Stem-completion priming in Alzheimer's disease: The importance of target word articulation

JOHN JOSEPH DOWNES, ERIC J. DAVIS, PAUL DE MORNAY DAVIES, TIMOTHY J. PERFECT, KEN WILSON, ANDREW R. MAYES and H. J. SAGAR

1Department of Psychology, University of Liverpool, P.O. Box 147, Liverpool L69 3BX, U.K.; 2Department of Clinical Psychology, 15 Hyde Terrace, University of Leeds, Leeds LS2 9LT, U.K.; 3Department of Psychology, Royal Holloway, University of London, Egham, Surrey TW20 0EX, U.K.; 4Department of Psychology, University of Bristol, 8 Woodland Road, Bristol BS8 1TN, U.K.; 5Department of Psychiatry, University of Liverpool, P.O. Box 147, Liverpool L69 3BX, U.K.; and 6Department of Clinical Neurology, Royal Hallamshire Hospital, Glossop Road, Sheffield S10 2JF, U.K.

(Received 25 August 1994; accepted 1 May 1995)

Abstract—Stem-completion priming performance in patients with Alzheimer's type dementia (DAT) was explored in three experiments in which both the standard repetition priming effect and a novel indirect form of priming, cohort priming, were measured. In the first experiment, in which study stimuli were words, both priming effects were found to be markedly attenuated in the DAT group. In the second experiment, the study stimuli were specially constructed nonwords, and it was found that cohort priming was present at normal levels in the DAT group. In a third experiment we tested the specific hypothesis that the requirement to overtly articulate target stimuli during the study phase was critical for the appearance of normal cohort priming in the DAT group in Experiment 2, and also for the normal levels of repetition priming which have been reported in some published studies. Two encoding conditions were compared, one in which subjects simply had to read aloud the target words and a second in which subjects were required to make evaluative (pleasantness) ratings for each of the target words (identical to that used in Experiment 1). Stem-completion priming performance following the latter condition was significantly attenuated in the DAT group relative to a healthy control group, but following the "read aloud" encoding condition, normal levels of repetition and cohort priming were observed. It is suggested that the most fruitful approach to understanding the performance of DAT subjects on lexical repetition priming tasks will involve a detailed analysis of language functions and how they interact with other, possibly mnemonic, processes in the generation of primed responses.

Key Words: stem-completion priming; Alzheimer's disease; target words.

Introduction

Stem-completion priming refers to the bias in completing initial letter stems (usually the first three letters) with words that have previously been studied. In the case of organic amnesia, there have been numerous demonstrations showing that stem-completion priming for familiar information is entirely normal in such patients [e.g. 16]. For patients with dementia of the Alzheimer type (DAT), however, the evidence seems to suggest an impairment in stem-completion priming [4, 5, 15, 20-22, 29, 31, 39, 40], although the effect is often not completely absent, and, importantly, not all studies have reported an impairment [e.g. 6, 7, 19, 28]. In the present experiments we examine the stem-completion priming performance of DAT patients in greater detail by focusing both on the standard repetition type of priming, which has been the dependent variable of all previous studies, and an indirect form of priming we have previously designated cohort priming [11].

The background to the study of repetition priming in both normal and pathological groups is such that it is now commonly interpreted as a memory phenomenon. Intact repetition priming in a memory-disordered group is therefore taken as prima facie evidence that certain underlying (memory) structures/processes are normal. Considering the evidence on different forms of implicit memory performance in neurodegenerative diseases, a multiple memory systems perspective is strongly supported. Thus, studies which have compared DAT with Parkinson’s disease and Huntington’s disease [e.g. 20] have found evidence of a double dissociation, with the latter two groups impaired on skill-learning or
procedural memory tasks such as pursuit rotor learning, and normal on stem-completion priming, whereas DAT patients exhibit normal learning curves on the pursuit rotor task but impaired performance on stem-completion priming. This strongly suggests that implicit memory performance is subserved by at least two distinct neuroanatomical systems: one, involving the basal ganglia and cerebellum, is related to action, and the other, involving temporal lobe structures, is concerned with lexical-semantic events.

Another view of priming, which applies particularly to lexical stimuli, posits that repetition effects are due to the functional properties of the relevant processing system. Rozin [32, see also 9] was probably the first to suggest this, and used the analogy of "hot tubes" to explain how a word's mental representation remains in an activated state after exposure to its orthographic form. A model which incorporates some of the features of the activation approach has been developed by Schacter [36], who proposed that the priming phenomena which have so far been described can be logically classified into two types: those that rely on some modification to existing (semantic) representations and their associations, and those that rely on activation in what he refers to as the perceptual representation system (PRS). The subsystems of the PRS are assumed to be pre-semantic and modular, and are thus able to deal with several classes of information, including visual word forms, structural descriptions, and auditory word forms [36]. However, a major departure from standard activation models is that the PRS is assumed capable of storing new, highly specific representations (although such specificity may not apply equally to all perceptual attributes [37]). The PRS might therefore be thought of as a memory system, although as stressed by Schacter, it does not store some of the types of information (e.g. semantic, contextual) that are assumed to contribute to the phenomenology of recollective experience.

The subsystems of Schacter's PRS model dealing with different types of lexical information, viz. the visual word form and auditory word form subsystems, are similar to the representation systems which have been postulated from research on the cognitive neuropsychology of language (see Ellis and Young [13] for a detailed review of this work). In previous work using this model of language processing [11], we successfully predicted a novel sublexical component to stem completion priming, designated as cohort priming. This cohort priming effect refers to the increased likelihood that for non-target completions, there is overlap with studied words in terms of the initial stem phonology. For example, if a subject has studied the word BLAME, evidence for the standard (direct) priming effect takes the form of an increased likelihood of responding BLAME to the stem BLA-, relative to an unstudied baseline condition. Evidence for this indirect priming effect takes the form of an increased likelihood of responding with (nontarget) words such as BLATENT or BLADE, as opposed to BLANK or BLACK. Cohort priming was also found to permeate across lexical boundaries, in that studied nonwords produced a similar bias in (word) stem-completion priming. The model accounting for these effects is considered in more detail in the Discussion. Below, we first describe two experiments examining cohort priming in a group of DAT patients and matched healthy control subjects. The main aim of these experiments is to determine if this newly identified component of stem-completion priming, like the standard repetition measure, is abnormal in DAT, and whether the observed pattern can illuminate the underlying deficit(s).

