no cerebral region was incorrectly normalized in the stereotaxic anatomical space. All scans were finally smoothed using an isotropic Gaussian filter 12 mm FWHM (Signorini et al, 1999) to increase the signal-to-noise ratio, to account for differences in gyral anatomy in the data set, and to conform more closely to a Gaussian field model (statistical inferences made in the following imaging analyses rely on the theory of Gaussian fields).

Analysis of PET Data

Proportional scaling was used to achieve global normalization of voxel values between scans mainly to remove effects of differences in global cerebral metabolism between scans. To test hypotheses about regionally specific effects, we estimated the parameters in the general linear model (Frackowiak, Friston, Frith, Dolan, & Mazzotta, 1997) at each and every voxel included in the analysis. The statistical program (SPM99) generated a group-specific adjusted mean value and an associated adjusted error variance for each voxel. The resulting statistics are expressed by the t distribution. The t values constituted an SPM(t). Results were thresholded to a t statistic corresponding to \( p < .001 \) to assess statistical significance. Accounting for multiple nonindependent comparisons inherent in the analyses, a corrected significance value of \( p < .05 \) was used as the final threshold for significance. To suppress the effects of age, we introduced this measure as a confounding variable in all SPM99 design matrices.

The statistical analysis performed in SPM99 was a correlational analysis between resting brain metabolism and performance during the recognition of familiar and novel words in AD patients. When an a priori hypothesis could be drawn from the literature, regional activation was reported for a significance value of \( p < .005 \), uncorrected.

Analysis of Neuropsychological Data

Different measures were computed from the data, consisting of the proportion of novel and familiar items identified by subjects as study-list items (both hits and false alarms), a recognition accuracy score for both novel words and familiar words (\( P_{hit} \) minus \( P_{false alarms} \)), a total corrected score reflecting accuracy of total recognition (\( P_{hit} \) minus \( P_{false alarms} \)), \( d' \) measure (sensitivity of discrimination) with a higher score corresponding to better signal-to-noise detection and \( C \) (decision bias) score from the signal-detection theory (MacMillan & Creelman, 1991). All these measures were compared between both groups using the \( t \) test. Repeated measures analyses of variance (ANOVAs) were also conducted for comparisons between groups in terms of hits and false-alarm proportions for novel and for familiar items. The mean judgment time was recorded during the learning of familiar and novel items of the study list. This measure was compared within each group between familiar and novel words and between groups. The goal of the judgment time comparison was to explore whether the recognition performance was influenced by a differential processing (longer rehearsal) for novel versus familiar words during the learning of the study list, as suggested by some authors (e.g., Greene, 1999).
Results

Neuropsychological Results

Novel words

A 2 x 2 repeated measure ANOVA (Group: NCs vs. AD Patients X Novel Words: Hits vs. False Alarms) was conducted. Although there was no significant main effect of group, F(1, 30) = 1.08, p = .31, this analysis revealed a significant main effect of response accuracy, F(1, 30) = 204.86, p < .0001. Both groups made more hits than false alarms for novel words. There was also a significant interaction, F(1, 30) = 19.83, p < .0005. Planned comparisons detailed this interaction, showing no difference between groups in terms of hits, F(1, 30) = .78, p = .382, but a significantly higher proportion of false alarms in AD patients, F(1, 30) = 8.27, p < .01.

Familiar words

A 2 x 2 repeated measure ANOVA (Group: NCs vs. AD Patients x Familiar Words Hits vs. False Alarms) was also conducted. There was no main effect of group, F(1, 30) = 2.71, p = .109, but a significant main effect of response accuracy, F(1, 30) = 12.08, p < .005, appeared, showing that both groups made more hits than false alarms for familiar words. However, no interaction emerged, F(1, 30) = .95, p = .336.

The main results are summarized in Table 1. The recognition accuracy for novel, but not for familiar words, was statistically different between the two groups: for novel words, t(30) = -4.45, p < .0005, and, for familiar words, t(30) = .97, p = .336. Given the very low level of recognition accuracy for familiar words, the recognition accuracy score for each group was compared with chance performance. The level of chance performance was estimated in terms of probabilities of making either hits (40/160 = .25) or false alarms (40/160 = .25) on familiar items.

