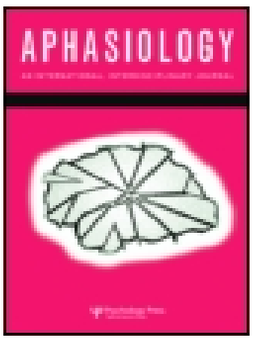

Aphasiology

ISSN: 0268-7038 (Print) 1464-5041 (Online) Journal homepage: http://www.tandfonline.com/loi/paph20

Performance on the Boston Cookie theft picture description task in patients with early dementia of the Alzheimer's type: Missing information

Elaine Giles , Karalyn Patterson & John R. Hodges

To cite this article: Elaine Giles , Karalyn Patterson & John R. Hodges (1996) Performance on the Boston Cookie theft picture description task in patients with early dementia of the Alzheimer's type: Missing information, Aphasiology, 10:4, 395-408, DOI: 10.1080/02687039608248419

To link to this article: https://doi.org/10.1080/02687039608248419

Performance on the Boston Cookie Theft picture description task in patients with early dementia of the Alzheimer's type: missing information

ELAINE GILES, KARALYN PATTERSON† and JOHN R. HODGES

Addenbrooke's Hospital, Cambridge, UK
†MRC Applied Psychology Unit, Cambridge, UK

(Received 9 November 1994; accepted 16 March 1995)

Abstract

We report on the spoken language of 48 patients with dementia of the Alzheimer's type (DAT) who were divided into three approximately equal subgroups on the basis of the Mini-Mental State Examination (minimal 24–29, mild 17–23, and moderate 3–16) and 18 matched controls. The Cookie Theft picture description task from the Boston Diagnostic Aphasia Examination was chosen because it is considered an ecologically valid approximation to spontaneous discourse. All subjects were also assessed on a battery of semantic memory and non-verbal tests. Our analysis of the discourse sample focused on quantity versus information content over time, and included measures of the total number of syllables produced, the total number of information units produced, and the total time taken to describe the picture. We found that the total number of information-carrying units was the most salient variable which differentiated controls from even the minimal DAT subgroup. Moreover, information content correlated significantly with measures of lexico-semantic processing, but not with performance on non-linguistically based tests.

Introduction

The diagnosis of Alzheimer's disease, which affects up to 5% of people over 65, remains essentially a pathological one, which is confirmed by the finding of extracellular neuritic plaques containing amyloid and neurofibrillary tangles. These changes are maximal in the hippocampal region but are found with increasing frequency in the temporal and other neocortical areas as the disease progresses (Braak and Braak 1991).

Although observers have frequently noted the presence of speech and language disturbances in dementia of the Alzheimer's type (DAT), it is only in the past two decades that neuropsychologists and linguists have turned their attention to this area. A large number of studies have looked at single-word production and comprehension (for a review see Hart 1988). These have established that lexico-semantic disorders

Address for correspondence: J. R. Hodges, University Neurology Unit, Addenbrooke's Hospital, Hill's Road, Cambridge CB2 2QQ, UK.

0268–7038/96 $12·00

occur from an early stage in DAT, and are manifest in reduced category-based verbal fluency (i.e. generation of words from a semantic category, e.g. animals, within a limited time period), impaired naming (e.g. Bayles 1982, Hodges *et al.* 1991), and impaired word–picture matching (e.g. Chertkow and Bub 1990, Martin and Fedio 1983).

More recently, there has been a growing realization that these traditional assessments of relatively discrete linguistic abilities are of limited application in that they fail to tap the complex interactive processes which go to make up connected speech. In an attempt to redress the balance, researchers have begun to focus on discourse abilities, and the development of measures of communicative effectiveness. Discourse entails a complex interaction of linguistic, communicative and other cognitive processes. Its production is, therefore, potentially sensitive to disturbance in any of the areas, and its measurement is useful in assessment of borderline dysfunction, as well as in monitoring progression in more advanced cases.

How should discourse abilities be measured? Two methodological issues deserve consideration. Firstly, what is the best sample on which to perform an analysis? Secondly, what variables should be measured, and how? In an attempt to answer the first question we will review the major methods used to evaluate discourse abilities. These include indirect methods such as subjective rating scales (usually completed by the spouse of clinician), direct methods such as conversational analysis, and the use of naturalistic scenarios, picture descriptions, narratives and procedures to assess spoken output.

The first approach involves the use of *subjective rating scales.* For example, the Edinburgh Functional Communication Profile (Wirz *et al.* 1990), based on speech act theory (Searle 1969), obtains ratings from the subject's partner of functional communication ability in different activities and contexts. Similarly, the Boston Diagnostic Aphasia Examination (Goodglass and Kaplan 1983) contains a clinician's rating scale for prosodic, phonological, lexical, and syntactic abilities. The most recent of these scales is the Communicative Effectiveness Index (Lomas *et al.* 1989) which involves a functional communication questionnaire given to the spouse. It is worth noting, however, that Green (1984) found that spouses tend to overestimate the affected person's abilities, while clinicians tend to underestimate them.

The second approach involves assessing periods of *free form conversation* between the subject and partner in a naturalistic setting. This produces a balanced and informal conversation which is likely to be more representative of everyday abilities. Conversational analysis was found by Crockford (1991) to be the most sensitive way of detecting change in communicative ability in aphasic stroke patients, and more recently Hamilton (1994) has published a detailed study of sociolinguistic interactions with a single DAT patient, charting her changing conversational ability over $4\frac{1}{2}$ years. Although they appear to be of great value in single-case studies, such methods do not lend themselves to standardization for within-group comparisons, since they involve multiple variables, the least quantifiable being pragmatic and psychosocial factors. The remaining methods constrain the desired output to an increasing degree.

