Aphasiology

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

Analysis of spontaneous, conversational speech in dementia of Alzheimer type: Evaluation of an objective technique for analysing lexical performance

R. S. Bucks , S. Singh , J. M. Cuerden & G. K. Wilcock

To cite this article: R. S. Bucks , S. Singh , J. M. Cuerden & G. K. Wilcock (2000) Analysis of spontaneous, conversational speech in dementia of Alzheimer type: Evaluation of an objective technique for analysing lexical performance, Aphasiology, 14:1, 71-91, DOI: 10.1080/026870300401603

To link to this article: https://doi.org/10.1080/026870300401603

Published online: 31 Aug 2010.

APHASIOLOGY, 2000, VOL. 14, NO. 1, 71–91

Analysis of spontaneous, conversational speech in dementia of Alzheimer type: Evaluation of an objective technique for analysing lexical performance

R. S. BUCKS[1], S. SINGH[2], J. M. CUERDEN[2] and G. K. WILCOCK[2]

[1] Clinical Research Centre and Memory Disorders Clinic, Blackberry Hill Hospital, Manor Road, Bristol, BS16 2EW, UK
[2] Department of Computer Science, University of Exeter, Exeter EX4 4PT, UK

(*Received 21 July 1998; accepted 8 March 1999*)

Abstract

Spontaneous, conversational speech in probable dementia of Alzheimer type (DAT) participants and healthy older controls was analysed using eight linguistic measures. These were evaluated for their usefulness in discriminating between healthy and demented individuals. The measures were; noun rate, pronoun rate, verb rate, adjective rate, clause-like semantic unit rate (all per 100 words), including three lexical richness measures; type token ratio (TTR), Brunét's Index (W) and Honoré's statistic (R). Results suggest that these measures offer a sensitive method of assessing spontaneous speech output in DAT. Comparison between DAT and healthy older participants demonstrates that these measures discriminate well between these groups. This method shows promise as a diagnostic and prognostic tool, and as a measure for use in clinical trials. Further validation in a large sample of patient versus control 'norms' in addition to evaluation in other types of dementia is considered.

Introduction

Dementia is characterized by the breakdown of intellectual and communicative functioning accompanied by personality change (DSM-IV, American Psychiatric Association 1994). Communication disorders are a common feature of dementia (Bayles *et al*. 1987, Kempler 1991), being present in 88–95% of sufferers (Thompson 1987). They are particularly pronounced in probable dementia of Alzheimer type (DAT, Cummings *et al*. 1985) and include word-finding deficits, paraphasias and comprehension impairment (Alzheimer 1907, Appell *et al*. 1982, Bayles 1982, Obler 1983, Irigaray 1973). More recent research has found additional evidence for impaired performance on verbal fluency tasks (Phillips *et al*. 1996, Becker *et al*. 1988), circumlocutory responses (Hodges *et al*. 1991) and impairments in discourse, which worsen over the course of the disease (Hutchinson and Jensen 1980, Ripich and Terrell 1988, Ulatowska and Chapman 1991). Phonemic and

* Please address correspondence to Romola Bucks, e-mail: romola.bucks@bristol.ac.uk

Aphasiology ISSN 0268–7038 print/ISSN 1464–5041 online © 2000 Psychology Press Ltd
http://www.psypress.co.uk

syntactic processes, however, have been shown to be relatively preserved (Hodges *et al*. 1991, Kertesz *et al*. 1986, Appell *et al*. 1982, Hier *et al*. 1985, Kempler and Zelinski 1994, Schwartz *et al*. 1979).

Assessment of aphasia in DAT, as in other disorders, can be difficult. For example, structured aphasia assessments can sometimes be insensitive to the early communication deficits experienced by sufferers and observed by their families in normal conversational situations (Crockford and Lesser 1994). Likewise, low scores on structured tests may misrepresent the true level of deficit present. Sabat (1994, p. 334) argues, in a critical review of language dysfunction research in DAT that:

> Many investigators who report studies on 'speech' (sic) use standard tests ... although these may be accurate tests of the instrumentalities of speech, these are not tests of natural speech. Rather, they tend to break the natural course of speech into its component parts and test the use of the parts in the absence of the larger social psychological contexts in which language and speech are used naturally.

Sabat further argues that low scores on such tests do not necessarily represent the participant's performance unless that same difficulty is observed in natural conversation.

Many 'structured' tasks have been used to assess individuals with DAT, such as confrontation naming (see e.g. Bayles *et al*. 1987, Bayles *et al*. 1989, Hodges *et al*. 1991), single word production (see e.g. Martin and Fedio 1993), or generation of words beginning with a certain letter (verbal fluency; see e.g. Phillips *et al*. 1996; Becker *et al*. 1988). However, although previous researchers have found deficits in qualitative aspects of DAT sufferer's conversation using discourse techniques (Ripich *et al*. 1991, Hutchinson and Jenson 1980, De Santi *et al*. 1994), a search of the literature yielded little previous research into quantitative characteristics of conversational, spontaneous speech in DAT. Many of the studies purportedly assessing spontaneous speech have actually measured picture description (Nicholas *et al*. 1985, Hier *et al*. 1985, Croisile *et al*. 1996). Three other studies were found which used spontaneous speech. The first used a semi-standardized interview to compare the performance of 10 DAT, 5 Wernicke's aphasics and 5 normal older controls on spontaneous speech (Blanken *et al*. 1987). Participants' speech was transcribed and a mixture of methods used to analyse the conversations, including measures of average sentence length (divided into simple and complex sentences), numbers of words in each class (nouns, verbs, adjectives and adverbs), type token ratio and instances of word finding difficulties. Significant differences between participant groups were found, though the study was compromised by a difference in the total length of speech recorded for each group and the relative brevity of the interviews; 5–10 minutes. The second, a study of 150 patients with DAT (Sevush *et al*. 1993), assessed language using spontaneous speech, comprehension, repetition, oral reading, writing and naming. Spontaneous speech was evaluated for fluency, syntax and paraphasias. Each participant's performance was graded as normal, mildly impaired, or markedly impaired. The authors found differences in early and late onset DAT patients' performance on these measures. Interestingly they found that object naming was worse in late onset DAT, but spontaneous speech was worse in early onset DAT participants. The third study followed 63 patients with Alzheimer disease rating their spontaneous speech during a 4 minute interview on 6 scales (communication, articulation and prosodics, automatic speech, semantic structure, phonematic structure and syntactic structure) (Romero and Kurz 1996). Despite large samples, these last two studies were limited by the use of qualitative rating scores, which might reduce the sensitivity of the measures employed.

The assessment of spontaneous speech appears to offer a more naturalistic assessment of language, which may prove to be a powerful method of assessing individuals with DAT, as well as other dementias. Such assessment might form the basis of new, more sensitive, assessment tools. These tools may allow earlier diagnosis of DAT as well as the evaluation of individuals with severe dementia who cannot tolerate formal psychological testing (Mohs *et al*. 1986). In addition, they may also offer accurate measurement of improvement with treatment. Experience with anticholinergic treatments suggests that the families of sufferers report subtle changes in communicative behaviour, which are not necessarily measured by many assessment tools often used in clinical trials. These changes in linguistic behaviour are often observed in the context of everyday conversation. Another potential benefit is the impact of any findings on diagnostic specificity. A number of researchers have posited the view that DAT can be divided into two subtypes. The principal difference between these subtypes being the age at onset and the severity of language deficits. For example, it has been suggested that DAT can be divided into early (young) versus late onset (older) (Koss *et al*. 1996). Whilst some researchers hold the view that younger onset sufferers experience greater language impairment (Seltzer and Sherwin 1983, Sevush *et al*. 1993, Binetti *et al*. 1993, Goldblum *et al*. 1994, Becker *et al*. 1988, Chui *et al*. 1985, Filley *et al*. 1986, Selnes *et al*. 1988), others hold that it is later onset sufferers who experience greater impairment (Bayles 1991, Bayles *et al*. 1992, 1993). Still others have found no relationship between language impairment and age at onset (Cummings *et al*. 1985, Martin *et al*. 1985, Grady *et al*. 1987). Sevush *et al*. (1993) criticise many of these studies for using relatively small sample sizes and some for depending on qualitative testing (e.g. Seltzer and Sherwin 1983). The difficulty in reconciling these findings may arise from the different language assessment methods used and the degree to which they may be affected by comprehension, attention, memory and other deficits (Kempler *et al*. 1995). Clearly, comprehension, concentration and memory also play important roles in conversation, in particular in helping the individual maintain his/her flow of speech and remember what they were trying to say. However, conversation is an over-learned skill, which may prove less susceptible to these difficulties than many traditional tests of language. Finally, evidence from formal testing indicates that the severity of language impairment is predictive of the rate of decline of DAT sufferers (Boller *et al*. 1991). Thus, the possibility of detecting deficits in language function before they are found on structured assessment has major implications for diagnosis, prognosis, planning interventions, measuring their outcome, and for service management.