**Experiments 1 and 2**

**Subjects**

The DAT group consisted of 7 patients drawn from the Bradford and Sheffield areas. All patients had undergone a thorough neurological examination, and were of mild to moderate severity, being tested within one year of their initial diagnosis. CT scans revealed generalized cerebral atrophy for all patients, and EEGs revealed a marked reduction in alpha rhythm. All subjects scored below 22 on the Mini-Mental State (MMS [14]; range 16–21). The assessment battery included the Verbal subtests of the Wechsler Adult Intelligence Scale (WAIS [42]), the National Adult Reading Test (NART [26]), and Warrington's Recognition Memory Tests (WRMT, [41]). Details of scores on these various measures are shown in Table 1.

An equal-sized group of 7 control subjects were selected from a panel at the University of Bradford, and individually matched on age and NART IQ to each of the DAT patients. Selection criteria included a Mini-Mental State score of greater than 28, and a medical history free from alcoholism, neurological illness, and head injury. They were also selected to be from approximately the same socio-economic back-

<table>
<thead>
<tr>
<th>Group</th>
<th>Age</th>
<th>VIQ</th>
<th>NART-V</th>
<th>WRMT-w</th>
<th>WRMT-f</th>
<th>MMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>DAT mean</td>
<td>67.56</td>
<td>120.43</td>
<td>118.00</td>
<td>29.43</td>
<td>24.86</td>
<td>17.86</td>
</tr>
<tr>
<td></td>
<td>(9.21)</td>
<td>(12.84)</td>
<td>(13.96)</td>
<td>(8.31)</td>
<td>(4.43)</td>
<td>(2.92)</td>
</tr>
<tr>
<td>CONTROL mean</td>
<td>66.12</td>
<td>123.75</td>
<td>119.54</td>
<td>45.64</td>
<td>40.13</td>
<td>29.29</td>
</tr>
<tr>
<td></td>
<td>(4.23)</td>
<td>(13.38)</td>
<td>(12.97)</td>
<td>(5.10)</td>
<td>(4.82)</td>
<td>(0.31)</td>
</tr>
</tbody>
</table>

Notes: VIQ: Verbal IQ derived from the Wechsler Adult Intelligence Scale; NART-V: Verbal IQ estimated from the National Adult Reading Test; WRMT-w and WRMT-f: Warrington Recognition Memory Test scores, words and faces, respectively; MMS: Mini-Mental State score. Standard deviations of all measures are shown in parentheses.
ground as the patients. Control subjects were given the same psychometric battery as the patients. Table 1 gives the summary statistics for these groups. The DAT patients were matched to their control subjects in terms of mean age, 67.56 (9.21) years and 66.12 (4.23) years, respectively, $t(6) = 0.37$, $P > 0.05$; NART estimated premorbid Verbal IQ, 118.00 (13.96) and 119.54 (12.97), respectively, $t(6) = 0.78$, $P > 0.05$; and also WAIS verbal IQ, 120.43 (12.84) and 123.75 (13.38), respectively, $t(6) = 0.47$. There were, however, significant group differences on the word and face version of the WRMT: DAT faces 29.43 (8.31), control faces 40.13 (4.82), $t(6) = 2.5$, $P < 0.05$, DAT words 24.86 (5.10), control words 45.64 (5.10), $t(6) = 8.3$, $P < 0.05$. A comparison of NART estimated premorbid verbal IQ and current verbal IQ as measured by the WAIS seemed to demonstrate an absence of significant intellectual decline in the DAT group. Indeed, there was a small trend in the opposite direction: WAIS IQ 120.4 (12.8), NART IQ 118.6 (13.9). There are probably two reasons for this pattern: first, because the ceiling estimate from the NART is 128, subjects with IQs higher than this, and who are deteriorating intellectually, would not initially show any evidence of this in terms of the NART–WAIS discrepancy scores; second, the DAT patients were selected to be early-in-the-course, and it is well documented that verbal abilities hold well in the early stages of the disease [35].

Experiment 1

Methods

Materials. Experimental stimuli were a subset of those previously described [11], and this original selection was based on a pilot study involving 100 students. Briefly, target words for the earlier study were selected according to the following three criteria. First, several alternative completions for the initial three-letter stems were possible, and additionally, there were subsets with different initial stem phonologies; second, the likelihood of completing the stem with the selected (target) word was between 0.10 and 0.15 (this range covers the baseline completion rates for most published studies); third, the selected words each began with an uncommon stem phoneme. For example, the word BRILLIANT was selected as an experimental stimulus: the likelihood of completing the stem BR- with BRILLIANT was 0.12. The most common completion to this stem was BRIGHT, at 0.41. This completion was also a member of the most common initial phoneme subset, which included, for example, the words BRIDLE and BRINE. The selected word belonged to the less common initial phoneme subset containing the words BRITTLE and BRIM. Selection of words in this way allowed for two influences to be evaluated: a bias towards completing stems with the studied words and secondly, within the set of nontarget completions, a bias towards completing stems with words that share the initial phonemic structure as the target. From this original stimulus set of 64, 32 were randomly selected for the present experiment.

Stimuli were laser-printed in letters 1 cm high and fixed to flash cards (dimensions: 6 cm x 4 cm). A five-point numerical scale (1–5), which was used for a rating judgement (described in the Procedure), was printed below each word. In addition, separate cards were prepared on which the first three letters of the target stimuli and distractors were printed (in the same typography). For the recognition test, targets were paired with single distractors (matched on word frequency and word length), and these test pairs were printed eight to a page, again in the same typography. A separate booklet, in which stem completion responses were recorded by the experimenter, was also prepared.