The third approach involves describing common scenarios to the subject and *eliciting responses to these everyday situations.* Ulatowska *et al.* (1992) describe a system for analysing the subjects' output in a role-play situation. In the Amsterdam Nijmegen Everyday Language Test (Blomert 1992) the examiner describes a number of scenarios such as 'Your neighbour's dog barks all day long. You are really tired of it. You want to talk to him about it. What do you say?'

More complex processes can be tapped by the use of the fourth and fifth methods, namely *narrative and procedural discourse elicitation tasks* (Osiejuk 1989, Ulatowska *et al.* 1988). These tasks require the inclusion of specific elements with an overall hierarchical structure, which necessitates more sophisticated planning and execution skills. Procedural discourse differs from narrative in that it involves describing how to carry out a behaviour sequence, such as making a cup of tea, or phoning someone from a public call-box. Some methods provide no picture cues, for example the Cinderella story task described by Saffran *et al.* (1989) in which the subject was required to retell the story from memory. Another option is a picture sequence such as the Aesop's Fables cat story (Ulatowska *et al.* 1983). Chenery and Murdoch (1944) used animated sequences which have the potential advantage of generating non-routine, brief samples, which are easy to standardize. However, the presentation of continuous rolling film to DAT patients has the potential problem of confounding linguistic and attentional and/or memory deficits.

Finally, the sixth method, the *description of static visually presented material*, has many advantages in focusing attention, and lessening interference from memory difficulties. Objects, pictures of objects, and composite pictures have all been employed, the two most commonly used examples being Norman Rockwell's 'Easter Morning' picture (Tomoeda and Bayles 1993, Ulatowska and Chapman 1991) and the Cookie Theft picture description task (Goodglass and Kaplan 1983). In the present study we chose the latter task as the most efficient way of obtaining a suitable discourse sample, which could be standardized across many subjects. It provides a clear pictorial focus, thus reducing ambiguity about the subject matter and enabling even those subjects with severe memory problems to stay on track more easily. As a composite picture it affords contrasts of person, time, and place. The key vocabulary is acquired early in life, which should make the nominal aspects of the task easier for subjects with word-finding difficulties (Hirsch and Ellis 1994), and the domestic scene has some degree of familiarity for all subjects.

With regard to the second question of which variables are most appropriate to measure, studies of discourse abilities have employed a wide range of analyses (for a review see Ulatowska and Chapman 1991), but for the purposes of the present study we have concentrated on robust and relatively simple measures which have been applied to patients with DAT.

One of the most consistent findings has been that DAT patients' discourse shows a reduction in information content with a relative increase in redundant phrases. Hier *et al.* (1985), for example, found that DAT subjects produced fewer relevant observations than did normal age-matched control subjects when describing the Cookie Theft picture. Similar findings have been reported by Swindell (1986) and Maxim (1991), though Smith *et al.* (1989) found only a non-significant tendency for their DAT patients to produce less total meaningful content than controls.

In an attempt to quantify the content of DAT patients' discourse, Bayles *et al.* (1989) developed a measure of 'information units'. These were defined as 'a relevant, truthful, non-redundant fact or plausible inference about the stimulus picture'. They found that out of eight discourse measures, information units were the most reliable means of measuring the effect of DAT on connected speech and, in addition, that the number of aborted phrases increased as the disease progressed. Repetition of ideas and revisions increased with moderate disease, then decreased in more severe cases.

Informativeness cannot, however, be assessed merely by quantifying the amount of language produced, since a single message can be concise or verbose. Indeed,

Ulatowska *et al.* (1988) showed that DAT subjects produced significantly less information than normals on narrative and procedural discourse tasks, despite a minimal difference in the total amount of language produced. The difference increased further still when subjects were asked to generate a summary and a moral to the story.

From reviewing the published literature on discourse abnormalities in patients with DAT it is apparent that increasing disease severity is associated with a decrease in the information content with a parallel increase in redundant clauses and aborted phrases. However, a number of issues remain unsettled. Firstly, how early in the course of the disease can such abnormalities be detected? Although the vast majority of patients with DAT present with failing memory for day-to-day events, subtle impairments in the lexico-semantic aspects of language production can be demonstrated on formal tests such as category fluency and confrontational naming tasks (e.g. Bayles *et al.* 1989, Hodges *et al.* 1991, 1992b). What is not clear is whether patients with very mild disease also have measurable deficits in discourse. Furthermore, it is uncertain whether the earliest change is a fall in the absolute information content, or alternatively a fall in the *relative* informativity due to an increase in redundancy and aborted phrases. Secondly, there is the related issue of whether deficits in discourse correlate with performance on more traditional tests of single-word production such as category fluency and naming.

The aims of our study were, therefore, to examine some aspects of discourse in patients with DAT ranging in severity from extremely mild to moderately advanced, and to correlate discourse performance with more conventional neuropsychological measures of linguistic and non-linguistic competence. The primary dependent variables for our analysis were the total number of information units, the number of syllables, the time taken and the relationships between these variables. The patients in the present investigation, in common with other contemporary studies, are described as having dementia of the Alzheimer's type (DAT) with the presumption that the majority would show Alzheimer pathology. Unlike some earlier studies of language competence in demented subjects, we have applied strict diagnostic criteria to exclude, as far as possible, patients with other causes of dementia.