It would seem important, therefore, to establish a method of analysing spontaneous speech in DAT that is objective, relatively easy to administer and interpret, and sensitive enough to measure changes. The method must also allow comparison between DAT patients and non-demented individuals (i.e. be sensitive to a wide range of performance) and should not be heavily dependent on memory or perception for test stimuli or instructions. A recently developed technique for the linguistic analysis of aphasic speech (Singh 1994, 1995, 1996, Holmes and Singh 1996) looked promising. Singh (1994) originally proposed the technique for quantifying conversational performance using word-frequency measures in aphasia. The technique uses a set of linguistic measures that are dependent on word-frequencies. This approach has been extensively tested on small samples (28 participants: Singh 1995) and on large samples (100 participants: Singh 1996, Holmes and Singh 1996)). The technique has also been used to discriminate between individuals with dysphasia and normal adults using a neural network (Singh 1997) and has been shown to be sensitive to change in patient progress in single case studies (Singh and Bookless 1997). Finally, test-retest reliability has been established (Singh 1996). The technique differs from conversational analysis, which has generally been used for quantifying the pragmatic aspects of conversation and works with turn-taking data (see Lesser and Milroy 1993). The technique used here requires interviewing and transcribing spontaneous speech recorded in the context of a conversational setting. These transcriptions are then analysed using a series of objective, clearly defined linguistic measures.

We report a study in which we analysed the structure of spontaneous, conversational speech of persons with a diagnosis of probable dementia of Alzheimer's type (DAT). The aims of the study were to establish the utility of this method for analysing spontaneous speech in a sample of DAT patients and to compare the performance of these participants with that of an age and gender matched group of healthy older adults. Should the method prove sensitive to differences between healthy older controls and individuals with an established diagnosis of DAT, then further investigation of its sensitivity in early or undiagnosed dementia would be warranted.

Power

On evidence available from Singh (1996), healthy individuals produced a mean Noun-rate of 15.0 and SD of 0.4. It was predicted that DAT patients would have a mean below and outside the range of normal subjects. This would suggest setting a minimum size effect of 4 standard deviations below the mean in normal subjects. To account for possible greater spread in healthy older adults, a standard deviation of twice the reported size was used in the power calculations. Using these values, on the assumption of a significance level of 5%, power of 95% and a sample size ratio of 2:1, groups of 12 and 6 would be sufficient (Dallal 1990). We already had available 16 healthy older participant controls and therefore sought a matching 8 DAT participants.

Sample

There were 24 participants: 8 individuals with a diagnosis of probable dementia of Alzheimer type (DAT: 4 female, 4 male) and 16 healthy older controls (HO: 7 male, 9 female). The DAT group comprised individuals suffering with mild to moderate-severe cognitive impairment as measured by the Mini-Mental State Examination (MMSE: Folstein *et al.* 1975). A diagnosis of probable DAT was made according to DSM-III-R and NINCDS-ADRDA criteria (American Psychiatric Association 1987, McKhann *et al.* 1984). Each patient attended a memory disorders clinic where they underwent thorough medical screening in order to rule out any other treatable pathology that could explain their impairment. This included neuropsychological assessment, laboratory blood testing and Computed Tomography (CT) scanning of the head. Patients with a score of 5 or

Table 1. Age, years of education, MMSE, duration of illness and handedness for healthy older participants and participants with dementia of alzheimer type

	Healthy older adults N = 16			Participants with DAT N = 8		
	M	SD	range	M	SD	range
Age	61.8	7.7	51–79	67.4	6.2	57–77
Years of education	15.1**	2.4	9–16	10.3	2.6	9–16
MMSE				15	6.8	3–24
Duration of illness (months)				48.1	13.2	31–70
Handedness	16 right			7 right, 1 left		

** $p < 0.01$.

greater on the Hachinski scale (Hachinski *et al.* 1975) were excluded to reduce the possibility of including vascular dementia. Any patient whose diagnosis was changed at follow up 6 months to 1 year later was removed from the study. HO participants were not suffering from any neurological or psychological illness likely to impair performance and were not complaining of memory or other cognitive difficulties. These participants had taken part in a larger study comparing Broca's dysphasics with healthy participants (Singh 1996). Table 1 shows the mean age, years of education and handedness for DAT and HO participants, with MMSE and duration of illness for DAT participants. There were no significant differences in the distribution of males and females ($p > .05$) or in mean age ($p > .05$). However, there was a significant difference in the mean number of years of education. This difference was controlled for statistically.

Method

All participants were interviewed using a semi-structured interview format in which they were encouraged to talk about themselves and their experiences. Participants were asked open ended questions, which did not restrict or control either the extent or the nature of their response. Responses were not corrected by the interviewer and no stimulus or interruption was provided unless the participant was clearly becoming distressed by his/her inability to respond. Clear interview guidelines were followed. Questions were asked slowly and repeated or reworded as necessary. Interviews were recorded with a tape recorder and a small microphone. Interviews lasted between 20 and 45 minutes allowing as much time as was needed to collect at least 1000 words of conversation from each participant. As recommended by Andreason and Pfohl (1976), each transcription consisted of approximately 1000 words (except in one DAT case where this was not possible). Interviews with controls took less time than those with patients.

Transcription

Interviews were transcribed according to agreed rules. Only words spoken by the participant were transcribed. As this was an exploration of the technique in DAT, initially all words were transcribed exactly as they had been spoken, including repetitions, incomplete words, interjections, paraphasias and mispronunciations.

Subsequently, a series of transcription guidelines were formulated and each transcript was reappraised according to these guidelines (see appendix). For each participant, transcription and lexical analysis were performed by the same researcher in order to maintain consistency. All transcriptions were recorded as ASCII text files for later analysis using the Oxford Concordance Program (OCP–Hockey and Martin 1988), a general purpose computer program designed to calculate word frequencies and to find the context in which a particular word occurs.

Stereotypical set phrases such as, 'you know', 'oh boy', 'to be honest', were excluded, because such expressions were not acceptable to us as proper clauses or full sentences, but had a significant impact on the measures used. Paraphasias and mispronunciations, however, were not deleted and were therefore counted in the analysis (e.g. 'dryven' instead of 'driven'). Words such as 'you've' and 'you have' were carefully transcribed separately: although transcribing these as a single word or two words does not affect the number of pronouns or verbs, it does alter the text length. Numbers were transcribed as words (e.g. 32 as thirty two (2 words)) and proper names were transcribed as single compound words (e.g. 'Home-and-Away'). Multiple attempts at the same word were only recorded once (e.g. 'I rig, ring the bell' was transcribed veridically, but scored as 'I ring the bell').

Linguistic measures

Eight linguistic measures were used. These measures had previously been used in combination in aphasia (Singh 1996, Holmes and Singh 1996), but had not been applied in dementia. They all depend on word frequencies of lexical items and can therefore be combined to generate a final index of performance which allows comparison between participants and comparison over time.

