The effects of very early Alzheimer's disease on the characteristics of writing by a renowned author

Peter Garrard,[1] Lisa M. Maloney,[2] John R. Hodges[3] and Karalyn Patterson[3]

[1]*Institute of Cognitive Neuroscience, London,* [2]*Defence Services Medical Rehabilitation Unit, Headley Court, Epsom, Surrey and* [3]*MRC Cognition and Brain Science Unit, Cambridge, UK*

Correspondence to: Dr. P. Garrard, Institute of Cognitive Neuroscience, 17 Queen Square, London WC1N 3AR. UK. E-mail: p.garrard@ucl.ac.uk and garrard@cnbc.cmu.edu

Summary

Iris Murdoch (I.M.) was among the most celebrated British writers of the post-war era. Her final novel, however, received a less than enthusiastic critical response on its publication in 1995. Not long afterwards, I.M. began to show signs of insidious cognitive decline, and received a diagnosis of Alzheimer's disease, which was confirmed histologically after her death in 1999. Anecdotal evidence, as well as the natural history of the condition, would suggest that the changes of Alzheimer's disease were already established in I.M. while she was writing her final work. The end product was unlikely, however, to have been influenced by the compensatory use of dictionaries or thesauri, let alone by later editorial interference. These facts present a unique opportunity to examine the effects of the early stages of Alzheimer's disease on spontaneous written output from an individual with exceptional expertise in this area. Techniques of automated textual analysis were used to obtain detailed comparisons among three of her novels: her first published work, a work written during the prime of her creative life and the final novel. Whilst there were few disparities at the levels of overall structure and syntax, measures of lexical diversity and the lexical characteristics of these three texts varied markedly and in a consistent fashion. This unique set of findings is discussed in the context of the debate as to whether syntax and semantics decline separately or in parallel in patients with Alzheimer's disease.

Introduction

The neuropsychological deficits attributable to neocortical involvement in the early stages of Alzheimer's disease have been documented extensively, allowing important clinical subgroups to be identified. These include patients with deficits in early visual processing, spatial skills and various components of language, implicating occipital, bi-parietal and left temporal regions, respectively (Galton *et al.*, 2000). The study of patients with such relatively modular deficits continues to contribute to our understanding of functional organization within a range of cognitive domains.

Language provides one of the best examples of how the study of early Alzheimer's disease has informed cognitive neuropsychology, with the analysis of individuals and groups showing disruption of word and sentence level production and comprehension, and disintegration of the vast, structured network of information that endows objects and words with meaning, i.e. semantic memory (Nebes, 1989). Many of these studies have employed assessment techniques based on standardized tasks such as word fluency and picture naming, providing evidence for the influence of lexical frequency, familiarity and age of acquisition, and other psycholinguistic variables, on performance.

Whilst these techniques are reliable, they have been criticized for their lack of 'real-world' validity. In response, a number of studies have provided quantitative analyses of more naturalistic activities such as the production of connected speech, either spontaneously or in response to a specific prompt (e.g. description of a complex scene, or narration of a well-known story). One research group was able to carry out this type of analysis longitudinally, spanning periods of ⩾50 years to include data from before and after the appearance of the earliest symptoms of Alzheimer's disease. The study in question was made possible by the cooperation of members of an enclosed religious community, who enrolled into a prospective study of incident dementia, and made available an archive of written documents dating back to the time

of each subject's entry into the order (Snowdon *et al.*, 2000). Samples of the written output of individuals at ages between 18 and 32 years were compared with samples 50 or 60 years later. The comparison revealed changes in syntactic complexity and semantic content in all subjects; but, of more interest, subgroup analysis identified reliable differences between the initial scores of individuals who went on to develop dementia and those whose cognitive abilities subsequently remained intact. The latter finding was interpreted as evidence for the influence of neurodevelopmental or other early determinants of cognitive capacity on the risks of developing Alzheimer's disease in later life.