Method

Subjects

Two groups, consisting of a total of 66 subjects, participated in the study: 48 patients with a diagnosis of probable DAT (33 females and 15 males) and 18 normal control subjects (15 females and three males).

The DAT group represents an unselected consecutive series of patients enrolled into a longitudinal study of semantic memory and related cognitive deficits in DAT. The diagnosis of probable DAT was made according to the criteria developed by the National Institute of Neurological and Communicative Disorders and Stroke (NINCDS) and the Alzheimer's Disease and Related Disorders Association (ADRDA), which consist of inclusion and exclusion criteria (McKhann *et al.* 1984). All patients presented with progressive cognitive impairment, predominantly affecting memory, with a history of between 1 and 4 years, and have been followed up for a period of 2–3 years to establish that they show a pattern of progressive decline in cognitive abilities.

Table 1. Basic demographic variables and performance on MMSE in control and DAT subjects: means (and standard deviations in parentheses)

		DAT subjects		
	Controls	Minimal	Mild	Moderate
n	18	16	15	17
Age	68·4(7·3)	72·8(7·1)	66·6(8·7)	63·3(9·1)
Education	11·1(2·5)	10·6(1·8)	11·8(3·7)	10·9(2·6)
MMSE score	29·1(1·1)	25·8(1·9)	21 (1·6)	10·3(4·3)

Patients with a past history of known or suspected transient cerebral ischaemic event or stroke, alcoholism, head injury or major medical illnesses (e.g. cancer, anaemia, thyroid dysfunction) were excluded, as were those with major depression, visual or auditory impairment. All patients were examined by a neurologist and underwent CT or MRI scanning, together with the usual battery of screening blood tests to exclude other treatable causes of dementia. Only subjects with normal-for-age hearing and vision were included.

Normal control subjects were members of the MRC Applied Psychology Unit subject panel, matched with the DAT patients on the basis of age and education level. Subjects with a history of alcoholism, drug abuse, learning disability, neurological or psychiatric illness were excluded.

For correlational analyses the DAT patients were considered as a whole group; to obtain measures of performance on the various tests at different stages of disease severity we also divided the 48 patients into three subgroups on the basis of their scores on the Mini Mental State Examination (MMSE; Folstein *et al.* 1975). As none of the subjects, even the most impaired, could be considered severely demented (for example, no patient was institutionalized or greater than grade 2 (moderate) on the Clinical Dementia Rating Scale (Berg 1988)), the three subgroups have been designated *minimal*, *mild* and *moderate dementia*, in keeping with the classification used in our previous studies (Patterson *et al.* 1994, Hodges and Patterson 1995). The MMSE score ranges for the three subgroups were 24–29 for *minimal*, 14–23 for *mild*, and 2–13 for *moderate*. The upper cut-off of 24 was chosen since this is conventionally regarded as the lower limit of normality. The lower cut-off (14) was chosen as the median value for those cases scoring below 24.

The number of subjects in each of these subgroups, and their mean ages, years of education and MMSE scores, are shown in Table 1, along with values for the same dimensions for the control group. By one-way ANOVA the groups differed significantly in age [$F(3,62) = 3{\cdot}912$, $p > 0{\cdot}05$]; and *post-hoc* analysis by Newman–Keuls tests for pairwise differences between groups showed a significant difference between the minimal and moderate DAT groups only. The fact that the moderate group are the youngest does at least have the following benefit: although some aspects of cognitive performance decline with age (especially in tasks with a large component of problem solving or speeded performance; see e.g. Cerella 1985), in this study any association between disease severity and task performance is very unlikely to be attributable to ageing. The groups were well matched for education (one-way ANOVA, $F(3,62) \leq 1$).

Discourse task

The Boston Cookie Theft picture description task (Goodglass and Kaplan 1983) was chosen to elicit the speech sample for several reasons, mentioned earlier. The samples were recorded on a Marantz tape-recorder using an external microphone, and transcribed by a qualified speech therapist (E.G.). They were obtained either in the subject's home, or at hospital, with only the experimenter and subject present. The instruction given to the subject was: 'Tell me everything you see going on in this picture.' The sample began when the instruction ended, and finished after the subject indicated that he or she had no more to say, or paused for 15 seconds. Interjections unrelated to the picture, such as 'I've seen this before', were excluded. The time taken was measured from the first sound uttered to the end of the last syllable in the sample. For the most part, subjects described the picture without seeking comment or assistance from the experimenter. Interruptions of a verbal nature were avoided by the experimenter; instead non-verbal encouragement was given if the subject looked up. In order to be sure of the correct referent, we noted pointing or gesturing in relation to items in the picture, and any verbal/pointing discrepancies.

Following Tomoeda and Bayles (1993), who found information units to be the most sensitive measure of change in subjects' descriptions of the Norman Rockwell 'Easter Morning' photograph, we chose to look at informativity, efficiency and conciseness. This was achieved by measuring three components in the picture descriptions obtained:

1. The *total number of seconds* (inclusive of all speech acts).
2. The *total number of syllables* uttered within that time frame.
3. The *total number of information units* (IU) produced (see Appendix).