Word classes

The first four word measures were, pronoun rate (closed class) , noun rate (open class), adjective rate (open class) and verb rate (open class), all per 100 words. When the same word could be counted in a number of different ways (e.g. *that* can be a conjunction, a pronoun and an adjective–*the hill that I climbed*, *I like that* and *that man*), concordance analysis (using OCP) was carried out to determine the correct word category. Noun rate (N-rate) is a simple measure of ability to use nouns which it was felt would be sensitive to the word finding difficulties commonly experienced by people with DAT. Pronoun rate (P-rate) contrasts well with N-rate and was chosen to quantify the use of indirect referencing. Adjective rate (A-rate) has not been used in most studies involving spontaneous speech but it seemed to be an important variable in characterising the quality of participants' speech (its colour). Verb rate (V-rate) is a commonly used linguistic variable as it reflects the flow of speech. These measures have previously been used in studies of aphasia and speech in affective disorders (Fillenbaum *et al*. 1961, Jones *et al*. 1963, Andreason and Pfohl 1976).

Measures of richness of vocabulary

Three measures of the richness of vocabulary were used: type token ratio (TTR); Brunét's index (W) and Honoré's Statistic (R).

Type-token ratio (TTR)

TTR represents the ratio of the total vocabulary (V) to the overall text length (N). It is a simple measure of vocabulary size, which is generally found to correlate positively with the length of text sampled (N). This was one of the important reasons for attempting to standardize the length of text sampled from each participant at ~ 1000 words. Although TTR can be computed for different classes of lexical items (e.g. verbs, adjectives etc. see Singh and Bookless 1997), TTR was computed in this study for the whole transcription. Type-token ratios have been used by a number of researchers (see e.g. Andreason and Pfohl 1976), including Blanken *et al*. (1987) who reported their use in DAT.

Two further lexical richness measures were borrowed from stylometric studies of text (see Holmes, 1992; 1994) and are novel in their application to dementia. Singh (1996) and Holmes and Singh (1996) have previously applied these measures in aphasia.

Brunét's index (W)

Brunét's index was included because, unlike type-token ratio, it quantifies lexical richness without being sensitive to text length (Brunét 1978). It is calculated according to the following equation: $W = {}_{N}V^{(-.165)}$ where N is the total text length and V is the total vocabulary used by the participant. This measure generally varies between 10 and 20. The lower the value, the richer the speech.

Honoré's statistic (R)

Honoré's statistic (Honoré 1979), is based on the notion that the larger the number of words used by a speaker that occur only once (hapax legomena; Holmes 1992), the richer the lexicon. Words spoken only once (V_1) and the total vocabulary used (V) have been shown to be linearly associated. Honoré's Statistic generates a lexical richness measure that establishes the number of words used only once by the participant as a proportion of the total number of words used, according to the following formula: $R = 100 \log N/(1 - V_1/V)$ where N is the total text length. The higher the value of R, the richer the vocabulary used by the participant.

Measure of semantic cohesion in phrases–CSU rate

Finally, a measure of semantic cohesion in phrases was also included called clause like semantic unit (CSU) rate. Previous studies have quantified phrase length in a number of different ways (see e.g. Goodglass *et al*. 1964, Prins *et al*. 1978, Easterbrook *et al*. 1982 and Miceli *et al*. 1984). However, persons with dementia often leave clauses unfinished, making it difficult to apply a strict linguistic definition of a clause when analysing their speech. In order to quantify the participant's ability to cluster words together therefore, a CSU measure was used. This measure characterizes the participant's ability to form noun and verb phrases and gives an indication of the flow of speech (Singh 1996). A CSU can be defined as a string of words that are semantically connected in a meaningful unit. Singh (1996) has developed a series of rules for calculating CSUs and argued its superiority over established measures attempting a similar approach to sentence segmentation. An appendix at the end of this paper summarizes the rules used to segment sentences into CSUs (see also Singh and Bookless 1997). After marking each CSU on the printed transcript, the number of words in each CSU was counted. CSU rate per 100 words was then calculated. The higher the frequency of CSUs with large numbers of words in them, the lower the CSU rate. A lower CSU rate therefore indicates better performance. Normal speakers are capable of clustering as many as 20 words in a single CSU, and this ability has been found to be much reduced in aphasia (Holmes and Singh, 1996). Indeed, Holmes and Singh found that CSU-rate was the most important discriminator between normal and dysphasic speech in their sample.

Analysis

Each of the 8 linguistic measures was calculated for each participant and between groups comparisons of the measures carried out using Mann-Whitney U analysis to test the null hypothesis that both samples (DAT and healthy) came from the same distribution. This test requires continuous data but does not assume that it is normally distributed. Analyses of covariance was used to control for differences in years of education. Principal components analysis was conducted to identify underlying components that might explain the pattern of correlations within the set of measures. Discriminant function analysis was carried out to examine the predictive power of the measures. Statistical analyses were conducted using SPSS (Norusis 1997) and an alpha level of .05 was employed.

Results

For DAT participants, MMSE score, age, years of education and duration of illness, did not correlate significantly with any of the linguistic measures. Despite the fact that some correlation coefficients were 0.6 or above, small sample sizes prevented statistical significance being achieved. This suggests that these variables may have been sensitive to the degree of cognitive impairment (see table 2.). For healthy older participants, age did not correlate significantly with any of the linguistic measures (all coefficients >0.3). Despite the significant difference between the groups in terms of their years of education, only pronoun rate in HO controls was a significant correlate (at the 5% level). Analysis of covariance conducted for each of the 8 linguistic measures with years of education as a covariate, revealed that education was not a significant covariate of any of the measures, nor did it affect any of the group differences. Whilst more careful matching of years of education would have been preferable, previous research with normal adults has also shown that these measures were not significantly affected by educational level (Singh 1996). Thus, it follows that any differences found between the groups on the linguistic measures cannot solely be a function of differences in educational attainment, and must therefore reflect changes in the DAT participants due to their dementia. Table 3 contains the means, standard deviations, and ranges for each of the linguistic measures used.

There were significant differences between individuals with DAT and HO participants on all measures (N-rate, P-rate, Honoré's R, Brunét's W and TTR (all $p < .01$) A-rate, V-rate (both $p < .05$)) except CSU-rate. This small observed difference in CSU rate of 0.04 yielded a post-hoc power of 6% for this sample. Of

Table 2. Correlations among variables for participants with dementia of Alzheimer's type and healthy older participants

	MMSE	Age (years)	Years of education	Duration of illness (months)
DAT participants	n = 8			
Noun-rate	0.68	0.48	− 0.11	0.51
Pronoun-rate	0.22	− 0.17	0.42	− 0.58
Adjective-rate	− 0.32	0.06	− 0.32	− 0.33
Verb− rate	0.40	− 0.07	0.27	− 0.24
CSU− rate	− 0.63	0.40	0.30	− 0.06
Honoré's-R	0.42	0.19	0.43	0.26
Brunét's-W	− 0.31	− 0.70	− 0.16	− 0.41
Type Token Ratio	0.16	0.66	0.09	0.44
Healthy Older Controls	n = 16			
Noun-rate	–	− .22	0.0	–
Pronoun-rate	–	.05	− 0.51*	–
Adjective-rate	–	− .06	0.01	–
Verb-rate	–	.09	− 0.36	–
CSU− rate	–	.23	− 0.39	–
Honoré's-R	–	.09	0.38	–
Brunét's-W	–	− .23	− 0.13	–
Type Token Ratio	–	.25	0.22	–

* $p < .05$; MMSE: Mini-Mental State Examination; noun-rate, pronoun-rate, adjective-rate, verb-rate all per 100 words, CSU-rate: clause like semantic unit rate per 100 words, Honoré's-R: Honoré's lexical richness statistic, Brunét's-W: Brunét's lexical richness index.