Using a one-sample t test (one-tailed), we compared the mean accuracy score compared with the estimated chance level (Pr__<sub>hits</sub> − Pr__<sub>false alarms</sub> = 0.00). These analyses revealed that the performance of both groups was superior to chance: in the elderly group, t(15) = 1.83, p < .05, and in the AD group, t(15) = 3.04, p < .005.

Contrary to the results of Tulving and Kroll (1995), but in accordance with those of Dobbins et al. (1998), Greene (1999), and Maddox and Estes (1997), the pattern of performance of both groups showed that hits and false alarms were higher for the familiar, compared with the novel, stimuli. The proportion of hits for studied familiar words was significantly greater than the proportion of hits for studied novel words: in the elderly group, t(15) = -4.41, p < .001, and in the AD group, t(15) = -7.05, p < .00001. The proportion of false alarms for nonstudied familiar words was also significantly greater than for nonstudied novel words: in the elderly group, t(15) = -27.93, p < .00001, and in the AD group, t(15) = 6.09, p < .00005 (see Table 1).

To explore whether better recognition of novel words, compared with familiar words, was due to an expanded judgment time for novel items, compared with familiar items, we conducted a 2 x 2 repeated measure ANOVA (Group: NCs vs. AD Patients x Judgment Time: Novel vs. Familiar Items) on the mean of median judgment time. There was a main effect of group, F(1, 30) = 16.90, p < .0005, but no main effect of judgment time, F(1, 30) = .06, p = .81. However, there was a significant interaction effect, F(1, 30) = 5.98, p < .05. Planned comparisons showed that there was no difference in judgment time between novel and familiar items in AD patients (p > .10), although in NCs, a difference approached the level of significance, F(1, 30) = 3.62, p = .067, with a greater judgment time for the familiar, compared with novel, items.

The mean total corrected score reflecting accuracy of total recognition (Pr__<sub>hits</sub> − Pr__<sub>false alarms</sub> for both novel and familiar word recognition) was significantly greater in NCs than in AD patients (M<sub>NC</sub> = .28, SD = .07; M<sub>AD</sub> = .17, SD = -.08), t(30) = -4.04, p < .0005. The d' measure was also better in NCs than in AD patients (M<sub>NC</sub> = .85, SD = .28; M<sub>AD</sub> = .46, SD = .25), t(30) = -4.08, p < .0005. Concerning the decision bias C, there was no significant difference between both groups (M<sub>NC</sub> = -2.00, SD = 3.63; M<sub>AD</sub> = -1.62, SD = 4.67), t(30) = .23, p = .819. The analysis in which age was included as a confounding covariable led to similar results.
Table 1 Proportion of Test Words Judged to Be Part of the Study List

<table>
<thead>
<tr>
<th>Words</th>
<th>Studied (hits)</th>
<th>Nonstudied (false alarms)</th>
<th>Recognition accuracy (hits minus false alarms)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M$</td>
<td>$SD$</td>
<td>$M$</td>
</tr>
<tr>
<td></td>
<td>$SD$</td>
<td>$SD$</td>
<td>$SD$</td>
</tr>
<tr>
<td>Novel</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NC</td>
<td>.67</td>
<td>.20</td>
<td>.15</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AD</td>
<td>.61</td>
<td>.18</td>
<td>.34</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Familiar</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NC</td>
<td>.89</td>
<td>.10</td>
<td>.86</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AD</td>
<td>.84</td>
<td>.09</td>
<td>.78</td>
</tr>
</tbody>
</table>

Note. NC = normal controls; AD = Alzheimer’s disease patients.