An information unit is the smallest non-redundant fact or interference about the picture, which may range in size from e.g. plural /s/ to a phrase. Repetitions of words or ideas were not included, and where a revision occurred, the first attempt was not counted unless it was clearer than the repair. The use of present tense was regarded as the norm, and any past or future tense counted as new information. Other researchers have employed relevance and plausibility in determining what does or does not count. We did not legislate as to the validity of the information; for example several subjects commented on the old woman lying down in the garden. Once you knew where to look, it was easy to see what they were referring to, albeit unlikely. No controls made this observation, but we decided to count it as information. If a new word for the same thing was used, this did not count as another information unit, e.g. tin for jar. See below for an example.

Sample transcript produced by a 75-year-old lady with DAT, and a MMSE score of 18, in response to the Cookie Theft picture

E: 'Tell me everything you see going on in this picture'

S: */oh yes there's some washing up going on/(laughs) yes/ ..... oh and the other/..... this little one is taking down the cookie jar/ and this little girl is waiting for it to come down so she'll have it/..... er this girl has got a good old splash/she's left the taps on (laughs) she's gone splash all down there/ um..... she's got splash all down there/**

. = approximately 1 second pause

• sample beginning and stopwatch started
** sample and stopwatch stopped
Time between *and **: 60 seconds
Number of syllables: 72
Number of information units—IU: 24.

To check reliability two other experimenters analysed the samples on five subjects. Using an unweighted Kappa test, raters reached 0·97 and 0·95 reliability (Kraemer 1982). Discrepancies were discussed and criteria clarified.

Other neuropsychological tests

In addition to the MMSE and the Cookie Theft picture description task, all subjects were given a comprehensive semantic memory test battery which has previously been described in detail (Hodges *et al.* 1992a, b). This battery of tests, all employing one consistent set of stimulus items, is designed to assess input to and output from central representational knowledge about the same group of items via different sensory modalities. It contains 48 items chosen to represent three categories of animals (land animals, sea creatures, birds) and three categories of man-made items (household items, vehicles and musical instruments) matched for category prototypicality and word frequency. The items were chosen from the corpus of line drawings by Snodgrass and Vanderwart (1980). The full battery consists of six subtests, but as the focus of the present study is on language production, only data from the first four will be presented.

Category fluency for four man-made (subordinate category boats included in addition to those mentioned above) and four living categories (subordinate category dogs included). One minute was allowed per category.

Naming of all 48 line drawings without cueing.

Naming in response to the verbal description (e.g. 'what do we call a two-wheeled vehicle propelled by the rider?')

Semantic feature questions: for each item we designed eight questions such that four questions explore knowledge of physical features (size, shape, colour, etc.) and the other four explore knowledge of non-perceptual attributes (habitat, diet, uses, etc.) One-half of the questions posed should receive 'yes' answers, and the other half 'no' answers (e.g. does a zebra have stripes?; does a zebra eat meat?)

As a measure of working memory we used *digit span* (WAIS-R), and to assess visuospatial and perceptual processing we employed the *Judgement of Line Orientation Test* (JLO) (Benton *et al.* 1978) and an 'unusual views' *object matching test* (Humphreys and Riddoch 1984). The JLO examines ability to judge and match angular relationships. The subject views pairs of lines with various orientations, and is required to match their angulation with reference to a display of 11 numbered lines spanning 180 degrees, which are viewed simultaneously with the target lines. In the object matching test the subject is presented with an array of three photographs consisting of a target object photographed from a conventional angle, presented above two response alternatives: (i) the same target object photographed from a different (unusual) view, and (ii) a photo of a different but visually similar object. The subject is asked to indicate which of the lower two photographs is of the same object as that shown above.

Table 2. Results of Cookie Theft analyses on the controls and DAT subjects: means (and standard deviations in parentheses)

		DAT subjects		
	Controls	Minimal	Mild	Moderate
Time (seconds)	94·2(57·5)	83·8(43)	78·5(49·1)	73·5(46·3)
Syllables	162·1(101·9)	143·7(52·1)	122·7(72·6)	89·6(71·3)
Info. units	46·6(23·8)	34·6(11·1)	28·9(14·4)	15·1(12·6)
Info/second	0·6(0·2)	0·5(0·2)	0·4(0·2)	0·2(0·2)
Syllables/second	1·9(0·7)	2·0(1·0)	1·8(0·9)	2·3(2·4)
Syllables/info.	3·6(1·2)	3·7(1·9)	4·3(1·5)	5·8(5·7)

Results

Cookie Theft analysis

As shown in Table 2 there was a decrease in the total time taken across the groups; but a one-way ANOVA showed no significant group difference [$F < 1$]. The total number of syllables produced was, however, significantly different [$F(3,62) = 2{\cdot}77$, $p < 0{\cdot}04$]; *post-hoc* pairwise comparisons revealed a significant difference between the controls and the moderate DAT subgroup only, with no evidence of a significant reduction in the two less impaired subgroups compared to controls (i.e. controls = minimal = mild > moderate).

For information units, a one-way ANOVA also revealed a highly significant main effect [$F(3,62) = 10{\cdot}79$, $p < 0{\cdot}0001$], and *post-hoc* comparisons showed significance pairwise differences between the controls and all three DAT subgroups, as well as between most pairs of subgroups (controls > minimal = mild > moderate). It is notable that even the minimally impaired patients showed a significant decline in the number of information units produced.

When analysed in terms of the relationships between variables, the data revealed a highly significant main effect for information units per second, a measure of efficiency [$F(3,62) = 6{\cdot}99$, $p < 0{\cdot}001$]; but only the moderate DAT subgroup differed significantly from the others (controls = minimal = mild > moderate). Thus, although the minimal and mild subgroups showed a statistically reliable decrement in information units uncorrected for time, relative to controls, the measure of information per time yielded only a non-significant trend in these two less impaired subgroups.