Design. The 32 words were divided into two sets of 16, which were used equally often as target words in the experimental list, and distractors in the stem-completion test, in which case subjects never saw the stimuli in full. The same semantic encoding condition was used for all target stimuli and by all subjects. Similarly, stem-completion and recognition tests were identical for all subjects.

Procedure. All subjects were tested individually and were instructed that they were participating in a memory experiment, and that they should try to remember the words subsequently to be shown on flash cards. Five practice trials were then given for the encoding task. Subjects were instructed to rate each word according to how much it was “liked,” or the pleasantness of the word or any associations that came to mind. For this purpose a five-point scale, with 1 representing “dislike” and 5 representing “like”, was printed below each word. Subjects were instructed to take no longer than 5 sec for each judgement, and if necessary, were prompted for a response if this time limit was reached (which was necessary on only a small number of trials). Shortly after the rating task was completed, subjects were given the stem completion test, which was disguised as a filler task. It was stressed that guessing stem completions was the most appropriate tactic because the experimenter’s interest was in the first words that came to mind. Stems were shown on flash cards, and the experimenter wrote down subjects’ responses in specially prepared test booklets. The recognition test followed immediately, and for this subjects were given test booklets and told to underline the word from each pair which they remembered seeing previously.

Results

Two dependent variables, repetition and cohort stem completion rates, were derived in the following way for both target and distractor stimulus sets. The repetition priming measure was computed in the standard way, by dividing the number of target completions by the total number of completions (which in all cases was 16). The cohort priming measure was computed by dividing the number of nontarget completions, which overlapped with target words in terms of initial stem pronunciation, by the total of all nontarget completions. For example, suppose a subject completed target-derived stems with X target words and Y cohort words, leaving $(16 - X - Y)$ alternative (i.e. nontarget, noncohort) completions. The repetition priming measure would have a value of $X/16$, whereas the cohort priming measure would have a value of $Y/(16-X)$. Baseline repetition and cohort completion rates were derived in a similar way. Means and standard deviations for these measures on all experimental conditions are shown in Table 2.

To determine if repetition and cohort completion rates for the target stimulus set were enhanced to an equivalent degree for the two groups, these rates were compared in an ANOVA with the rates for the unstudied stimulus set. The ANOVA design was three-way split-plot factorial, with one between-subjects factor, group (DAT and control), and two within-subjects factors: stem-type (studied and unstudied), and
Table 2. Experiment 1 mean repetition and cohort priming completion rates for studied and unstudied words in DAT and matched control subjects

<table>
<thead>
<tr>
<th>Group</th>
<th>Studied Repetition</th>
<th>Studied Cohort</th>
<th>Unstudied Repetition</th>
<th>Unstudied Cohort</th>
</tr>
</thead>
<tbody>
<tr>
<td>DAT</td>
<td>0.17 (0.10)</td>
<td>0.43 (0.08)</td>
<td>0.15 (0.08)</td>
<td>0.39 (0.11)</td>
</tr>
<tr>
<td>Control</td>
<td>0.30 (0.08)</td>
<td>0.61 (0.13)</td>
<td>0.09 (0.09)</td>
<td>0.32 (0.08)</td>
</tr>
</tbody>
</table>

Notes: Standard deviations of all measures are shown in parentheses.

completion-type (repetition and cohort). Of the main effects, there was no significant overall difference between the groups, controls 0.32 (0.19) DAT 0.27 (0.15), \( F(1, 12) = 2.63, P = 0.13 \). As expected, there was a significant main effect of stem-type, that is, a significant priming effect, studied 0.35 (0.17) unstudied 0.24 (0.15), \( F(1, 12) = 15.95 \). Finally, the main effect of completion-type was also significant, repetition 0.18 (0.11) cohort 0.41 (0.14), \( F(1, 12) = 64.74, P < 0.01 \). In other words, completion rates for cohort words were far higher than completion rates for identical (target) words, which concurs with our previous results [11].

Of the interaction terms, the only one to emerge was the group by stem-type interaction, \( F(1, 12) = 12.74, P < 0.01 \) which can be seen clearly in Fig. 1. Post hoc tests (Tukeys) showed that overall priming, that is, the difference between targets and distractors, was significant in the control group, whereas the markedly attenuated priming effect for the DAT group was nonsignificant. In other words, overall priming, averaged across repetition and cohort types, was absent in the DAT group, and there was no additional evidence, in the form of a significant group by stem-type by completion-type interaction, that repetition or cohort priming was differentially affected. Neither of the two remaining interactions was significant: group by completion-type, \( F(1, 12) < 1, P > 0.05 \); stem-type by completing-type, \( F(1, 12) < 1, P > 0.05 \).

Recognition memory performance in the two groups was compared by a \( t \) test, which revealed that control subjects recognized a significantly higher proportion of the target words, 0.95 (0.09), than the DAT subjects, 0.62 (0.14), \( t(6) = 7.49, P < 0.01 \).

Discussion

The results of this experiment replicate the findings of several other studies [e.g. 37] in showing that DAT patients are impaired on stem-completion priming. The novel finding is that cohort priming is similarly affected. There was no evidence, in the form of a three-way interaction between group, stem-type, and completion-type, that the DAT group was any less impaired on cohort than on repetition priming. In our previous work, the sublexical nature of cohort priming was verified by showing that a similar effect was present even after studying specially constructed nonwords [11, Experiment 2]. In that experiment, subjects were exposed to nonwords which were assembled from the beginnings and endings of real words, but in such a way that a particular pronunciation of the initial three-letter stem was made highly probable. In a subsequent (word) stem-completion test, subjects' word solutions were much more likely to share the pronunciations given to the three-letter stems of the nonwords compared to a baseline condition. In Experiment 2 we were interested to determine if the global stem-completion priming deficit observed for the DAT group in Experiment 1 applied to all conditions which produce a cohort priming effect. Specifically, after studying nonwords, will the word stem-completions of DAT patients be subject to the same cohort bias shown by healthy control subjects?