Correlations Between Brain Metabolism and Performances of AD Patients During the Verbal Recognition Memory Task

These analyses were carried out in the group of 16 AD patients and in a subgroup of 9 elderly subjects. They were designed to assess the involvement of some regions described in the literature concerning familiarity and novelty networks (Saykin et al., 1999; Tulving et al., 1996; Tulving, Markowitsch, et al., 1994). As frequently observed in the literature, no significant correlation was found in the elderly subjects, owing to the low level of metabolic variability in these subjects (Eustache, Desgranges, & Baron, 1995) or possibly to the small sample of elderly subjects. However, significant results were found in the sample of 16 AD patients (see Table 2), which persisted in two randomly selected subsets of 9 AD patients.

Concerning the recognition of novel words, activity in a cluster located in the right hippocampal area was significantly related to the correct rejection of nonstudied novel words. The conjunction between performance for both studied novel words (hits) and nonstudied novel words (correct rejections; with a masking procedure using both simple contrasts) also showed a positive correlation with metabolic activity in the right hippocampal formation (see Figure 1). The masking procedure ensured that the brain area was significantly correlated to both individual measures. The hippocampal region also appeared significantly when the recognition accuracy score (hits minus false alarms) for novel words was used as a variable of interest (all $p < .005$, noncorrected for multiple comparison but a priori hypothesized).

To isolate the neural substrate of recognition performance for familiar words, we contrasted a clinicometabolic correlation using performance for familiar words with performance for novel words (emphasizing brain regions where activity is related more to familiar, than to novel, word processing).

A conjunction analysis was performed on both studied and nonstudied familiar words using hits and correct rejections, respectively. The results highlighted a significant positive correlation with activity in the left posterior orbito-frontal region (BA 47/11; see Figure 2).
Table 2 Correlations Between Brain Metabolism and Recognition Performance for Novel Versus Familiar Words in Alzheimer’s Disease Patients

<table>
<thead>
<tr>
<th></th>
<th>Right hippocampal Formation coordinates</th>
<th>Familiar words</th>
</tr>
</thead>
<tbody>
<tr>
<td>Novel words</td>
<td>x: 26, y: -6, z: -26, Z score: 3.52</td>
<td>Conjunction of studied and nonstudied familiar words versus novel words correlations^c^</td>
</tr>
<tr>
<td>Nonstudied novel words (correct rejections)^a^</td>
<td>x: 24, y: -8, z: -24, Z score: 4.18</td>
<td>x: -30, y: 12, z: -18, Z score: 4.86</td>
</tr>
<tr>
<td>Conjunction between studied (hits) and nonstudied (correct rejections) novel words^b^</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recognition accuracy score for novel words^b^</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. In the study, age was taken as a confounding covariate for all analyses. Coordinates are in the stereotactical space of Talairach and Tournoux (1988). Recognition accuracy score for novel words = hits minus false alarms; BA = Broadmann’s area.

^a^Cluster level, p < .005, uncorrected. ^b^Voxel level, p < .005, uncorrected. ^c^Voxel level, p < .001, uncorrected.

Figure 1. Results from SPM99 (the statistical parametric mapping software) analysis. Significant positive correlation between novel words recognition score (conjunction of both hits and correct rejections) and regional mean consumption of glucose. Significant correlation with the right hippocampal formation is shown as white voxels projected on corresponding TI-weighted magnetic resonance imaging of 1 patient with Alzheimer’s disease.
Figure 2. Results from SPM99 (the statistical parametric mapping software) analysis. Recognition score for familiar words (compared with novel words) was positively correlated to a left posterior orbitofrontal region (Broadmann’s area 47/11). Significant correlation with this orbitofrontal region is shown as white voxels projected on corresponding T1-weighted magnetic resonance imaging of 1 patient with Alzheimer’s disease.