One-way ANOVAs showed no significant differences between the controls and DAT groups on syllables per second (speech rate, inclusive of pauses), or syllables per information unit (verbosity) [$F(3,62) = 0{\cdot}32$ and $1{\cdot}6$ respectively, $p > 0{\cdot}05$].

In order to compare the relative performance of the patient groups on the variables of interest, we normalized the patients' scores as percentages of the control means. Figure 1 illustrates the marked decline in information units with increasing disease severity, and the less precipitous reduction in the other parameters.

Correlation with other neuropsychological measures

Performance on the various neuropsychological measures is shown in Table 3. For all tests there was an overall group difference [$F(3,62)$ ranging from $15{\cdot}1$ to $37{\cdot}9$, $p < 0{\cdot}001$] and *post-hoc* pairwise comparisons revealed that the minimal subgroup were impaired,

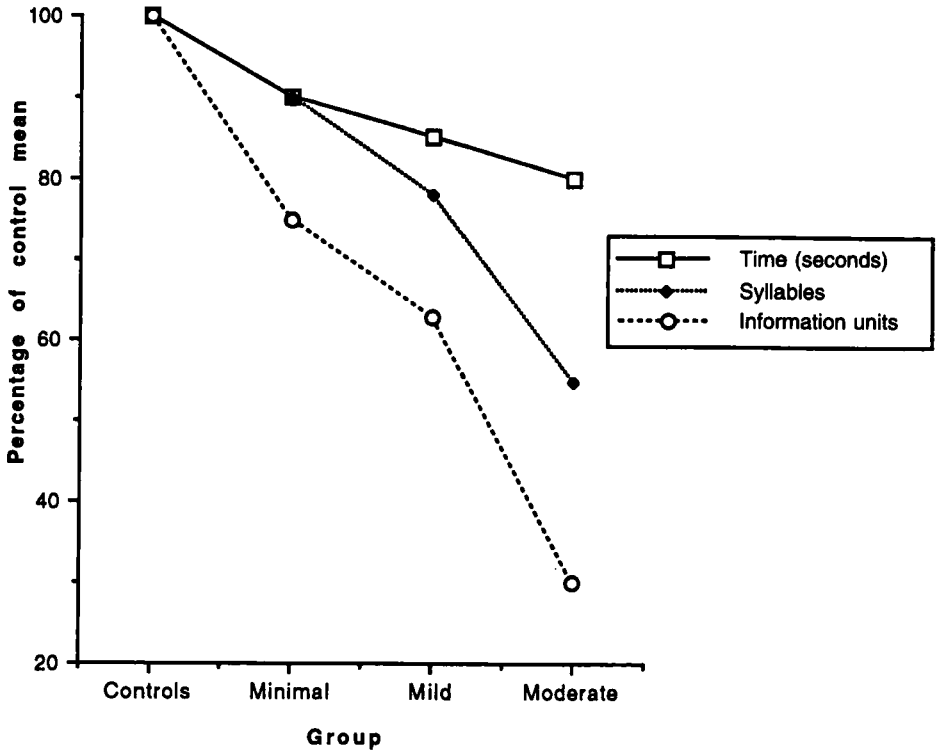

Figure 1. Normalized results of the DAT subgroups' performance describing the Boston Cookie Theft picture.

Table 3. Neuropsychological test scores of the controls and minimal, mild and moderate DAT groups: means (and standard deviations in parentheses)

		DAT subjects		
Test	Controls	Minimal	Mild	Moderate
Digit span F	6·7(1·1)	6·8(1·2)	5·7(1·1)	3·9(1·8)
Digit span B	4·9(1·2)	5(1·2)	3·4(0·7)	2(1·2)
Line orientation/30	27·9(2·8)	25·6(5)	21·5(4·6)	4·3(8·2)
Object matching/40	38·2(1·7)	35·5(2·6)	35·5(3·6)	28·6(4·4)
Pyramids and palm trees/52	51·3(0·9)	47·8(4·4)	46·1(4·8)	37·8(5·4)
Fluency: living (four categories in each)	59·7(11·5)	35·1(12·7)	26·7(12·7)	10·3(9)
Fluency: man-made	56·7(7·7)	33·3(10·2)	28·7(11·3)	9·9(7·4)
Naming/48	43·7(2·4)	38·9(4·8)	37·7(7·2)	26·4(9·2)
Naming to description/24	22·6(1)	19(3)	16·3(5·8)	7·5(6·3)
Semantic feature questions/186	178·9(4·6)	169(11·2)	164·3(17·9)	136·6(18·2)

relative to controls, on all components of the semantic memory test battery—category fluency, naming, naming to description and the semantic feature questions—but not on other tasks. The mild DAT subgroup were impaired on all tests except object matching, while the moderate DAT subgroup showed universal impairment.