Table 3. Linguistic measures for participants with dementia of Alzheimer type and healthy older participants

	Healthy older participants N = 16			Participants with DAT N = 8		
	M	SD	range	M	SD	range
Noun-rate	15.9	2.1	11.1–19.1	9.2	2.0	6.9–13.0
Pronoun-rate	15.8	1.8	13.2–19.0	23.7	4.1	14.3–27.2
Adjective-rate	7.2*	1.4	5.1–11.3	10.5	3.1	5.6–15.1
Verb-rate	21.0*	2.5	16.6–27.4	25.2	5.8	15.0–30.5
Type Token Ratio	0.32**	0.02	0.29–0.35	0.26	0.04	0.22–0.33
CSU-rate	15.2	2.2	11.0–17.9	15.6	1.4	13.5–17.3
Brunét's-W	14.4**	0.3	13.8–14.8	15.8	0.9	14.6–16.9
Honoré's-R	1613.8	136.6	1401.6–1863.7	1380.8	128.1	1140.9–1516.9

Mann-Whitney U Test, $p < .05$*, $p < .01$**; MMSE: Mini-Mental State Examination; Noun-rate, Pronoun-rate, Adjective-rate, Verb-rate all per 100 words, CSU-rate: clause like semantic unit rate per 100 words, Honoré's-R: Honoré's lexical richness statistic, Brunét's-W: Brunét's lexical richness index.

the four lexical measures (N, P, A and V–rate) DAT participants had *higher* mean P-rate, A-rate and V-rate scores, but *lower* N-rate scores compared with normal older controls. A closer inspection of the data revealed that DAT participants used pronouns 1.6 to 3.7 times as often as they used nouns in their conversation (mean 2.7, SD 0.8). Of those pronouns, between 29 and 54% of them were 'I' (mean 35%, SD8%). Normal controls used fewer pronouns than DAT participants (0.8

Table 4. Principal components analysis for healthy older participants and participants with dementia of Alzheimer yype

Measure	PC1 Loadings 'Lexical richness'	PC2 Loadings 'Sentence making ability'
Adjective-rate	0.66	− 0.43
CSU-rate	0.22	0.94
Noun-rate	− 0.83	− 0.05
Pronoun-rate	0.95	− 0.01
Honoré's-R	− 0.87	− 0.31
Type Token Ratio	− 0.93	− 0.06
Verb-rate	0.73	0.01
Brunét's − W	0.95	− 0.10

MMSE: Mini-Mental State Examination; noun-rate, pronoun-rate, adjective-rate, verb-rate all per 100 words, CSU-rate: clause like semantic unit rate per 100 words, Honoré's-R: Honoré's lexical richness statistic, Brunét's-W: Brunét's lexical richness index.

to 1.7 times the number of nouns, mean 1.0, SD 0.3) and this difference was statistically significant (Mann-Whitney U (corrected for ties); $U = 1.0$, $p = 0.0001$). Of these pronouns, between 11 and 51% were 'I' in normal participants (mean 3%, SD 1%). Calculating the number of 'I's as a proportion of all other pronouns, and comparing DAT and normal older participants' scores revealed a trend towards a greater proportion of 'I's in DAT participants, though the difference did not reach significance at the 5% level (Mann-Whitney U (corrected for ties); $U = 33.0$, $p = 0.0576$).

Of the three lexical richness measures, Brunét's index (W) showed the expected pattern of richer speech (*lower* mean values) in healthy older controls. Likewise, both Honoré's statistic (R) and TTR showed that most of the normal controls had lexically richer speech than DAT participants (*higher* mean values). Mean CSU-rates for these two groups did not differ, however, suggesting that whilst there were clear differences between DAT and healthy older participants in lexical items and in lexical richness measures, DAT participants did not differ significantly from normal older adults in their ability to form noun and verb phrases.

Principal Components Analysis

Principal Components Analysis (PCA) was conducted using all eight linguistic measures in order to reduce the dimensionality of the data set and to explore the possibility of fewer underlying variables or components, which might explain the data more efficiently. Two principal components were found with eigenvalues greater than 1, explaining 78.8% of the variance in the original data set (see table 4). Figure 1 shows Principal Component 1 (PC1) plotted against Principal Component 2 (PC2). The two groups, i.e. normal and DAT, are well separated on the PC1 axis, which explains 64.1% of the variance. According to the sign of the coefficients (positive or negative), PC1 contrasts N-rate, TTR, R and W with A-rate, P-rate and V-rate. This component appears to represent a lexical richness and phrase making factor, which was clearly different for the two groups. Most DAT participants scored positively on PC1 (mean 1.13, SD 0.84, range − 0.48 to 2.25),

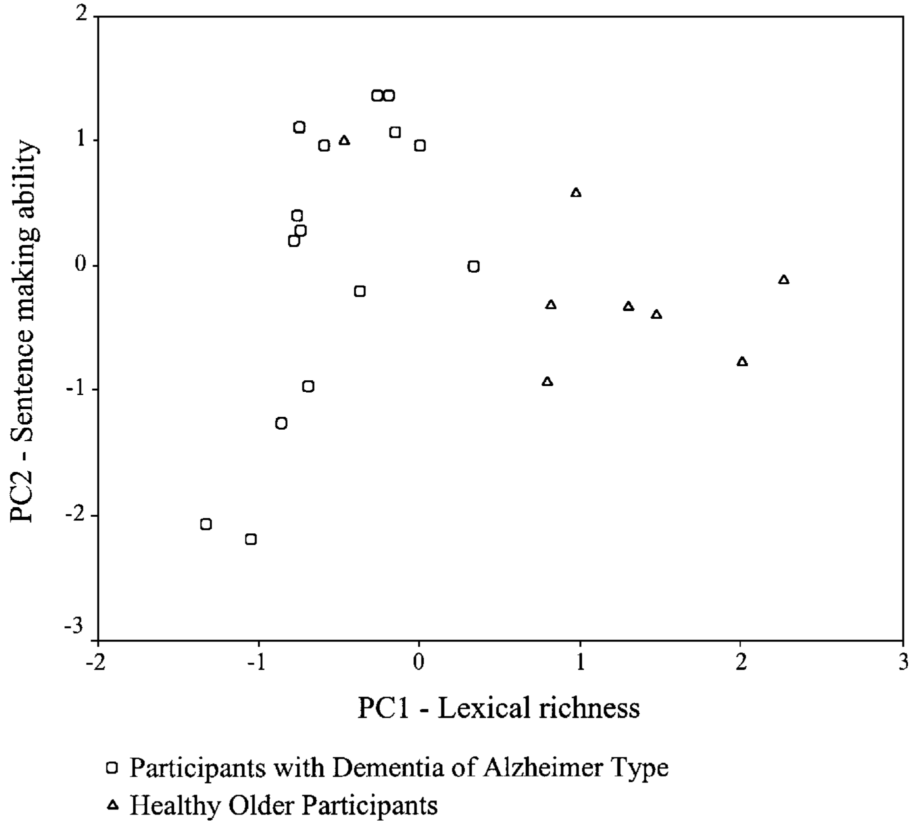

Figure 1. Principal components analysis: plot of PC1 'Lexical Richness' against PC2– 'Sentence Making Ability' for participants with dementia of Alzheimer type and healthy older participants.

whilst normal participants generally scored negatively (mean − .57, SD .42, range − 1.32 to 0.32). The PCA finding confirms that these two groups also differ in terms of A, P and V-rate. Principal component 2 (PC2) has a high positive loading only on CSU-rate. Adjective rate (A-rate), N-rate, P-rate, TTR, Honoré's statistic (R), Brunét's index (W) and V-rate contribute little to explaining further variance. PC2 account for an additional 14.7% of the variance. Healthy participants were more separated on this axis compared to DAT participants. This component appears to represent sentence making ability. Healthy participants differed more as a group, in their sentence making abilities than DAT participants (HO; mean 0.08, SD 1.15, range − 2.18 to 1.36: DAT; mean − 0.16, SD 0.65, range − .93 to 1.0). Since PC2 is dominated by CSU rate the similarity in PC2 scores between HO and DAT participants confirms the earlier finding that CSU-rate is not statistically different across the two groups.