Experiment 2

Methods

Materials. Thirty-two three-letter word stems, different from those used in Experiment 1, were chosen according to the criteria described above. These stems formed the first three
letters of the nonwords which were used in the present experiment and were constructed according to the following rules [see 11, Experiment 2]: 1. They did not rhyme with real words; 2. They were between 4 and 6 letters in length; 3. Each initial stem was unique in the set used; 4. The ending of the nonword was carefully chosen to ensure that the initial stem was pronounced in the desired manner. For example, to ensure that the stem BLA- was pronounced as in BLACK but not as in BLAME, the ending chosen was -LG, giving BLALG, rather than other potential endings such as -YG, giving BLAYG, which would have induced an inappropriate pronunciation. Stimuli were prepared as in Experiment 1, with a scale of 1–5 printed below each nonword to guide the rating task (which was different from Experiment 1, see Procedure below). For the test phase, flash cards with the initial stems and recognition test booklets were prepared.

Design. As with Experiment 1 nonwords were assigned to one of two sets, matched in terms of mean length, which were used equally often as targets and (the stems) as distractors in the stem-completion phase of the experiment, with additional balancing across groups. For the recognition test, a two-alternative forced choice procedure was used, with each target paired with a distractor of equal length.

Procedure. Subjects were introduced to the nonword stimuli of Experiment 2 by being shown a set of three examples. As with Experiment 1 the experiment was introduced as a memory test. It was explained that the stimuli were not real words and that the task was to read out the stimulus shown on each card and then give a rating on a dimension of “word-likeness”, with the object of remembering as many as possible. One of the practice stimuli was the pseudohomophone CLAME, and it was explained that this should be given a high rating because it looked and sounded like a real word. Another of the practice stimuli, GLOMP, was selected as an example of a stimulus which would receive a low rating. The actual instructions used varied across patients and depended on their initial level of comprehension. If subjects seemed unclear about the task, the instructions and examples were given again. This was judged successful in that all subjects read correctly all of the nonwords and appeared to use the scale appropriately during the experimental phase. After being shown all the nonword stimuli, subjects were given the stem-completion test in which they were shown the set of 32 stems on flash cards and encouraged to produce word completions as fast as they could. This was followed by the two-alternative forced choice recognition test in which subjects were asked to select the nonword from each pair which they remembered seeing in the earlier study phase. When unsure, subjects were encouraged to guess.

Results

For Experiment 2, only a single measure of cohort priming was derived, because subjects studied only nonwords, and, by definition, there could be no (word) repetition priming. Therefore, all word completions which overlapped with the pronunciations of the nonword stems were included as cohort completions. Data were analysed using a two-way split-plot factorial ANOVA, with group as a between-subjects factor and stem-type (studied and unstudied) as a within-subjects factor. This revealed a significant main effect of stem-type, studied 0.65 (0.11) unstudied 0.42 (0.08), F (1, 12) = 73.19, P<0.01, that is, significant priming. The main effect of group and the group by stem-type interaction were both nonsignificant. Figure 2 shows the relative priming effects for the control and DAT groups.

Recognition memory performance in the two groups was compared using a t test. This revealed a significant main effect of group, t (6) = 7.61, P<0.01, explained by the superior performance of the control group, 0.93 (0.07), relative to the DAT group, who performed at chance levels, 0.48 (0.12).

Discussion

In both Experiments 1 and 2, explicit recognition memory performance was impaired in the DAT subjects. In contrast, a measure of implicit memory performance, cohort priming, was normal in Experiment 2 when the study material was nonwords, yet was impaired in Experiment 1 when the study material was words, as was the standard measure of implicit memory, repetition priming. Caution is always necessary when interpreting the results of implicit memory experiments which appear to show deficits for a group which is known to be memory-impaired because, as several commentators have noted, the measures used can never be assumed to be “process-pure” [38]. In other words, the control subjects may be supplementing performance on the implicit memory test with explicit
memory processes. However, this is a most unlikely confound in the present study because cohort priming is an indirect form of priming, and, by definition, the relevant words contributing to the dependent variable were not studied, and, therefore, could not be "remembered".

As noted in the Introduction, there are several studies which have reported an impairment in stem-completion priming for DAT subjects, and the finding of Experiment 1, therefore, concurs with these. Although the pattern of cohort priming observed across the two experiments might, ostensibly, appear difficult to explain, as noted in the Introduction, there is a similar inconsistency in the literature concerning repetition priming in DAT. That is, in addition to findings of impaired stem-completion priming [4, 5, 15, 20–22, 29, 31, 33, 39, 40], a significant number of studies have also reported normal levels of this form of repetition priming in DAT [6, 7, 9, 10, 19, 23, 24, 28]. Also, a substantial number of other studies using different indices of (lexical) repetition priming [1, 3, 6, 18, 21, 30, 39] have shown normal levels of priming for familiar words in groups of DAT patients. It is unlikely that disease severity can explain either the results from the first two experiments or the pattern observed across previous studies because in the present study the same subjects participated in both experiments, and in a number of previous studies the same group of DAT subjects have been found to be abnormal on stem-completion priming but normal on some other measure of repetition priming, for example, anagram solution [39] or perceptual identification [21]. However, this does not rule out the possibility, if not likelihood, that the magnitude of different types of priming will decrease as disease progresses in DAT. It is possible that the small group sizes of the present study may have introduced a significant degree of unreliability for the dependent variables used, but we believe that the pattern observed here and across all previous studies can be explained by the same set of mechanisms.