For the DAT group as a whole there were highly significant correlations ($p < 0{\cdot}01$) between each of the major discourse variables, with the exception of time taken (i.e. syllables produced, information units and information/time) and *all* of the other neuropsychological variables. Examination of the data suggested that this was almost certainly due to the fact that the moderately demented subgroup showed severe impairment on every test. We decided, therefore, to exclude the moderate subgroup of 17 patients and to confine the correlation analyses to the minimal and mild subgroups (combined group size = 31). Interestingly, when the more impaired patients were excluded, there was no significant correlation between MMSE score and any of the discourse variables, showing that at the earlier stages of dementia there is no clear correspondence between dementia level and ability to convey information in this kind of picture description task. There was, however, a significant correlation between information units produced and performance on category fluency ($r = 0{\cdot}38$, d.f. = 29, $p < 0{\cdot}05$), naming ($r = 0{\cdot}35$, d.f. = 27, $p < 0{\cdot}05$) and semantic feature questions ($r = 0{\cdot}4$, d.f. = 27, $p < 0{\cdot}005$), but not on naming to description. Performance on the discourse measures did not correlate significantly with digit span, the JLO, or object matching.

In summary, the reduction in information units seen in the more mildly demented patients correlated with performance on the category fluency, semantic feature questions and picture naming subtests of the semantic memory battery, but not with overall level of dementia or with non-linguistic measures.

Discussion

Our original questions concerned how early in the disease process evidence of discourse impairments might be detectable, and the changing interrelationships of discourse and other language use over time. Our findings confirm the existence of a significant reduction in the information content of discourse in DAT patients (Hier *et al.* 1985, Swindell 1986, Bayles *et al.* 1989, Ulatowska *et al.* 1988). We have extended this work by showing that such changes can be detected very early in the course of DAT; in our study, patients in the minimal subgroup, whose MMSE scores were above the traditional cut-off point of 24, demonstrated a significant reduction in the number of information units produced. The diagnosis in the minimal subgroup has been substantiated by the demonstration of deterioration in two or more cognitive domains over a 2–3-year period of follow-up. Within the minimal and mild subgroups there was, however, no clear relationship between the discourse measures and the level of dementia, implying that considerable heterogeneity exists in the earlier stages of DAT.

We have also extended previous work in this area by showing that the deficits involving discourse variables correlate with more traditional measures of lexico-semantic competence, particularly category fluency and naming, but do not correlate with other non-linguistic variables. The finding of reduced information content was predicted since, as mentioned in the Introduction, this has been the outcome in the majority of studies to date. For instance, Hier *et al.* (1985) found that DAT patients produced fewer relevant observations than normal age-matched controls. Similarly,

Bayles *et al.* (1989) analysed DAT patients' oral and written descriptions of Norman Rockwell's picture of Easter Morning and a story retelling task, and found a significant reduction in the number of information units as compared to normal controls on all tasks, with delayed story retelling being the most discriminating condition. Ulatowska *et al.* (1988) demonstrated the effect of cognitive complexity on informativity: as well as narrative and procedural discourse tasks, DAT subjects were required to provide a summary and a moral for one story. Compared to normals, the DAT subjects produced fewer *a-priori* propositions on all the tasks, with an even greater differential on the moral and summary task which requires both memory and convergent thinking.

The positive correlation between the amount of information relayed in the discourse task and performance on the semantic battery tests suggests that these two deficits may be linked. Impaired lexico-semantic processing has been consistently found in patients with DAT using single-word-based tests of language production (e.g. category fluency, picture naming, etc.) and comprehension (e.g. word–picture matching) (for review see Hart 1988). Although the underlying cause of these deficits has been a topic of controversy, most of the evidence favours a loss of semantic information rather than a disorder of access (Chertkow and Bub 1990, Hodges *et al.* 1992; but for a counter-argument see Nebes 1989). Recently it has been shown that breakdown in semantic abilities can be found very early in the course of DAT (Hodges and Patterson 1995). Moreover, these findings concur with the known distribution of pathology in DAT: the medial temporal complex bears the brunt in the initial stages, followed by more lateral neocortical structures (Braak and Braak 1991) which are thought to be critical for semantic processes (Patterson and Hodges 1995). It is clear from the present study that deficits in lexico-semantic processing are correlated with deficits on a task—description of the Cookie Theft picture—which measures more everyday language abilities.

A less expected finding of our study was the relatively unaffected performance of the minimal and mild subgroups on measures of the number of syllables produced, and the total time taken, while the information content of the descriptions dropped significantly. Thus, speech rate (syllables per second) remained much the same for each group, while the number of syllables per information unit was progressively greater with more advanced dementia. It is unclear whether this tendency to increase output is a cause or an effect of the reduced information. The types of gross changes in verbal output reported by other workers—logorrhoea (press of speech), palilalia (involuntary repetitions of segments), perservation (inability to shift to another behaviour) and echolalia (involuntary repetition of a stimulus)—were present in some of our moderately demented group (MMSE 2–13) who were perhaps more similar to those included in previous studies, although it is difficult to be certain since comparable scores on dementia rating scales have not always been reported. These features are likely to reflect the involvement of diffuse brain regions—including prefrontal areas—found in patients with end-stage DAT.

Our study of information units supplements previous detailed investigations of coherence and cohesion (Ulatowka *et al.* 1988, Ripich and Terrell 1988) and self-monitoring and repair (McNamara *et al.* 1992). One key difference between DAT and elderly subjects is the pattern of repair and self-monitoring. McNamara *et al.* (1992) assessed speech output monitoring in the Boston Cookie Theft picture descriptions of patients with DAT, Parkinson's disease (PD) and healthy subjects. They found significant differences between the three groups: the DAT group corrected only a quarter of their total speech errors as contrasted with 82% for the controls. There was

also a qualitative difference in the pattern of their errors. In the healthy elderly subjects there were more *lemma* repairs, which are single-word substitutions (e.g. 'chair, er, stool'). The younger healthy subjects tended to produce approximately equal numbers of lemma and *reformulation* repairs (e.g. 'he fell over the, is falling over'), which altered the syntax. The DAT patients had a greater propensity for error production, and were capable of a degree of self-monitoring, but could not apply the process of repair as often as necessary.