Discriminant analysis

Linear discriminant analysis was carried out to establish how well the proposed linguistic measures predicted participant group membership. Discrimination across the two groups was excellent; group membership was correctly predicted for all 24

Table 5. Pooled within group correlations between discriminating variables and canonical discriminant functions (ordered by size) for healthy older and DAT participants

Measure	Discriminant Loadings
Noun-rate	− 0.69
Pronoun-rate	− 0.60
Brunét's-W	0.52
Type Token Ratio	− 0.41
Honoré's-R	− 0.37
Adjective-rate	0.33
Verb-rate	0.23
CSU-rate	0.04

MMSE: Mini-Mental State Examination; Noun-rate, Pronoun-rate, Adjective-rate, Verb-rate all per 100 words, CSU-rate: Clause like semantic unit rate per 100 words, Honoré's-R: Honoré's lexical richness statistic, Brunét's-W: Brunét's lexical richness index.

Table 6. Classification summary with cross-validation

		Predicted group membership	
Actual group	No. of cases	DAT	Healthy older participants
Participants with DAT	8	6 (75 %)	2 (25 %)
Healthy older Participants	16	1 (6.7 %)	15 (93.8 %)

N = 24, N correct = 21, Proportion correct = 87.5 %.

participants (100 % correct). The most important measures for discriminating between normal and DAT participants were N-rate, P-rate and Brunét's Index (W). Adjective rate, V-rate and CSU-rate correlated least well with the function and were therefore the least important discriminators between DAT and HO participants (see table 5). DAT participants generally had positive scores on the discriminant function, whilst HO participants had negative scores (DAT participants; mean 3.18, SD 1.3: normal participants; mean − 1.59, SD 0.8).

However, this analysis may have resulted in over-optimistic results since the discriminant function was tested on the same data that was used to produce it. Cross-validation (see Fu 1994) is a technique which can be used to avoid this circularity. The leave-one-out method of cross validation takes subsets of data for training and testing. It involves generating the discriminant function on all but one of the participants (n− 1) and then testing for group membership on that participant. The process is repeated for each participant (n times) and the percentage of correct classifications generated through averaging for the n trials. The classification summary following this method can be seen in table 6. Using this more rigorous classification technique, 87.5% correct classification was achieved. One HO and 2 DAT participants were misclassified. The HO participant was a 62 year old male with a small negative discrimination index score of −.59 partly due to a relatively low N-rate of 11.1. The AD participants were a 77 year old female with a MMSE of 21 and a 69 year old male with a low MMSE of 8. Both had low positive discriminant index scores of 1.45 and 1.47 respectively. The first of the two had a relatively high N-rate of 13.0, which actually fell in the range of performance from the healthy participants. The second of these AD participants only managed to record 532 text words due to a marked expressive language deficit. As Andreason and Pfohl (1976) and Singh (1996) have shown, the measures reach stability for speech samples of at least 1000 words. Thus, this misclassification reflects the importance of recording adequate text. These results strongly support those of the principal components analysis and the Mann-Whitney U tests, which suggest that the proposed linguistic measures discriminate well between DAT and normal older participants.

Discussion

This study has demonstrated that there are significant, objectively measurable lexical differences in the spontaneous, conversational speech of individuals with a diagnosis of probable dementia of Alzheimer type and healthy older participants. As expected, DAT participants produced significantly lower mean numbers of nouns per 100 words (N-rate), reflecting their anomia. Future analysis might usefully investigate whether this difference in noun rates applies to all classes of nouns or just to some (e.g. animate, inanimate, mass, count, abstract etc.). For example, a number of researchers have proposed that there is selective loss of semantic information regarding natural kinds, relative to man-made objects (see e.g. Garrard *et al*. 1998). Others have argued that this finding is an artefact of the overall naming difficulty of living and nonliving items (Tippett *et al*. 1996), or that familiarity facilitates word finding (see also Funnell and Sheridan 1992). Spontaneous speech may offer a method for exploring this debate.

Additionally, in this study DAT participants produced higher mean adjective, verb and pronoun rates than healthy older participants, though the differences between groups in adjective and verb rates were relatively small. Though these findings concur with those of Blanken *et al*. (1987), they are in contrast to Irigaray (1973) who found reduced adjective and verb rates in DAT participants by comparison with normal controls. Similarly, other researchers have found significantly poorer verb comprehension performance in AD by comparison with healthy controls (Robinson *et al*. 1996, Grossman *et al*. 1996) and in metalinguistic judgements about verbs (Kemper 1997). These differences in findings may reflect the different analyses used.

Pronoun rates were also significantly different; DAT participants used a mean of 23.7 pronouns per 100 words, whereas healthy older participants only used a mean of 15.8. A preliminary analysis of these transcripts revealed that there was a non significant trend for DAT participants to use 'I' more often than any other pronoun and more often than normal older participants, a finding reported elsewhere (e.g. Kempler and Zelinski 1994). One hypothesis for this is that DAT sufferers substitute pronouns for nouns because of word finding difficulties. A similar explanation of the higher incidence of pronouns in DAT participants can be posited to explain differences in adjective and verb rates. In any analysis calculating rates of words in conversation, which specifies the number of words to be sampled, a reduction in one class of word often results in an increase in another. Notwithstanding this, Nicholas *et al.* (1985) investigated the contribution of word finding difficulties to discourse incoherence in DAT. These authors found a significant negative correlation between scores of DAT participants on the BNT and the use of indefinite words (e.g. thing, stuff), but they found no significant correlation between the overuse of pronouns and test scores. Nor did they find any correlations between anomia and the use of pronouns without antecedents. They concluded that the naming deficit did not explain incoherent or empty discourse and that overuse of pronouns may be due to other impairments, such as problems with reference (Ulatowska *et al.* 1988).

As anticipated, healthy older adults produced richer speech on all three lexical richness measures (W-Brunét's, R-Honoré's and TTR). Concentrating on lexical richness measures for different classes of lexical units (e.g. nouns, verbs etc.) might also be a useful way of investigating the significant differences found between participant groups in lexical items (N, V, P and A-rate). Current research into the deficits associated with DAT has focused on the proposed breakdown of semantic memory. Whilst there is no doubt that in severe DAT many aspects of cognition are impaired, and this is likely to include semantic representations, the stage at which semantic memory begins to break down is contested (see e.g. Rossor and Hodges 1994, Martin 1992; Hodges and Patterson 1995 Vs Nebes and Brady 1988, Barr and Brandt 1996, Nicholas *et al.* 1996, Johnson *et al.* 1997). One reason for this may be variations in the demands each experimental task places on the participants (Nebes 1989, 1994a). AD patients' impaired performance may reflect an impaired ability to maintain all the information in working memory that is necessary for successful task performance; for example, for the successful ranking of attribute properties or for semantic fluency tasks (e.g. Smith *et al.* 1995, Johnson *et al.* 1997). Linguistic analysis of different classes of nouns used in spontaneous conversation might allow exploration of the relationship between lexical function and semantic memory breakdown, without the confounding effects of task difficulty and unnatural testing situations. Additionally, studying the effect on lexical richness and open class lexical items of varying the subject matter under discussion may also help elucidate the semantic memory breakdown debate.

In this study, Principal Components Analysis revealed two components. The first, a composite characterized by high positive loadings for A-rate, P-rate, V-rate and high negative loadings for N-rate, TTR, R and W, seems to represent the contrasting of lexical richness against phrase making or sentence making ability. A participant with a high PC1 score, would have problems with word finding and would show poverty of lexical richness, but would have no difficulty generating adjectives, pronouns or verbs. This pattern was characteristic of the DAT participants. The second PC (PC2) was characterized by high positive loading on CSU-rate alone. PC2 seems to represent the ability to make noun and verb phrases. Participants scoring highly on PC2 would score well in their ability to construct sentences. Indeed, participants from both groups demonstrated good scores on CSU-rate. Confirmation of the sensitivity of the linguistic measures in discriminating between DAT and normal participants, was provided by the discriminant function analysis. Even using cross-validation, this analysis resulted in 87.5% correct classification of participants, i.e. three participants were misclassified; one normal and two DAT. It is possible that the linear-discriminant method used may have given rise to these misclassifications since it is well known that linear methods have limitations whilst establishing function boundaries between different classes of non-linear data (Fu 1994).