There are several obvious differences between the stimuli and procedures used in Experiments 1 and 2 which might provide clues about the discrepancy in the performance of the DAT subjects across the two experiments. For example, the lexical status of target stimuli (real words in Experiment 1, nonwords in Experiment 2), access to semantics afforded by those stimuli, the judgements required the encoding phase (pleasantness rating in Experiment 1 and word-likeness rating in Experiment 2), and the fact that articulation of the stimuli was not required during encoding in Experiment 1 but was during Experiment 2. Deciding which of these factors, either alone or in combination, was responsible for the different outcomes would require a complex series of experiments. However, another rich source of evidence which can be consulted exists in already published studies. As noted above, not all of these have reported abnormal levels of priming in DAT, and in Table 3 published studies of stem-completion priming in DAT are categorized according to outcome, that is, normal or abnormal effects for the DAT group.

In all, 21 studies have been reported, 11 of which have reported a stem-completion priming deficit in DAT with the remainder reporting no group differences between DAT and healthy control subjects. The actual figures when separate experimental conditions are considered are 14 and 11, respectively. Also shown in Table 3 are the encoding tasks employed in each study. Although a wide range have been employed there is no immediately apparent difference between the two sets of studies when the more obvious methods of classifying encoding conditions are applied. For example, using a levels of processing framework, it can be seen that encoding tasks requiring either shallow or deep levels of processing are represented in both subsets. However, if we again consider the differences, listed earlier, between Experi-

<table>
<thead>
<tr>
<th>Study</th>
<th>Abnormal effects</th>
<th>Encoding task</th>
<th>Normal effects</th>
<th>Encoding task</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gabrieli et al. (1994)</td>
<td>Pleasantness rating</td>
<td></td>
<td>Christensen et al. (1992)</td>
<td>Sentence generation</td>
</tr>
<tr>
<td>Shimamura et al. (1987)</td>
<td>Pleasantness rating</td>
<td></td>
<td>Partridge et al. (1991)</td>
<td>1. read aloud with pleasantness rating</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Russo and Spinnler (1994)</td>
<td>2. read aloud with definition</td>
</tr>
</tbody>
</table>

Table 3. Published studies of stem-completion priming in Alzheimer's type dementia classified according to outcome and encoding task.
ments 1 and 2, one factor, the requirement to articulate or read aloud the target stimuli, correctly classifies almost 85% of the studies shown in Table 3. That is, in 12 of the 14 experimental conditions which have produced DAT related deficits in stem-completion priming, successful fulfilment of the encoding task did not require the subjects to overtly generate target word phonologies, whereas in 9 out of 11 experimental conditions in which the control-DAT comparison was nonsignificant, subjects were required to read aloud the target words, if only in partial fulfilment of the encoding task. In Experiment 3, therefore, we test the hypothesis that the level of stem-completion priming observed for DAT subjects can be modulated by varying the requirement to read aloud the target stimuli. Two encoding conditions are compared, the pleasantness rating task used in Experiment 1 (and several other studies which have reported a DAT related deficit, see Table 3), and a condition in which subjects were simply required to read aloud the target words. Although this hypothesis is not theoretically driven, in the General Discussion we offer some speculations about how the results obtained, and the more general pattern of findings, can be accommodated by the model of stem-completion priming recently developed by our group [11].

Experiment 3

Methods

Subjects. A total of 28 subjects, comprising equal numbers of healthy control and DAT subjects, took part in Experiment 3. DAT patients were attendees at a Psychogeriatric Day Hospital Service in the Liverpool region, and had undergone a thorough clinical examination to exclude other treatable causes of dementia. The control group was recruited from a Local Authority Social Services Day Centre. For both groups, exclusion criteria included history of stroke or other neurological illness, psychiatric illness, or head injury. CT scans for all DAT cases showed evidence of diffuse cerebral atrophy. All subjects were assessed using the Mini-Mental State, the Vocabulary subtest of the WAIS, and the NART to provide a measure of pre-morbid IQ (see Table 4). The two groups were matched on age, DAT 67.50 (4.03), control 67.79 (3.89), t (26) = 0.19, P > 0.05, and NART estimated pre-morbid Verbal IQ; DAT 98.57 (6.90), control 100.50 (7.61), t (26) = 0.70, P > 0.05. The DAT group were, however, significantly worse than the control group on the WAIS Vocabulary subtest (expressed as raw scores), DAT 24.86 (13.18), control 41.14 (9.14), t (26) = 3.80, P < 0.01, and on the MMS, DAT 18.36 (3.34), control 26.86 (2.83). All DAT subjects scored below, and all control subjects above 22 on the MMS.

Materials. Eighty words were selected (from a dictionary) which satisfied the following criteria: first, that there were at least 10 other words beginning with the same three letters; second, that the initial phonemic stems were not all pronounced in the same way; third, that the selected word was not a member of the most common cohort (i.e. stem pronunciation). Two lists of 40 words each, A and B, were formed, and each list was further divided into two subsets, giving A1, A2 and B1, B2, which were matched approximately for word frequency and word length. A further two sets of 20 words, again matched with the A and B subsets on word length and word frequency, were employed as foils on the recognition memory tests.

The A and B subsets were printed on flash cards in two ways—either as isolated words, or each word printed above a five-point scale. In addition, flash cards showing the initial three-letter stems of all the words were prepared. Finally, 4 versions of the 20 two-alternative forced-choice recognition test were printed, with 10 items to a page. All words and their corresponding stems and relevant recognition test items were printed in the same case, font, and point size.

Design. All subjects completed two versions of the memory test using the two lists A and B. Equal numbers of subjects in each group studied subsets A1, A2, B1, and B2, with the constraint that each subject studied one A and one B list. Also, both lists and study subsets were, as far as possible, balanced across the two encoding conditions, and order of presentation of the two versions was balanced across groups. The stem-completion test comprised stems derived from the studied words and from the matching (unstudied) subset; for example, if a subject studied subset A1, the stem-completion test comprised stems derived from A1 and A2 words. Orders of presentation for words in the study subsets, stems in the implicit memory test, and the two alternative test items of the recognition test were fixed across subjects.