Discussion

The goal of this study was to explore recognition memory performance in AD patients and NCs for novel versus familiar words, using an adaptation of the Tulving and Kroll (1995) paradigm. For familiar words, recognition would essentially be based on familiarity processes. In line with this suggestion and in agreement with previous studies (Chalmers & Humphreys, 1998; Greene, 1999; Maddox & Estes, 1997), both groups exhibited a concordant pattern of recognition, that is, more hits and false alarms for familiar words, compared with novel words. In agreement with the hypothesis of preserved familiarity processes in AD, be-tween-group comparisons showed that both groups demonstrated a similar pattern of recognition for familiar items. However, AD patients were impaired in the recognition of novel words. More specifically, they tended to erroneously recognize more novel distractor items than NCs. Clinico-metabolic correlations in AD patients related the recognition performance for novel words to metabolic levels in the right hippocampal area, whereas recognition for familiar words was preferentially associated with activity in the left orbitofrontal gyrus.

Effects of Alzheimer’s Disease on the Recognition of Novel Words

First, both groups exhibited greater recognition accuracy for novel words, than for familiar words. However, in agreement with Greene (1999), this superiority for novel words was not due to a greater rehearsal of these words compared with familiar words. Indeed, although the presentation rate was self-paced, subjects did not spend more time on the processing of novel words. However, AD patients expended more time for the processing of both novel and familiar words than NCs. This could be explained by a general slowing of information processing in AD patients, as has been recently demonstrated for both lexical and nonlexical tasks (Myerson, Lawrence, Hale, Jenkins, & Chen, 1998). However, it should be noted that NCs spent more time in processing familiar, compared with novel,
items, whereas there was no difference in AD patients. We suggest that this observation could reflect a greater consciousness of NCs than AD patients about repetition of items. Given this, it is possible that NCs spent more time in processing familiar words because they attempted to distinguish them from previous familiar nonstudied items. This suggestion is in keeping with a recent study by Moulin, Perfect, and Jones (2000) showing that AD patients were not aware of the repetition of items during encoding, suggesting a difficulty in consciously controlling memory during verbal learning tasks.

Second, recognition accuracy for novel words was lower in the AD group than in the NC group. This pattern of performance appeared to be related to the higher level of false alarms for the nonstudied novel words by AD patients, compared with NCs (despite the lack of a significant differences for hits). High levels of false alarms have been frequently observed in recognition performances of AD patients (Balota et al., 1999, 2002; Budson et al., 2000). Because AD patients made the same level of hits as NCs in the recognition of novel words, an encoding impairment hypothesis cannot fully explain the higher level of false alarms in AD patients. However, this pattern of performance could be more clearly understood in terms of impaired recollection processes in AD (Balota et al., 1999, 2002; Budson et al., 2000; Dalla Barba, 1997).

Our suggestion is that NCs and AD patients use different processes for the recognition of novel words. Whereas NCs are able to efficiently use recollection processes, AD patients may only be able to use familiarity processes. Consequently, recognition solely based on familiarity processes would be sufficient to correctly recognize novel hits (given their high distinctiveness) but would not be sufficient to correctly reject novel distractors. In fact, recollection processes are needed to prevent familiarity processes to act on semantic attributes of the distractor items. These results are in agreement with those found in a study exploring false recognition memory in young and elderly healthy subjects alongside AD patients (Budson et al., 2000). The authors used an adaptation of the Deese and Roediger McDermott paradigm (Deese, 1959; Roediger & McDermott, 1995), which consisted of exposing subjects to lists of items in which each item converged to a semantically related, but nonpresented, critical word. The authors recorded hits and false recognitions after single and multiple exposures to word lists. They found that after a single list exposure, AD patients made significantly more false alarms to unrelated lures than older subjects, despite lack of statistical differences in the recognition of hits.

In the present study, data obtained from the clinico-metabolic correlations in AD patients showed a positive correlation between right hippocampal metabolism and the recognition accuracy score for the novel words. This is consistent with previous studies indicating that the hippocampus is associated with the conscious recollection of episodic memories (Eldridge, Knowlton, Furmanski, Book-heimer, & Engel, 2000; Mishkin, Vargha-Khadem, & Gadian, 1998; Reed & Squire, 1997). In addition, this observation could also account for the impairment in recollection processes in AD, given the well-described hippocampal pathology present from the very first stages of the disease (Fox, Warrington, & Freeborough, 1996; Koc'hler et al., 1998). Given these considerations, we suggest that recognition performance in AD patients for novel words is related to residual metabolic activity in the hippocampus, allowing recollection processes to operate.