In conclusion, this study demonstrates the sensitivity of even a quick and simple picture description task in detecting deficits in complex integrated communication abilities. Ulatowska *et al.* (1988) made the point that monologue language tasks may help conceal patients' deficits through avoiding pragmatic and sociolinguistic constraints found in interaction. Our finding does not contradict this, since it is possible that impairment may have been detectable via conversational analysis even earlier in our subjects. Having established the existence of a significant impairment in meaningful, informative communication in the very early stages of Alzheimer's disease, we note the need to investigate further the nature of the mechanisms leading to this impairment, and management strategies which help minimize dysfunction.

Acknowledgements

We are indebted to all the families who took part in this research. Thanks also to Naida Graham and Lindsey Stuart-Green for assistance in testing patients, and Hilary Baddeley for testing control subjects. Naida Graham also provided invaluable statistical input. This research was supported in part by a research grant from the Medical Research Council, UK, to J.R.H.

References

BAYLES, K. A. (1982) Language function in senile dementia. *Brain and Language*, **16**, 265–280.

BAYLES, K. A., BOONE, D. R., TOMOEDA, C. K., SLAUSON, T. J. and KASZNIAK, A. (1989) Differentiating Alzheimer patients from the normal elderly and stroke patients with aphasia. *Journal of Speech and Hearing Disorders*, **54**, 74–87.

BAYLES, K. A., TOMOEDA, C. K. and TROSSET, M. W. (1992) Relation of linguistic communication abilities of Alzheimer's patients to stage of dementia. *Brain and Language*, **42**, 454–472.

BENTON, A. L., VARNEY, N. R. and HAMSHER, K. (1978) Visuospatial judgement: a clinical test. *Archives of Neurology*, **35**, 364–367.

BERG, L. (1988) Clinical dementia rating (CDR). *Psychopharmacology Bulletin*, **24**, 637–639.

BLOMERT. L. (1994) The Amsterdam–Nijmegen Everyday Language Test (ANELT). In L. Blomert (Ed.) *Assessment and Recovery of Verbal Communication in Aphasia* (Koninklijke Bibliothek, Den Haag), pp. 9–34.

BRAAK, H. and BRAAK, E. (1991) Neuropathological stageing of Alzheimer-related changes. *Acta Neuropathologica*, **82**, 239–259.

CERELLA, J. (1985) Information processing rates in the elderly. *Psychological Bulletin*, **98**, 67–83.

CHENERY, H. J. and MURDOCH, B. E. (1994) The production of narrative discourse in response to animations in persons with dementia of the Alzheimer's type: preliminary findings. *Aphasiology*, **8**, 159–171.

CHERTKOW, H. and BUB, D. (1990) Semantic memory loss in Alzheimer's type: what do various measures measure? *Brain*, **113**, 397–417.

CROCKFORD, C. (1991) Assessing functional communication in aphasic adults: a comparison of three methods. Unpublished dissertation, University of Newcastle upon Tyne.

FOLSTEIN, M. F., FOLSTEIN, S. E. and MCHUGH, P. R. (1975) 'Mini-mental' state: a practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, **12**, 189–198.

GOODGLASS, H. and KAPLAN, E. (1983) *The Assessment of Aphasia and Related Disorders*, 2nd edn (Lea & Febiger, Philadelphia, PA) .

GREEN, G. (1984) Communication in aphasia therapy: some of the procedures and issues involved. *British Journal of Disorders of Communication*, **19**, 35–36.

HAMILTON, H. E. (1994) *Conversations with an Alzheimer's Patient: an Interactional Sociolinguistic Study* (Cambridge University Press, Cambridge).

HART, S. (1988) Language and dementia: A review. *Psychological Medicine*, **18**, 99–112.

HIER, D. B., HAGENLOCKER, D. and SCHINDLER, A. G. (1985) Language disintegration in dementia: effects of aetiology and severity. *Brain and Language*, **25**, 117–133.

HIRSCH, K. W. and ELLIS, A. W. (1994) Age of acquisition and lexical processing in aphasia: a case study. *Cognitive Neuropsychology*, **11**, 435–458.

HODGES, J. R. and PATTERSON, K. (1995) Is semantic memory consistently impaired early in the course of Alzheimer's disease? Neuroanatomical and diagnostic implications. *Neuropsychologia* (in press) .

HODGES, J. R., SALMON, D. P. and BUTTERS, N. (1991) The nature of the naming deficit in Alzheimer's and Huntingdon's disease. *Brain*, **114**, 1547–1558.

HODGES, J. R., PATTERSON, K., OXBURY, S. and FUNNELL, E. (1992a) Semantic dementia: progressive fluent aphasia with temporal lobe atrophy. *Brain*, **115**, 1783–1806.

HODGES, J. R., SALMON, D. P. and BUTTERS, N. (1992b) Semantic memory impairment in Alzheimer's disease: failure of access or degraded knowledge? *Neuropsychologia*, **30**, 301–314.

HUMPHREYS, G. W. and RIDDOCH, M. J. (1984) Roots to object constancy: implications for neurological impairment of object constancy. *Quarterly Journal of Experimental Psychology*, **36**, 384–415.

KRAEMER, H. C. (1982) Kappa co-efficient. In S. Kotz and N. L. Johnson (Eds) *Encyclopaedia of Statistical Sciences* (Wiley & Sons, New York), pp. 352–354.