Procedure. Subjects were seen on at least two occasions, the two test versions being run on separate occasions with at least 1 week delay separating them. The basic procedure was as follows. Subjects were first informed that the objective of the task was to evaluate certain aspects of their memory performance, but were also told that they would not be expected to "recall" the words, only to pick out the correct ones from two alternatives. The specific encoding instructions were then explained to the subject and some examples given. For the read-encoding orienting task they were simply required to read each word out aloud once and try to commit it to memory. For the like-encoding orienting task they were instructed on how to use the five-point scale, and were told to instruct on how to use the five-point scale, and were told

Table 4. Psychometric characteristics for subject groups who took part in Experiment 3

<table>
<thead>
<tr>
<th>Group</th>
<th>Age</th>
<th>Vocab</th>
<th>NART-V</th>
<th>MMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>DAT mean</td>
<td>67.50</td>
<td>24.86</td>
<td>98.57</td>
<td>18.36</td>
</tr>
<tr>
<td></td>
<td>(4.03)</td>
<td>(13.18)</td>
<td>(6.90)</td>
<td>(3.34)</td>
</tr>
<tr>
<td>CONTROL mean</td>
<td>67.79</td>
<td>41.14</td>
<td>100.50</td>
<td>26.86</td>
</tr>
<tr>
<td></td>
<td>(3.89)</td>
<td>(9.14)</td>
<td>(7.61)</td>
<td>(2.83)</td>
</tr>
</tbody>
</table>

Notes: Vocab: Vocabulary subtest raw score from the Wechsler Adult Intelligence Scale; NART-V: Verbal IQ estimated from the National Adult Reading Test; MMS: Mini-Mental State score. Standard deviations of all measures are shown in parentheses.
to indicate how much they "liked" each of the study words, ranging from 1 "not at all" to 5 "very much". Words were then shown at a rate of one every 4 sec. After being shown all 20 words, they were engaged in conversation for a couple of minutes before being given the implicit memory test. This was introduced as something to fill in the time before the memory test was started. Prior to being shown the test items, subjects were first given a couple of examples of stems. The flash cards on which the test items were printed were shown to the subjects up to a maximum of 10 sec. If a response was not given in that time, the subject was told not to worry and the experimenter moved on to the next item. Responses were recorded on specially prepared sheets. On completion of the implicit memory test, subjects then moved onto the explicit recognition memory test, and for this they were given the test booklets with instructions to indicate by underlining for each two-alternative test item, which of the two words they remembered seeing earlier.

Results

Two dependent variables, repetition and cohort stem-completion rates, computed separately for both studied and unstudied stimulus sets, were derived in the manner described in Experiment 1. All means and standard deviations are shown in Table 5.

To examine stem-completion priming performance, a four-way split-plot factorial ANOVA was used with group (DAT and control) as the first, between-subjects factor, and encoding task (read and like), completion-type (repetition and cohort), and stem-type (studied and unstudied) as the three within-subjects factors. Of the main effects, neither the group main effect: DAT 0.28 (0.20), control 0.28 (0.20), $F (1, 26)=0.01$, $P>0.05$, nor the encoding task main effect: read-encoding 0.29 (0.20), like-encoding 0.27 (0.21), $F (1, 26)=0.39$, $P>0.05$, was significant. Significant effects did emerge, however, for completion-type, $F (1, 26)=13.52$, $P<0.01$, and stem-type, $F (1, 26)=16.32$. The completion-type main effect indicates that, overall, the proportion of cohort completions, 0.32 (0.20), was higher than the proportion of target (repetition) completions, 0.24 (0.19), which concurs with the results of Experiment 1 and the results previously reported in our study using young healthy subjects [11]. The stem-type main effect indicates that the experiment was successful in producing significant priming effects, with completion rates for the studied stimuli, 0.34 (0.21), being reliably higher than completion rates for the unstudied stimuli, 0.22 (0.17). There was also a significant completion-type by stem-type interaction, $F (1, 26)=7.05$, $P=0.01$, which is explained by the larger priming effect for target completions, studied 0.33 (0.20) vs unstudied 0.14 (0.13), than for cohort completions, studied 0.35 (0.22) vs unstudied 0.29 (0.18).

Of most interest was the reliability of the interaction terms involving group and encoding task, since these related directly to the experimental hypothesis. The two-way group by encoding task interaction was nonsignificant, $F (1, 26)=1.38$, $P=0.25$, but the related three-way interaction involving stem-type was marginally nonsignificant, $F (1, 26)=3.57$, $P<0.07$. As Fig. 3 indicates, under like-encoding instructions, the DAT subjects show little evidence of priming, studied 0.27 (0.16), unstudied 0.25 (0.16), but this increases to normal levels under the read-encoding instructions: studied 0.39 (0.22), unstudied 0.22 (0.15). In contrast, the control subjects show equivalent levels of priming under both like-encoding instructions: studied 0.37 (0.21), unstudied 0.20 (0.12), and read-encoding instructions: studied 0.34 (0.20), unstudied 0.20 (0.11). The hypothesis that the study orienting task should influence the level of priming in DAT but not control subjects was stated a priori, so, despite the nonsignificant interaction term, priming in the two groups and for each encoding condition could be tested using planned comparisons (Tukey's HSD). This revealed significant and equivalent levels of priming under both encoding conditions for the control group, and significant priming for the DAT group with read-encoding, which was also at a level

<table>
<thead>
<tr>
<th>Group</th>
<th>Repetition</th>
<th>Cohort</th>
<th>Repetition</th>
<th>Cohort</th>
</tr>
</thead>
<tbody>
<tr>
<td>DAT</td>
<td>0.25</td>
<td>0.29</td>
<td>0.15</td>
<td>0.34</td>
</tr>
<tr>
<td></td>
<td>(0.14)</td>
<td>(0.18)</td>
<td>(0.17)</td>
<td>(0.26)</td>
</tr>
<tr>
<td>Control</td>
<td>0.36</td>
<td>0.38</td>
<td>0.12</td>
<td>0.28</td>
</tr>
<tr>
<td></td>
<td>(0.21)</td>
<td>(0.22)</td>
<td>(0.15)</td>
<td>(0.17)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Group</th>
<th>Repetition</th>
<th>Cohort</th>
<th>Repetition</th>
<th>Cohort</th>
</tr>
</thead>
<tbody>
<tr>
<td>DAT</td>
<td>0.35</td>
<td>0.42</td>
<td>0.14</td>
<td>0.30</td>
</tr>
<tr>
<td></td>
<td>(0.21)</td>
<td>(0.24)</td>
<td>(0.10)</td>
<td>(0.15)</td>
</tr>
<tr>
<td>Control</td>
<td>0.37</td>
<td>0.31</td>
<td>0.15</td>
<td>0.24</td>
</tr>
<tr>
<td></td>
<td>(0.21)</td>
<td>(0.25)</td>
<td>(0.12)</td>
<td>(0.10)</td>
</tr>
</tbody>
</table>
equivalent to that observed for the control group, but no evidence of priming for the DAT group under the like-encoding condition.