The possibility of examining, retrospectively, the products of cognitive operations at work during the pre-symptomatic period of a neurodegenerative disease occurs rarely but, as the above example illustrates, may provide information of exceptional interest. The value of such observations stems principally from the facts (i) that the task in question is undertaken voluntarily and presumably comes naturally to the subject; (ii) that the subject is unaware of the incipient disease, which eliminates negative emotional and compensatory strategic effects on performance; and (iii) that the availability of material from periods both before and after disease onset provides the opportunity for within-patient comparison.

An opportunity of this kind has presented itself recently via the writings of the novelist and philosopher Dame Iris Murdoch (I.M.), in whom Alzheimer's disease was diagnosed at the age of 76 years, not long after she finished writing the novel *Jackson's Dilemma,* her final published work. The clinical diagnosis was confirmed at post-mortem 3 years later, and anecdotal evidence (presented below) leaves little doubt that cognitive impairment was present in the months, or more probably years, leading up to the diagnosis. This implies that the composition of her final written text was taking place at a time when Alzheimer pathology was already beginning to disrupt her cognitive abilities. Although this fact was unrecognized at the time, there are clear indications from the reviews of this final book that critics and experts found it disappointingly different from her earlier works. A. S. Byatt, for instance, compares the structure of the novel to '... an Indian Rope Trick ... in which all the people have no selves and therefore there is no story and no novel', while Penelope Fitzgerald noted that the economy of the writing made it appear '... as though Murdoch had let her fiction wear through almost to transparency'. Other critics were less oblique: Kate Kellaway found *Jackson's Dilemma* '... not a perfect novel: the narrative itself is, at times, a little distrait: like Jackson [a central character], it often moves with scant explanation', and Hugo Barnacle described it as reading '...like the work of a 13-year-old schoolgirl who doesn't get out enough'. [All extracts taken from Porlock, 1995.]

Any piece of continuous discourse, whether written or spoken, is clearly a highly complex structure, and thus open to analysis at many different levels. In the case of *Jackson's Dilemma*, early cognitive deterioration might have affected any (or all) of these levels sufficiently to provoke

the uncharacteristically lukewarm critical responses just cited. At the most global level, the development and organization of the ideas and episodes of the book, as well as aspects of the literary 'style', might be assessed, though it is hard to see how either might be achieved in any objective or quantitative way. The more fine-grained the focus of the description, however, the more tractable it becomes. The Nun Study cited above, for instance, used measures of grammatical complexity and 'idea density' to quantify both syntactic and semantic content of a selection of written sentences, and Crosile *et al.* (1996) compared the Cookie Theft descriptions given by Alzheimer's disease patients in oral and written modalities. Both studies reported reduced syntactic complexity in patients compared with age-matched controls, though it remains unclear whether this finding reflects a specific linguistic deficit, diminished working memory capacity or some more general resource deficit (Grossman and Rhee, 2001). Spoken narratives were also used to analyse the lexical characteristics of the language of a group of patients with isolated progressive semantic impairment (semantic dementia), and suggested not only that the number of words from lower frequency bands used by the patients diminished with progression of the semantic deficit, but that this tended progressively to favour the production of verbs over nouns and of lower over higher imageability words—findings that accord with clinical observations in this group of patients (Bird *et al.*, 2000).

To the extent that difficulties with name production in Alzheimer's disease generally reflect disruption of the semantic system (see, for example, Hodges *et al.*, 1992; Garrard *et al.*, 2004), picture-naming performance in this patient group might be expected to be influenced by the same factors that affect word production in semantic dementia. Surprisingly, however, there have been few large-scale population-based studies of the influence of these variables *per se* on naming in this patient group. Kremin *et al.* (2001) found, in a multiple regression analysis on a small group of Alzheimer's disease patients, that age of acquisition, name agreement, frequency, familiarity and semantic category significantly predicted picture naming, but that word length had no effect. Astell and Harley (1996) showed that tip-of-the-tongue states occurred more often in Alzheimer's disease when the targets were words of low frequency and imageability.