LOMAS, J., PICKARD, L., BESTER, S., ELNARD, H., KINLAYSON, A. and ZOGHAIB, C. (1989) The Communicative Effectiveness Index: development and effectiveness of psychometric evaluation of a functional communication measure for adult aphasia. *Journal of Speech and Hearing Disorders*, **54**, 113–124.

MCKHANN, G., DRACHMAN, D., FOLSTEIN, M., KATZMAN, R., PRICE, D. and STADLAN, E. M. (1984) Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA work group under the auspices of the Department of Health and Human Services Task Force on Alzheimer's disease. *Neurology*, **34**, 939–944.

MCNAMARA, P., OBLER, T. K., AU, R., DURSO, R. and ALBERT, M. L. (1992) Speech monitoring skills in Alzheimer's disease, Parkinson's disease and normal ageing. *Brain and Language*, **42**, 38–51.

MARTIN, A. and FEDIO, P. (1983) Word production and comprehension in Alzheimer's disease: the breakdown of semantic knowledge. *Brain and Language*, **19**, 124–141.

MAXIM, J. (1991) Can elicited language be used to diagnose dementia? A comparison of demented, depressed, aphasic and normal elderly subjects. Work in Progress. National Hospitals College of Speech Sciences.

NEBES, R. D. (1989) Semantic memory in Alzheimer's disease. *Psychological Bulletin*, **106**, 377–394.

OSIEJUK, E. (1989) Discourse analysis in aphasics and right-brain-damaged patients. *Polish Psychological Bulletin*, **20**, 139–146.

PATTERSON, K., GRAHAM, N. and HODGES, J. R. (1994) Reading in dementia of the Alzheimer's type: a preserved ability? *Neuropsychology*, **8**, 395–407.

RIPICH, D. and TERRELL, B. (1988) Patterns of discourse cohesion and coherence in Alzheimer's disease. *Journal of Speech and Hearing Disorders*, **53**, 8–15.

SAFFRAN, E. M., BERNDT, R. S. and SCHWARTZ, M. F. (1989) The quantitative analysis of agrammatic production: procedure and data. *Brain and Language*, **3**, 440–479.

SEARLE, J. (1969) *An Essay in the Philosophy of Language* (Cambridge University Press, London).

SMITH, S. R., CHENERY, H. and MURDOCH, B. (1989) Semantic abilities in dementia of the Alzheimer type II: grammatical semantics. *Brain and Language*, **36**, 533–542.

SNODGRASS, J. G. and VANDERWART, M. (1980) A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, **6**, 174–215.

SWINDELL, C. S. (1986) Expressive language disturbance in probable Alzheimer's disease. Unpublished PhD thesis, University of Pittsburgh.

TOMOEDA, C. K. and BAYLES, K. A. (1993) Longitudinal effects of Alzheimer's disease on discourse production. *Alzheimer's Disease and Associated Disorders*, **4**, 223–236.

ULATOWSKA, H. K. and CHAPMAN, S. B. (1991) Discourse studies. In R. B. Lubinski (Ed.) *Dementia and Communication* (Mosby, St Louis), pp. 115–131.

ULATOWSKA, H., ALLARD, L., DONNELL, A., BRISTOW, J., HAYES, S., FLOWER, A. and NORTH, A. (1988) Discourse performance in subjects with dementia of the Alzheimer's type. In H. Whitaker (Ed.) *Neuropsychological Studies of Nonfocal Brain Damage* (Springer Verlag, New York), pp. 108–131.

ULATOWSKA, H. K., ALLARD, L., REYES, B. A., FORD, J. and CHAPMAN, S. (1992) Conversational discourse in aphasia. *Aphasiology*, **6**, 325–331.

ULATOWSKA, H. K., FREEDMAN-STERN, R., DOYEL, A. W. and MACULOSO-HAYNES, S. (1983) Production of narrative discourse in aphasia. *Brain and Language*, **19**, 317–334.

WIRZ, S. L., SKINNER, C. and DEAN, E. (1990) The revised Edinburgh Functional Communication Profile, Communication Skills Builders, Arizona.

Appendix: Information units

An information unit is the smallest non-redundant meaningful fact or inference. In this analysis, only discourse that was describing the picture was included.

One unit may range in size from a plural /s/ to a phrase.

With verbs, where past or future tense is indicated, an extra point is scored.

The same word may score more than one point if on the second occasion it is used in a new context. For example, if the subject has already described how the mother is wiping up, then gone on to talk about the children, the second reference to the mother in the context of the sink overflowing would count as another information unit.

Repetitions of the same information (even if using different vocabulary) were discounted. For example, some subjects referred to the cookie jar as a tin on a second occasion.

Where a revision (self-correction) occurred, the repaired words or utterances were scored rather than the initial version unless the first version was clearer.

Other researchers have employed relevance and plausibility criteria in determining what does or does not count. We did not legislate as to the validity of the information, but only gave points to information that was coherent. For example several subjects commented on the old woman lying down in the garden. Once you knew where to look, it was easy to see what they were referring to, albeit unlikely. No controls made this observation, but we decided it counted as information.

Examples of information units:
the woman = 1
the *young woman* = 2
she has washed up = 3 (person, action, tense)
she was staring out of the window = 4 (person, action, direction, tense)
in there = 1
in the cupboard = 2
holding out her *hand* = 2 (it couldn't be anyone else's hand)