Whilst test-retest reliability and sensitivity to change have been established in healthy participants (Singh 1996) the measurement of such reliability is still required in DAT. Further, it was beyond the scope of this study to consider inter-rater and intra-rater reliability. However, since 7 of the 8 measures are entirely objective, being generated either by computer or identified using dictionary definitions using a concordance programme which yielded the same results each time the analysis was run for each subject, we anticipate that the scope for inter and intra-rater error would be minimal. Finally, additional, normative data from a large sample of healthy older adults is required, in order to explore the sensitivity of these measures to very early language symptoms in DAT and in other dementing disorders such as Diffuse Lewy Body Disease (McKeith *et al*. 1996) and Vascular Dementia (Roman *et al*. 1993). Once a set of linguistic measures has been established further studies become possible.

One such study could explore the proposed gender differences in DAT patients, where more severe language deficits have been found in women, and women have been shown to decline faster than men in their language skills (Ripich *et al*. 1995, Henderson and Buckwalter 1994). The measures may be of help in further investigation, especially since they have shown significant gender differences in 30 healthy controls (Singh 1996). Longitudinal studies of changes in spontaneous speech that will help elucidate the progressive nature of the disorder also become possible. One of the major difficulties with longitudinal research is finding measures which can be repeated without contamination by practice and which will also be sensitive to deficits over the full range of the disease process: a problem experienced by Bayles *et al*. (1992) in her study to relate linguistic communication abilities to the stage of DAT. More severely demented patients are often untestable on formal language measures and there is also an ethical dilemma regarding the appropriateness of requiring them to attempt these sorts of assessments (Mohs *et al*. 1986, Post *et al*. 1994). Similarly, issues regarding repeated assessment of patients are also relevant for the evaluation of the effectiveness of anti-Alzheimer treatments, such as anti-cholinesterases.

In addition to using linguistic analysis of conversation to further understanding of dementia, these measures may also be used as a basis for developing new tests of language function. There is a need for clinical tools which can be administered more easily by psychologists and speech and language therapists, and which may help to improve diagnostic and prognostic accuracy. With the development of computer technology, new software is currently being designed to aid in the analysis of conversation, thus increasing the likelihood that linguistic measures will be used in clinical practice (see Singh 1996 for a discussion on computerised techniques for analysing speech transcripts). In addition, techniques for pattern classification, such as neural networks, may be used in the future when sufficient data of this type are available. These techniques may help in classifying different types of DAT patients based on conversational parameters; a development of work already underway in aphasia (Code *et al*. 1994, Singh 1997).

In summary, significant differences in the performance of DAT and healthy older participants have been demonstrated using linguistic measures of conversational, spontaneous speech. These measures discriminate well between participant groups, are objective, continuous and sensitive to differences in lexical richness and grammatic structure. An additional advantage of the method used in this study is that all measures are word-frequency dependent and can therefore be combined to yield a final index of performance (for example, using discriminant scores). This index of performance should lend itself well to use by clinicians. This study has identified the relative importance of variables in discriminating across DAT participants and normal controls, which is important to an understanding of qualitative deficits in DAT. We have also shown that variables such as CSU-rate and A-rate, considered most important in discriminating between dysphasic and normal speech (Holmes and Singh 1996), are less important in DAT. This suggests that the spontaneous speech of dysphasic and DAT patients differs qualitatively and that aphasia models should be used with caution in DAT studies. Further work is now necessary to explore additional measures of linguistic competence which may add to the sensitivity of the technique, and to investigate the relationship between these measures and other aspects of cognitive functioning such as semantic and episodic memory. Once a battery of measures has been established, opportunities for applying these to elucidate the nature of language breakdown in Alzheimer's disease and other dementias are obvious and exciting.

Acknowledgments

The authors wish to thank Jane Gilliard (School for Policy Studies, University of Bristol) for her assistance with the study, Anthony O. Hughes (Department of Social Medicine, University of Bristol) for his advice on statistical contents and Professor Pam Enderby (Speech and Language Research Unit, Frenchay NHS Trust) for her support. Completion of this study was facilitated by a grant from the Nuffield Provincial Hospitals Trust and The Charity, Bristol Research into Alzheimer's and Care of the Elderly. Finally, our thanks to the participants who gave their time for this study, and to the staff of The Bristol Memory Disorders Clinic for their assistance.

References

Alzheimer, A. 1907, Of a particular disease of the cerebral cortex. *Zentralblatt vur Nervenheilkunde und Psychiatrie*, **30**, 177–179.

American Psychiatric Association 1987, *Diagnostic and Statistical Manual of Mental Disorders—Revised* (Washington DC: American Psychiatric Association).

American Psychiatric Association 1994, *Diagnostic and Statistical Manual of Mental Disorders* (Washington DC: American Psychiatric Assocciation).

Andreason, N. J. and Pfohl, B. 1976, Linguistic analysis of speech in affective disorders. *Archives of General Psychiatry*, **33**, 1361–1367.

Appell, J., Kertesz, A. and Fisman, M. 1982, A study of language functioning in Alzheimer patients. *Brain and Language*, **17**, 73–91.

Barr, A. and Brandt, J. 1996, Word-list generation deficits in dementia. *Journal of Clinical and Experimental Neuropsychology*, **18**, 810–822.

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

Bayles, K. A., Kaszniak, A. W. and Tomoeda, C. K. 1987, *Communication and Cognition in Normal Ageing and Dementia* (Boston MA, USA: College-Hill Press/Little, Brown and Co.).

Bayles, K. A., Salmon, D. P., Tomoeda, C. K. and Jacobs, D. 1989, Semantic and letter category naming in Alzheimer's patients: a predictable difference. *Developmental Neuropsychology*, **54**, 335–347.

Bayles, K. A. 1991, Age at onset of Alzheimer's disease: relation to language dysfunction. *Archives of Neurology*, **48**, 155–159.

Bayles, K. A., Tomoeda, C. K. and Trosset, M. W. 1992, Relation of linguistic communication abilities of Alzheimer's patients to stage of disease. *Brain and Language*, **42**, 455–473.

Bayles, K. A., Tomoeda, C. K. and Trosset, M. W. 1993, Alzheimer's disease: effects on language. *Developmental Neuropsychology*, **9**, 131–160.

Becker, J. T., Huff, F. J., Nebes, R. D., Holland, A. L. and Boller, F. 1988, Neuropsychological function in Alzheimer's disease: pattern of impairment and rates of progression. *Archives of Neurology*, **45**, 263–268.

Binetti, G., Magni, E., Padovani, A. and Cappa, S. F. 1993, Neuropsychological heterogeneity in mild Alzheimer's disease. *Dementia*, **4**, 321–326.

Blanken, G., Dittman, J., Haas, J-C. and Wallesch, C-W. 1987, Spontaneous speech in senile dementia and aphasia. Implications for a neurolinguistic model of language production. *Cognition*, **27**, 247–274.

Boller, F., Becker, J. T., Holland, A. L., and Forbes, M. M. 1991, Predictors of decline in Alzheimer's disease. *Cortex*, **27**, 9–17.

Brunét, E. 1978, Le Vocabulaire de Jean Giraudoux. *Structure et Evolution* (Genéve: Slatkine).

Chui, H. C., Teng, E. L., Henderson, V. W. and Moy, A. C. 1985, Clinical subtypes of dementia of the Alzheimer type. *Neurology*, **35**, 1544–50.

Code, C., Rowley, D. and Kertesz, A. 1994, Predicting recovery from aphasia with connectionist networks: preliminary comparisons with multiple regression, *Cortex*, **30**, 527–532.

Crockford, C. and Lesser, R. 1994, Assessing functional communication in aphasia: clinical utility and time demands of three methods. *European Journal of Disorders of Communication*, **29**, 165–182.

Croisile, B., Ska, B., Brabant, M. J., Duchene, A., Lepage, Y., Aimard, G. and Trillet, M. 1996, Comparative study of oral and written picture description in patients with Alzheimer's disease. *Brain and Language*, **53**, 1–19.

Cummings, J. L., Benson, D. F., Hill, M. A. and Read, S. 1985, Aphasia in dementia of the Alzheimer type. *Neurology*, **29**, 315–323.

Dallal, G. E. 1990, PC-SIZE: consultant–a program for sample size determinations. *American Statistician*, **44**, 243.