None of the other interaction terms was significant. The nonsignificance of the four-way interaction, group by encoding task by completion-type by stem-type, suggests that both repetition and cohort priming components were affected similarly. However, for the DAT group, the size of the repetition priming effect reduces from 0.21, which is significant, under read-encoding instructions to 0.10, a nonsignificant effect, under like-encoding instructions, whereas the values for the cohort priming effect are 0.12 and −0.05, respectively. In other words, although the significant reduction in the size of the priming effect is of the same order of magnitude for both components of priming, the cohort completion rate for studied stimuli under like-encoding instructions is actually less than that for unstudied stimuli.

Performance on the recognition tests was evaluated using proportion correct as the dependent variable. As with the previous experiments the DAT group was significantly impaired relative to the control group.

Discussion

The finding that stem-completion priming is significantly impaired in DAT following a pleasantness rating encoding task concurs with the results of Experiment 1 and several other published experiments which have used the same encoding task. The relevant cohort priming result of Experiment 3 also replicates that of Experiment 1, in that this type of priming was similarly attenuated in the DAT group following the pleasantness rating task. The reversal of the repetition and cohort priming deficits following an encoding task which required subjects to read aloud the target words during the study episode conforms our suspicions that this was at least a contributory factor in producing nonsignificant and significant levels of cohort priming for the DAT group in Experiments 1 and 2, respectively. More generally, it strongly supports the hypothesis that the same factor underlies the pattern of impaired and intact repetition priming for DAT subjects apparent in different subsets of published studies. Not all studies conform to this pattern, so the target articulation hypothesis cannot provide a complete explanation of the known variance in outcomes, but the fact that this single factor can lead to a correct classification of almost 85% of published studies is impressive, and means that whatever subtle language-related impairment underlies it must be incorporated into any model which attempts to explain the complex pattern of implicit memory performance which has been observed in DAT.

Most discussions of stem-completion priming in DAT highlight those studies which have reported deficits and tend to ignore those that have found normal levels of priming. In fact, if all studies which have examined repetition priming in DAT are considered together, that is, including those which have used other priming paradigms such as lexical decision and perceptual identification, than the total which have reported normal levels of priming outnumbers the subset of studies which have reported deficits. It is also noteworthy that the only repetition priming on which DAT
related deficits have emerged is the stem-completion task. Any complete theory of priming in DAT must, therefore, explain two related patterns. First, the finding that all priming tasks apart from stem-completion priming have consistently revealed normal performance levels in DAT subjects, and second, for stem-completion priming, not all studies have reported DAT related deficits.

Of the explanations which have been put forward to account for these patterns the most theoretically inspired approach is found in the work of Gabrieli and his colleagues [15, 21]. They have been particularly concerned with explaining the first pattern described above, that is, DAT related deficits on stem-completion priming alongside normal performance levels on other repetition priming tasks. In one of the few studies which have examined different forms of priming in a within subjects design, Gabrieli and colleagues [21] reported that their DAT subjects were impaired on stem-completion priming but showed normal levels of performance on a perceptual identification task. To explain this pattern, it was claimed that lexical priming tasks vary in the degree to which visual perceptual and lexical semantic representations underlie performance. Tasks like perceptual identification and lexical decision depend more on visuo-perceptual features of stimuli, whereas stem-completion is assumed to involve both visuo-perceptual and conceptual/semantic components. This classification corresponds well with what is known about the neural organization of these functions and the neuropathology of DAT, because a plausible locus of visuo-perceptual priming is the occipital cortex, particularly prestriate cortex, which is relatively spared in DAT, whereas lexical-semantic priming effects are likely to be subserved by temporal-parietal association areas which, of course, are significantly compromised in DAT [2]. The results of Experiments 1 and 2 would also seem to fit this scheme well because normal cohort priming was only found after the study of nonwords which, of course, would not involve access to semantics.

There are, however, two major problems with this interpretation. First, it depends critically on the assumption that stem-completion priming involves a conceptual component. Second, it fails to address those studies which have reported normal levels of stem-completion priming in DAT, to which can now be added the results of Experiment 3. The reason why it is assumed that stem-completion priming involves a conceptual component is that, under cross-modal study-test conditions, significant priming is found, although the level is attenuated relative to same modality study-test conditions [e.g. 16]. Levels of stem-completion priming in DAT are reduced, therefore, by an amount which reflects the conceptual component. However, if the DAT deficit is related to some difficulty accessing semantic information during study, which is a corollary to this hypothesis, then encoding tasks which encourage semantic processing should lead to near normal levels of stem-completion priming. But, as noted above, the main distinction between those studies that have found normal stem-completion priming in DAT and those which have reported deficits is the requirement to specify the phonology of the stimuli during study. Many of the studies which have found abnormal priming in DAT have used a “deep” levels of processing encoding task, whereas an equal number of studies which have found normal levels of priming have used a “shallow” levels of processing encoding task. And, most convincingly, the results of Experiment 3 show that the DAT-related deficit in stem-completion priming can be reversed using an encoding task which, in levels of processing terms, is not as deep as the encoding task which produced the deficit, rating the pleasantness of words.