To the best of our knowledge, no similar lexical level analyses have been carried out on bodies of written texts in Alzheimer's disease. The main question guiding the textual analyses described in this study was as follows: do the psycholinguistic properties of the vocabulary used in *Jackson's Dilemma*—which, we argue, was written during the early stages of Alzheimer's disease—differ significantly from those of the vocabulary used in books written at earlier stages of Murdoch's literary career? In addition to lexical factors, we examined the texts for evidence of differences in a number of additional features, including overall structure, extent of shared vocabulary and grammatical class (comparative

frequency of different word types). We also looked for differences in syntactic complexity, by comparing sentence lengths and density of subordinate clauses, both of which have been found to correlate with ratings of syntactic complexity derived from time of acquisition during language development (Cheung and Kemper, 1992).

Before presenting the findings of these analyses, we should emphasize that, throughout her writing career, I.M. regarded the manuscripts that she sent to her publisher as representing her work in its finished form, and resisted any suggestions of alterations to the text (Bayley, 1998; Conradi, 2001; Wilson, 2003). As a consequence, it seems reasonable to assume that any differences identified between the books analysed here can reliably be attributed to the author herself rather than to any later editorial interference.

Case report

Jean Iris Murdoch was one of the most acclaimed writers of the 20th century. Born into a middle-class family in 1919, she was educated at Badminton School and then Somerville College Oxford, where she took a first in Greats in 1942. Following a wartime spell in the Civil Service, and a post-war attachment to the United Nations, she turned to philosophy, first at Cambridge and later Oxford, where she became Fellow and Tutor at St Anne's College in 1948.

I.M. wrote several works of fiction, most of which she later destroyed, before the publication of her first novel, *Under the Net*—a first-person narrative comically chronicling the adventures of a down-at-heel writer/philosopher in post-war London. Her reputation as both novelist and philosopher grew steadily, and in 1963 she retired from teaching and devoted the rest of her life to writing. I.M.'s approach to writing was highly individual: she would carefully work out characters and plots for up to 8 months before spending 6 months writing out the book in longhand (she never used a typewriter, let alone a word processor). There is no evidence that she agonized over choice of words, indulged in repeated revisions of passages, or made extensive use of a dictionary or thesaurus. Work in progress would occupy the pages of a large pad, which she carried around and added to whenever and wherever she found herself unoccupied (Wilson, 2003). Her publishers would be sent longhand manuscripts (and often complained that they could not read her handwriting) but she eschewed any editorial interference (Conradi, 2001).

Her work attracted outstanding critical acclaim throughout her life: in 1978, she was awarded the Booker prize for *The Sea, The Sea*, and in 1987 was created Dame Commander of the British Empire in recognition of her contribution to British literary life. Her final novel, *Jackson's Dilemma*—a story about the lives and love affairs of a group of friends, which has the manservant Jackson as a shadowy, behind-the-scenes protagonist—was published in 1995. *Jackson's Dilemma* was received 'respectfully' but without enthusiasm, and I.M. would later reveal that she had been dogged by an intense and distressing 'writer's block' while working on it.

A year earlier she had become uncharacteristically inarticulate while taking part in an unscripted question-and-answer session about her work at a conference (Bayley, 1998), and diary entries from 1993 are noted by her biographer as being reduced to 'heart-rending simplicity' (Conradi, 2003).

The progressive nature of her condition soon became obvious, and she underwent neurological and neuropsychological assessment between November 1996 and June 1997. During this interval, her condition deteriorated markedly: in early November 1996, she achieved 20/30 on a Folstein Mini-Mental State Examination (MMSE), with points lost mainly on tests of attention and orientation. By the following summer, her MMSE score had dropped to 10/30: she could identify her surroundings only by the name of the city (Oxford), was unable to perform serial subtraction or reversed spelling, and could register but not recall a three-item word list.

Neuropsychological evaluation

Table 1 documents I.M.'s scores on a range of neuropsychological tests administered (over three assessment occasions) during the early period of decline. Particularly noteworthy are her performances on picture naming, in which she produced mainly circumlocutory errors (BUS: 'something carried along'; SLEDGE: 'thing to be going up and down'; KANGAROO: 'beautiful creature that jumps') and a test of spelling, on which she produced numerous regularization errors (e.g. 'cruise' →crewse and 'kaleidoscope' →colidascope).