De Santi, S., Koenig, L., Obler, L. K. and Goldberger, J. 1994, Cohesive devices and conversational discourse in Alzheimer's disease. In R. L. Bloom, L. K. Obler, S. De Santi and J. S. Erhlich (Eds) *Discourse Analysis and Applications: Studies in Adult Clinical Populations* (Hillsdale NJ, USA: Lawrence Erlbaum Associates).

Easterbrook, A., Brown, B. B. and Pereka, K. 1982, A comparison of the speech of adult aphasics in spontaneous and structured interactions. *British Journal of Disorders of Communication*, **17**, 93–106.

Fillenbaum, S., Jones, L. V. and Wepman, J. M. 1961, Some linguistic features of speech from aphasic speech. *Language and Speech*, **4**, 91–108.

Filley, C. M., Kelly, J. and Heaton, R. K. 1986, Neuropsychological features of early and late-onset Alzheimer's disease. *Annals of Neurology*, **23**, 365–70.

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.

Fu, L. 1994, *Neural Networks in Computer Intelligence* (Singapore: McGraw-Hill).

Funnell, E. and Sheridan, J. 1992, Categories of knowledge? Unfamiliar aspects of living and non-living things. *Cognitive Neuropsychology*, **9**, 135–153.

Garrard, P., Patterson, K., Watson, P. C. and Hodges, J. R., 1998, Category specific semantic loss in dementia of Alzheimer's type. Functional-anatomical correlations from cross-sectional analyses. *Brain*, **121**, 633–646.

Goldblum, M. C., Tzortzis, C., Michot, J. L., Panisset, M. and Boller, F. 1994, Language impairment and rate of cognitive decline in Alzheimer's disease. *Dementia*, **5**, 334–338.

Goodglass, H., Quadfasel, F.A. and Timberlake, W. H. 1964, Phrase length and type and severity of aphasia. *Cortex*, **1**, 133–153.

Grady, C. L., Haxby, J. V., Horwitz, B., Berg, G. and Rapoport, S. I. 1987, Neuropsychological and cerebral metabolic function in early vs. late onset dementia of the Alzheimer type. *Neuropsychologia*, **25**, 807–16.

Grossman, M., Mickanin, J., Onishi, K. and Hughes, E. 1996, Verb comprehension deficits in probable Alzheimer's disease. *Brain and Language*, **53**, 369–389.

Hachinski, V. C., Iliff, L. D., Zilhka, E., Du Boulay, G. H., McAllister, V. L., Marshall, J. Russell, R. W. R. and Symon, L. 1975, Cerebral blood flow in dementia. *Archives of Neurology*, 32, 632–637.

Henderson, V. W. and Buckwalter, J. G. 1994, Cognitive deficits of men and women with Alzheimer's disease. *Neurology*, **44**, 90–96.

Hier, D. B., Hagenlocker, K. and Shindler, A. G. 1985, Language disintegration in dementia: Effects of etiology and severity. *Brain and Language*, **25**, 117–123.

Hockey, S. and Martin, J. 1988, *OCP–User's Manual* (Oxford, UK: Oxford University Computing Service).

Hodges, J. and Patterson, K.. 1995, Is semantic memory consistently impaired early in the course of Alzheimer's disease? Neuroanatomical and diagnostic implications. *Neuropsychologia*, **33**, 441–459.

Hodges, J. R., Salmon, D. P. and Butters, N. 1991, The nature of the naming deficit in Alzheimer's and Huntington's disease. *Brain*, **114**, 1547–1558.

Holmes, D. I. 1992, A stylometric analysis of mormon scripture and related texts. *Journal of the Royal Statistical Society (A)*, **155**, 91–120.

Holmes, D. I. 1994, Authorship attribution. Computers and the Humanities, **28**, 87–106.

Holmes, D. I. and Singh, S. 1996, A stylometric analysis of conversational speech of aphasic patients. *Literary and Linguistic Computing*, **11**, 45–60.

Honoré, A. 1979, Some Simple Measures of Richness of Vocabulary. *Association of Literary and Linguistic Computing Bulletin*, **7**, 172–177.

Hutchinson, J. M. and Jensen, M. 1980, A pragmatic evaluation of discourse communication in normal and senile older in a nursing home. In L. Obler and M. Albert (Eds) *Language and Communication in the Older* (Boston DC: Heath and Co.).

Irigaray, L. 1973, *Le language des dements* (The Hague: Mouton).

Jones, L. V., Goodman, L. F. and Wepman, J. M. 1963, The classification of parts of speech for the characterization of aphasia. *Language and Speech*, **6**, 94–107

Johnson, M. K., Bonilla, J. L. and Hermann, A. M. 1997, Effects of relatedness and number of distractors on attribute judgements in Alzheimer's disease. *Neuropsychology*, **11**, 392–399.

Kemper, S. 1997, Metalinguistic judgements in normal aging and Alzheimer's disease. *Journals of Gerontology, Series B, Psychological Science and Social*, **3**, 147–155.

Kempler, D. 1991, Language changes in dementia of the Alzheimer type. In R. Lubinski (Ed.) *Dementia and Communication* (Philadelphia: B. C. Decker, Inc.).

Kempler, D., Andersen, E. S. and Henderson, V. W. 1995, Linguistic and attentional contributions to anomia in Alzheimer's disease. *Journal of Neuropsychiatry, Neuropsychology and Behavioural Neurology*, **8**, 33–37.

Kempler, D. and Zelinski, E. M. 1994, Language in dementia and normal aging. In F. A. Huppert, C. Brayne and D. W. O'Connor (Eds.) *Dementia and normal aging* (Cambridge: Cambridge University Press).

Kertesz, A., Appell, J. and Fisman, M. 1986, The dissolution of language in Alzheimer's disease. *Canadian Journal of Neurological Sciences*, **13**, 415–418.

Koss, E., Edland, S., Fillenbaum, G., Mohs, R. C., Clark, C., Galasko, D., *et al.* 1996, Clinical and neuropsychological differences between patients with earlier and later onset of Alzheimer's disease. Part XII. A CERDAT analysis. *Neurology*, **46**, 136--141.

Lesser, R. and Milroy, L. 1993, *Linguistics and Aphasia* (London: Longman).

Martin, A. 1992, Semantic knowledge in patients with Alzheimer's disease: evidence for degraded representations. In L. Backman (Ed.) *Memory functioning in dementia* (Amsterdam: North-Holland).

Martin, A., Brouwers, P., Lalonde, F., Cox, C., Teleska, F. and Fedio, F. 1985, Towards a behavioural typology of Alzheimer's patients. *Neurology*, **35**, 394–7.

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.

McKhann, G., Drachman, D., Folstein, M. F., Katzman, R., Price, D. and Stadlan, E. M. 1984, Clinical diagnosis of Alzheimer's disease: report of the NINCDS-DATRDA work group under the auspices of Department of Health and Human Services task force on Alzheimer's Disease. *Neurology*, **34**, 939–944.

McKeith, I. G., Galasko, D., Kosaka, K., Perry, E. K., Dickson, D. W., Hansen, L. A., et al. 1996, Consensus guidelines for the clinical and pathological diagnosis of dementia with Lewy bodies (DLB): report of the consortium on DLB international workshop. *Neurology*, 47, 1113--1124.

Miceli, G., Silveri, M. C., Villa, G. and Caramassa, A. 1984, On the basis for the agrammatic difficulty in producing main verbs. *Cortex*, **20**, 207–220.

Mohs, R. C., Kim, Y., Johns, C. A., Dunn, D.D. and Davis, K. L. 1986, Assessing changes in Alzheimer's disease: memory and language. In L. Poon, T. Crook, B. J. Gurland, K. L. Davis, A. W. Kaszniak, C. Eisdorfer, and L. W. Thompson (Eds.) *Clinical Memory Assessment of Older adults*. Washington: APA.

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

Nebes, R. D. 1994, Semantic dysfunction in Alzheimer's disease. In F. I. M. Craik and T. A. Salthouse (Eds). *The Handbook of Ageing and Cognition* (Hillsdale, NJ: Lawrence Erlbaum.