Moreover, the involvement of the hippocampal formation is also in agreement with recent functional brain imaging studies suggesting that the hippocampus is part of a novelty-detection network for both verbal and nonverbal stimuli (see Tulving, Kapur, et al., 1994; Tulving, Markowitsch, et al., 1994, in which coordinates are very close to those found in our study). Our results pointed to the right hippocampal region. This finding is also consistent with previous studies that stressed preferential right hippocampal involvement in the novelty detection network, for both verbal and visual material (Martin, 1999; Martin et al., 1997; Tulving et al., 1996; Tulving, Markowitsch, et al., 1994). This observation supplies an additional factor, suggesting that the right side of hippocampal involvement is not dependent on the visual nature of the stimuli but could in fact reflect transmodal novelty assessment (Tulving et al., 1996). According to Martin (1999), the major role of the right hemisphere network in novelty detection is to maintain tonic attention and arousal.

Moreover, our results could also be interpreted in the light of the retrieval success hypothesis concerning the hippocampal formation. Retrieval success is related to the processes associated to ephory, that is, to the fact that a memory trace becomes available to consciousness (Rugg & Wilding, 2000). The behavioral data agree with this assumption, given that they showed that retrieval success was much greater for novel word recognition than for familiar word recognition. Consequently, our results suggest that the successful retrieval of novel words is related (in part) to hippocampal functioning. This observation is in line with some neuroimaging findings of hippocampal involvement in retrieval success (Heckers et al., 1998; Nyberg, McIntosh, Houle, Nilsson, & Tulving, 1996; Rugg,
Effects of Alzheimer’s Disease on the Recognition of Familiar Words

Both AD patients and NCs demonstrated the same level of recognition accuracy (hits minus false alarms) for familiar words. This lack of difference between both groups may be because all subjects were obliged to use familiarity judgments (Dobbins et al., 1998), which have been shown to be preserved in AD (Balota et al., 1999; Badson et al., 2000; Dalla Barba, 1997; Fabrigoule et al., 1998; Jorm, 1986; Knight, 1998; Koivisto et al., 1998; Watson, Balota, & Sergent-Marshall, 2001).

The level of recognition accuracy for familiar words was close to floor level in both groups of subjects but was superior to chance performance. This low memory performance is voluntarily induced by the Tulving and Kroll (1995) paradigm. In terms of cognitive processes, this procedure allowed us to tap mainly familiarity processes, although we cannot exclude that other recognition processes were implicated. According to Dobbins et al. (1998), however, this weak discrimination index seems to reflect recognition in which recollection processes did not have the possibility to operate efficiently.

The clinico-metabolic correlations in AD patients revealed that recognition for familiar words was related to metabolic activity in the left posterior orbitofrontal region. An extensive pathology of the orbitofrontal cortex (OFC) has been reported in Alzheimer’s disease. This pathology has been described in terms of neurofibrillary tangles, with posterior and medial areas of this region being the most severely affected (Van Hoesen, Parviz, & Chu, 2000).

Several neuroimaging studies have reported frontal activation in familiarity networks (Saykin et al., 1999; Tulving et al., 1996; Tulving, Markowitsch, et al., 1994), but none of them have described an orbital site of activation. However, this discrepancy is not entirely unexpected if we consider the differences between the experimental paradigm used in our study and those used in previous studies. In our study, the orbitofrontal region was related to an accuracy measure of recognition memory for familiar studied items. However, in previous studies (Saykin et al., 1999; Tulving, Marko-witsch, et al., 1994; Tulving et al., 1996), frontal activation was found when subjects were required only to look at visual or auditory material, made familiar either by repetition or by prior learning, without any constraint on memory retrieval. These methodological differences across studies are likely to involve different cognitive processes and have consequently led to different results.