Anterograde episodic memory and, to a lesser extent, visuospatial skills were also deficient from an early stage. Her retrograde/semantic memory, in the form of knowledge of famous people, was profoundly impaired; in this test, the participant is shown a random series of famous and unknown faces and asked first whether the person is famous and then (if the reply is 'yes') for his/her name or, failing the name, for specific identifying information ('what is he/she famous for?'). I.M.'s personal autobiographical/semantic memory was tested by two assessments specially constructed by our research group to probe her knowledge of her own books. In the first of these, she was shown the printed names of two real book titles, one by her and another by a different 20th century author, and asked which one she had written; in the second, she was shown the title of one of her books together with the names of two of her principal fictional characters, one from the target book and the other from a different one of her own books, and asked which character appeared in the target book. Her score was well above chance for the first of these tests (though pre-morbid performance would presumably have been errorless) but, on the second test, she was very slow to make her choices and scored rather poorly; her husband's score on the same test was better and achieved much more rapidly.

Short-term memory (as measured by digit span) was mildly impaired. Connected narrative speech (description of the 'Cookie Theft' scene from the Boston Diagnostic Aphasia Examination), although mostly grammatically well formed,

Table 1 *Neuropsychological test results from early and middle stages of IM's illness*

Domain	Test	Score	Normative values (mean ± SD)	When tested
General	MMSE	20/30	28.8 ± 0.5	November 1996
		10/30		June 1997
Working memory	Digit span	Forwards: 4	6.8 ± 0.9	November 1996
		Backwards: 2	4.7 ± 1.2	
Anterograde episodic memory	Logical memory*	Story 1: 0.5/24	<5th percentile	
		Story 2: 1/23		
Autobiographical memory	Title recognition[+]	25/29	29/29 (husband)	April 1997
	Character–book matching[+]	12/18	17/18 (husband)	
Visuoperceptual skills	VOSP[†]	Incomplete letters: 17/20	19.2 ± 0.8	November 1996
		Object decision: 11/20	17.7 ± 1.9	
		Dot counting: 10/10	9.9 ± 0.3	
		Position discrimination: 17/20	19.8 ± 0.6	
	Rey figure copy	22/36	29.6 ± 3.4	
Language and semantics	Graded naming test	0/30	<5th percentile	September 1997
	Pyramids and Palm Trees	Pictures: 41/52	51.2 ± 1.4	
	Cambridge semantic battery[‡]	Naming: 30/64	62.8 ± 1.2	
		Word–picture matching: 54/64	64 ± 0	
	Verbal fluency	FAS: 20 (excluding 6 intrusion errors)	34.8 ± 12.8	
		Category[§]: 3 (excluding 5 perseverative and 3 intrusion errors)	31.9 ± 19.9	
	Reading	Regular: 124/126	NA	
		Irregular: 119/126		
		Non-words: 39/40		
	Graded spelling test	17/30		

*From the Wechsler memory scale; [+] see text for description; [†] Visual Object and Space Perception Battery (Warrington and James, 1991); [‡] Bozeat *et al*. (2000); [§] Total of animals, water creatures, birds and breeds of dog (1 min per category). MMSE = Mini-Mental State Examination.

was piecemeal, circumlocutory, and clearly deficient in content words:

> The girl is just holding a plate and various pieces of . . . well . . . something useful . . . standing at a window . . . whether the window is open is not quite clear to me. The thing where the water is running out. The girl doesn't bother. The window is open. Plate and two cups.

An MRI brain scan was performed in June 1997 (see Fig. 1), and showed changes of global atrophy with profound bilateral hippocampal shrinkage and a suggestion of temporal lobe predominance. I.M. died in February 1999, aged 79 years. The presumptive diagnosis of Alzheimer's disease was confirmed at post-mortem, graded Braak stage 4–5. Microscopical examination demonstrated features of Alzheimer's disease throughout the brain, with disproportionate deposition of plaques and tangles, as well as gliosis and spongiosis, in temporal regions.