Nebes, R. D., and Brady, C. B. 1988, Integrity of semantic fields in Alzheimer's disease. *Cortex*, **25**, 305–315.

Nicholas, M., Obler, L., Albert, M. and Helm-Estabrooks, N. 1985. Empty speech in Alzheimer's disease and fluent aphasia. *Journal of Speech and Hearing Research*, **28**, 405–10.

Nicholas, M., Obler, L. K., Au, R. and Albert, M.L. 1996, On the nature of the naming errors in ageing and dementia: a study of semantic relatedness. *Brain and Language*, **54**, 184–195.

Norusis, M.J. 1997, *SPSS for Windows: Advanced Statistics* (Chicago: SPSS Inc.).

Obler, L. K., and Albert, M. L. 1981, Language in the older aphasic and in the dementing patient. In M. T. Sarno (Ed.) *Acquired aphasia* (New York: Academic Press).

Obler, L. K. 1983, Language and brain dysfunction in dementia. In S. Segalowitz (Ed.) *Language functions and brain organization* (New York: Academic Press).

Phillips, L. H., Della Sala, S. and Trivelli, C. 1996, Fluency deficits in patients with Alzheimer's disease and frontal lobe lesions. *European Journal of Neurology*, **3**, 102–108.

Post, S. G., Ripich, D. N. and Whitehouse, P. J. 1994, Discourse ethics: research, dementia, and communication. *Alzheimer Disease and Associated Disorders*, 8 Suppl. **4**, 58–65.

Prins, R. S., Snow, C. and Wagenaar, E. 1978, Recovery from aphasia: Spontaneous speech versus language comprehension. *Brain and Language*, **6**, 192–211.

Ripich, D. N. and Terrell, B. Y. 1988, Cohesion and coherence in Alzheimer's disease. *Journal of Speech and Hearing Disorders*, **53**, 8–14.

Ripich, D. N., Vertes, D., Whitehouse, P., Fulton, S. and Ekelman, B. 1991, Turn taking and speech act patterns in the discourse of senile dementia of the Alzheimer's type patients. *Brain and Language*, **40**, 330–43.

Ripich, D. N., Petrill, S., Whitehouse, P.J. and Ziol, E.W. 1995, Gender differences in language of DAT patients: a longitudinal study. *Neurology*, **45**, 299–302.

Robinson, K. M., Grossman, M., White-Devine, T. and D'Esposito, M. 1996, Category-specific difficulty with naming verbs in Alzheimer's disease. *Neurology*, **47**, 178–182.

Roman, G. C., Tatemichi, T. K., Erkinjuntti, T., Cummings, J. L., Masdeu, J. C., Garcia, J. H. *et al*. 1993, Vascular dementia: diagnostic criteria for research studies. Report of the NINDS-AIREN International Workshop. *Neurology*, **43**, 250--260.

Romero, B. and Kurz, A. 1996, Deterioration in spontaneous speech in AD patients during a 1-year follow up: Homogeneity of profiles and factors associated with progression. *Dementia*, **7**, 35–40.

Rosser, A. and Hodges, J. R. 1994, Initial letter and semantic category fluency in Alzheimer's disease, Huntington's disease, and progressive supranuclear palsy. *Journal of Neurology, Neurosurgery, and Psychiatry*, **57**, 1389–1394.

Sabat, S. R. 1994, Language function in Alzheimer's disease: a critical review of selected literature. *Language and Communication*, **14**, 331–351.

Schwartz, M. F., Marin, O. S. M. and Saffran, E. M. 1979, Dissociations of language function in dementia: a case study. *Brain and Language*, **7**, 277–306.

Selnes, O. A., Carson, K., Rovner, B. and Gordon, B. 1988, Language dysfunction in early and late-onset possible Alzheimer's disease. *Neurology*, **38**, 1053–56.

Seltzer, B. and Sherwin, I. (1983). A comparison of clinical features in early- and late-onset primary degenerative dementia. *Archives of Neurology*, **40**, 143–146.

Sevush, S., Leve, N., and Brickman, A. 1993, Age at disease onset and pattern of cognitive impairment in probable Alzheimer's disease. Journal of Neuropsychiatry and Clinical Neurosciences, **5**, 66–72.

Singh, S. 1994, Linguistic computing in speech and language disorders, *Proceedings of the 5th International Conference on Speech Science and Technology*, Dec. 5–8, Perth, 486–491.

Singh, S. 1995, Computational linguistics for analysing conversation in speech and language disorders. *Proceedings of the 3rd International Conference on Statistical Analysis of Textual Data*, Rome, 11–13 December, 355–362.

Singh, S. 1996, *Computational analysis of conversational speech in dysphasic patients*. PhD thesis. University of the West of England, UK.

Singh, S. 1997, Quantitative classification of conversational language using artificial neural networks. *Aphasiology*, **11**, 829–844.

Singh, S. and Bookless, T. 1997, Analyzing spontaneous speech in dysphasic adults. *International Journal of Applied Linguistics*, **7**, 165–182.

Smith, S., Faust, M., Beeman, M., Kennedy, L. and Perry, D. 1995, A property level analysis of lexical semantic representation in Alzheimer's disease. *Brain and Language*, **49**, 263–279.

Tippett, L. J., Grossman, M. and Farah, M. J. 1996, The semantic memory impairment of Alzheimer's disease: Category specific? *Cortex*, **32**, 143–153.

Thompson, I. M. 1987, Language in dementia. *International Journal of Geriatric Psychiatry*, **2**, 145–161.

Ulatowska, H. K. and Chapman, S. B. 1991, Discourse studies. In R. Lubinski (Ed.) *Dementia and communication* (Philadelphia: B. D. Decker, Inc.).

Ulatowska, H., Allard, L., Donnell, A., Bristow, J., Haynes, S. M., Flower, A. and North, A. J. 1988, Discourse performance in participants with dementia of the Alzheimer type. In H. Whitaker (Ed.) *Neuropsychological studies in non focal brain damage* (New York: Springer-Verlag), pp. 108–31.

Appendix: CSU computation rules summary

Rule 1. Most of the CSUs will have boundaries at conjunction words such as *but*, *and*, *since*, *so*, *because*, etc.

example: 'I went to the market where I met my friend / but I didn't recognize him at first / I was hungry / and we had the money / so we went to a restaurant / '.

Rule 2. As long as the segment is cohesive, do not divide it.

example: 'I went to this place when I knew that I was not getting a job / but then there was no option /and I had to go for it /no matter what would happen / '.

Rule 3. Include incomplete segments as long as they are syntactically well-formed.

example: 'I was new in there / but I worked for / I liked it / and there was __ / '

Rule 4. When a number of objects or names are uttered separated by *and*, do not segment.

example: 'It was Tom and David / and they had a spade and a shovel / '.

Rule 5. When in above (rule 4) the words are not separated by *and*, segment them individually.

example: 'It was hot / humid / rainy / '.

Rule 6. Do not include words such as *Well*, *Yes*, *No*, *Blimey*, etc. in the count.

example: '[Well Yes] actually I did like it / [Oh God] / I loved it / '.

Rule 7. Do not count the same word if it appears in succession.

example: 'We [we we] did go there / but [but] we didn't like it / '.

Rule 8. If the participant rephrases during a sentence, take out the words that are not in logical sequence.

example: '[I was] She was [much sick no] much tired after [the holiday no no] the work / '.
Rule 9. As in rule 8, especially for numbers and dates, accept only the last one.
example: 'I have [one two three four Oh God] three children / '.
Rule 10. Ignore automatisms.
example: 'I have [one one] [Oh dear dear dear dear] one child / [you know]'
Rule 11. Automatisms that are actually a part of the sentence, where the sentence becomes meaningless without them, should not be taken off.
example: 'I hope *you know* what I mean / '.
Rule 12. Include greetings and single word responses.
example: 'Thank you / Good / Very Good / '.
Rule 13. Treat many words, which are part of a single name, as one entity.
example: 'I like Coronation-Street / Emmerdale-Farm / Home-and-Away / '.