Relatively few memory neuroimaging studies have described activation of this orbitofrontal region, apart from an activation PET study by Schneider, Treyer, and Buck (2000). Their experimental design consisted of repeatedly presenting the same set of pictures to subjects, albeit appearing in a different order each time. Subjects were required to detect recurrences within each run. The first run was a measure of new learning and activated medial temporal structures, mainly on the right side. During the subsequent runs, subjects had to discriminate picture repetitions within the current run, from picture presentations in previous runs. In these runs, all pictures were already familiar, and responses based only on familiarity no longer remained correct. Results revealed that this situation of correct recognition activated the posterior medial OFC. In agreement with previous studies demonstrating that spontaneous confabulators (showing mainly orbitofrontal and basal forebrain damage) were impaired on this test (Schneider & Ptak, 1999), the authors concluded that this orbitofrontal region may be important for the selection of currently relevant memories.

This interpretation could also be applied to our results because, as in the Schnider et al. (2000) study, stimuli were first made familiar through repetition, and second, stimuli had to be recognized according to their correct temporal attribution (“when” they appeared or not) or currently correct relevance. The young volunteers of Schneider et al.’s study were able to make this judgment without difficulty, because their levels of performance were near the ceiling effect. However, the experimental paradigm used in our study induced a relatively low level of recognition accuracy for familiar words in AD patients as well as in NCs. Because the OFC seems to be implied in two different memory paradigms, inducing opposite levels of recognition accuracy, we suggest that the involvement of the orbitofrontal region may not be related to the excessive difficulty or easiness of the cognitive task.

Schneider et al. (2000) hypothesized that OFC is more likely related to the suppression of the interference by items seen in the preceding runs, than to the selection of repeated items within runs. Our data showed that the orbitofrontal region was related to performances for both the correct recognition of studied familiar items (selection process) and for the correct rejection of nonstudied familiar items (suppression). This latter observation has led us to suggest that the posterior (OFC) is not solely implicated in the suppression of currently irrelevant items but is also implicated in correctly selecting currently relevant items, which could be seen as a more general monitoring function.
Except in memory studies, an increased activation of the left orbitofrontal region has been found by manipulating the number of alternative outcomes in guessing tasks (Elliott, Rees, & Dolan, 1999). However, it is unlikely that the involvement of the OFC in our study was simply a reflection of guessing, given that its involvement was highlighted in relation to correct recognition performance, which was shown to be greater than chance. Consequently, we suggest that the role of the OFC may be understood more clearly in relation to the “making choice” processes implied in this difficult situation of recognition, when dealing with uncertainty in more complex situations.

More generally, in the absence of significant correlation in elderly subjects, one could ask whether the cerebral areas highlighted in the correlational analysis of AD patients reflect an abnormal processing. Nevertheless, we showed previously that the correlations observed in AD patients were very compatible with neuroimaging data reported in normal subjects. Consequently, it seems plausible to consider, in agreement with studies using the cognitivo-meta-bolic correlational method, that AD patients used similar cerebral regions as those recruited by normal subjects during memory tasks (Desgranges et al., 1998; Eustache, Des-granges, Giffard, de la Sayette, & Baron, 2001; Perani et al., 1993). However, other neuroimaging studies using an activation paradigm are needed to explore the cerebral regions differentially implicated in memory tasks in normal subjects and AD patients. These kinds of studies will help to better understand the functional connectivity in AD (Grady, Furey, Pietrini, Horwitz, & Rapoport, 2001).

Acknowledgments

This work was supported by the Interuniversity Pole of Attraction (IAP) Programme P5/04, Belgian State, Prime Minister’s Office; Federal Office for Scientific, Technical and Cultural Affairs; the Belgian National Fund for Scientific Research; the Research Fund of Liège University; and the Fondation Medical Reine Elisabeth. Françoise Lekeu was supported by the IAP. We would like to thank Maria Binasco and Sylvie Willems for their assistance in the selection and testing of subjects. We also thank Frank Laroi for his generous help in the revision of the English version of this article as well as David A. Balota for very helpful comments on an earlier version of this article.

References