In previous work [11] we have described a model of stem-completion priming based on existing models of the architecture of the language processing system and an analysis of the particular demands of the stem-completion task. According to this, subjects complete words in the stem-completion test by first generating an appropriate phonological interpretation of the stem and then using that as a basis for search and retrieval from a phonological word-form representation system. Because the initial three letters of many words are ambiguous with respect to phonology, the likelihood of selecting target (studied) words as completions is dependent on the prior specification of the appropriate stem phonology. Our work suggests that after studying target words, a bias is established in the procedure for translating letter segments into phonologies (the assembled reading route) such that given a stem from a studied word, the phonological translation of it is more likely to overlap with the pronunciation of the original word compared to baseline conditions when target words are unstudied. The stem-completion task is therefore viewed as a two-stage process involving translation of the stem into some phonological form and subsequent selection of an appropriate lexical entry from the phonological word-form representation system. Both stages can be independently biased, which leads to the emergence of cohort priming when the set of nontarget completions is analysed. Also, because the first stage, and, therefore, cohort priming, is sublexically based, then this can also result from the study of nonword stimuli in addition to words, as shown in Experiment 2 here and our previous study with healthy student subjects [11]. In normal subjects it is assumed that all routes to phonology (that is, the assembled reading route and the direct or addressed route) are automatically activated on presentation of word stimuli, which is why, even under silent study conditions, significant levels of cohort priming are observed.

The above also helps to explain why, for control subjects, the level of cohort priming (0.29) was greater than the level of repetition priming (0.21) in Experiment
1, whereas in Experiment 3 cohort priming (0.09) was less than repetition priming (0.23). Because there are two sources of activational influence, these can be independently biased, which means that there need not be any consistent relationship between levels of cohort and repetition priming across different word-sets. Also, there is some dependency between the measures used in that with the same cohort bias increasing the lexical bias has the effect of reducing the estimate of cohort priming. Thus, the cohort measure used provides a conservative estimate of cohort priming which is why, in some experiments conducted in our laboratory, the cohort priming effect has failed to reach significant levels (a lengthier discussion of these issues can be found in [11]).

According to this model, the reason why significant (but attenuated) cross-modal priming occurs is that auditory presentation results in the activation of phonological word-form representations. Levels of priming are higher under same modality study-test conditions because of the additional bias in the assembled reading route which can only result from visual presentation of words. Additional work from our laboratory [12] supporting this position has shown that cross-modal stem-completion priming can be eliminated by a secondary phonological task (articulatory suppression) during the test phase. Although, as stated above, the finding of normal cohort priming in DAT following nonword study could be taken as support of Gabrieli and colleagues' hypothesis, according to our model, the cohort priming effect following both word and nonword study is due to the same underlying mechanism, a biasing of procedures in the assembled reading route. That is, semantic or conceptual factors are not relevant for the normal production of this priming effect.

It is clear that existing theories have difficulty explaining the complete pattern of normal and impaired performance on stem-completion and other priming tasks. The ideas of Gabrieli and colleagues go some way to achieving this in that they can explain those findings of abnormal stem-completion priming and normal perceptual identification (and other forms of) priming. But they fail because, according to their hypothesis, stem-completion priming must always be abnormal because of damage to temporal lobe structures which are assumed to store lexical-semantic knowledge. Whilst there is plenty of additional evidence testifying to the breakdown of semantic knowledge structures in DAT [cf. 25], our own analysis suggests that stem-completion priming is a phonologically-driven, rather than a conceptually-driven task, and there are, in any case, a significant number of published studies which have reported normal levels of stem-completion priming in DAT and which must be accounted for.

The above considerations suggest that a more fruitful approach to understanding the pattern of preserved and impaired priming performance on stem-completion and other repetition priming tasks in DAT will involve a detailed analysis of language functions and how they interact with other, possibly mnemonic, processes in the generation of primed responses. At the very least, the results of the present study and our classification of previous studies implicate a subtle impairment in the procedures linking orthography and phonology. Stem-completion priming is dependent on adequate phonological specification of the target stimuli, which, in normals, occurs automatically regardless of task demands. For DAT subjects, however, when articulation is not explicitly required, activation of the relevant conversion procedures in the assembled reading route and/or activation of the phonological word-form representations in the phonological lexicon is weak, but the resultant deficit in repetition and cohort priming can be reversed by ensuring that stimuli are read aloud.

This hypothesis, although underspecified, suggests why stem-completion priming is sometimes present at normal levels and sometimes attenuated in DAT, but it does not explain why all other priming tasks have consistently shown DAT subjects to be unimpaired. It needs to be stressed at this point that our claim that priming is phonologically-driven applies only to stem-completion priming, and it does not follow that other priming tasks will be similar in this respect. Priming tasks such as word-fragment completion and perceptual identification are much less likely to involve a phonological selection route because the information provided by test stimuli, nonsequential letter strings, cannot be easily parsed into a phonological form. Lexical selection in these cases is, therefore, more likely to be orthographically-driven. Thus, our position differs from that of Gabrieli and colleagues because we are suggesting what marks stem-completion priming as different from other priming tasks is its reliance on derived phonology in the lexical selection process, rather than the involvement of lexical semantics. Both positions can explain the pattern of impaired performance on stem-completion priming and preserved performance on other priming tasks, but, as noted earlier, there are actually two patterns that require explanation. Postulating a deficit in lexical semantics does not explain why encoding tasks which ensure target word articulation generally produce normal levels of stem-completion priming, whereas those that do not generally lead to attenuated stem-completion priming in DAT.

Acknowledgements—Experiments 1 and 2 were submitted as partial fulfilment for the degree of Ph.D to the University of Bradford by Eric J. Davis. The financial support of the Research Fund of Professors B. Costall and R. Naylor of the Department of Pharmacology, University of Bradford, the Research Development Fund of the University of Liverpool, and the Sandoz Foundation for Gerontological Research, Belgium, is gratefully acknowledged.
References