Material and methods

The complete texts of I.M.'s first and final published works—*Under the Net* (1954) and *Jackson's Dilemma* (1995)—were converted to digital format using commercial optical character recognition software. The same procedure was performed on the first 100 pages of *The Sea, The Sea* (1978), a book written at the height of her literary career, and widely regarded as one of her most accomplished works. All formatting, other than line and paragraph breaks, was removed, and passages consisting of dialogue or direct quotation were identified and marked with special characters.

Concordances of two types were constructed from the resulting text files: (i) Concordance software (Watt, 2002) was used to transform the complete texts of all three works into alphabetical word lists, showing the frequency of each word by type, and displaying the context of every occurrence; and (ii) the same software was configured to select random samples of 100 words from each book, to produce similar (though much shorter) alphabetical lists.

Procedure (ii) was conducted five times with each text, giving 15 word lists, from which the few proper nouns that occurred were deleted. Because Concordance samples word tokens rather than unique word types, these lists contained <100 different words: across the five lists, the numbers of word types selected totalled 379 for *Under the Net,* 373 for *The Sea, The Sea* and 352 for *Jackson's Dilemma*. All listed words were classified as nouns, verbs, descriptors (adjectives and adverbs) or function words (e.g. conjunctions and pronouns). In the very few cases of grammatical class ambiguity (e.g. <HAND> which can be used as either a noun or a verb), the more typical reading was assumed (in the example, the word was classified as a noun). For each word in the list, word length was marked, and values of lexical variables of interest (frequency, familiarity, imageability) retrieved, if available, from the MRC online psycholinguistic database www.psy.uwa.edu.au/MRCDataBase/uwa_mrc.htm

The word lists and lexical data were employed in three types of analyses, conducted with a view to examining similarities and differences between the three novels in: (i) overall structure and
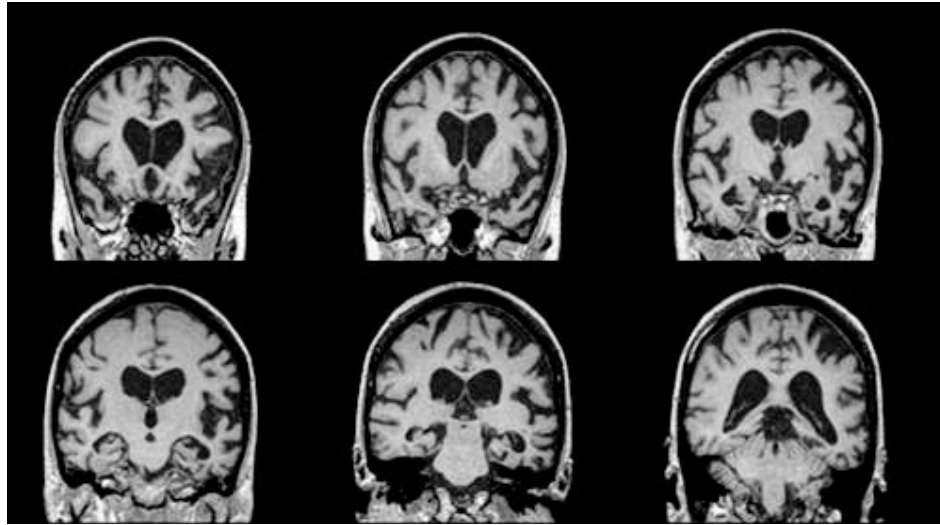
Fig. 1 Sample T1-weighted images from I.M.'s MRI (left hemisphere on right) showing bilateral hippocampal atrophy, moderate to severe global volume loss, with a temporal and parietal predominance.

Table 2 *General characteristics of the three books analysed*

	Mean (SD) of all 26 novels	*Under the Net* (1954)	*The Sea, The Sea* (1978)	*Jackson's Dilemma* (1995)
No. of chapters	28.5 (6.6)	20	8	13
No. of words	148 971 (47 714)*	101 456	209 085 [+]	90 004
No. of *dramatis personae*[†]	NA	11	≥17[‡]	19
Ratio of narrative to dialogue	NA	0.18	0.13[‡]	0.26

*Based on estimated mean numbers of words per page; [+] projected from a word count of 41 817 words in the first 100 pages; [†]based on the number of unique person names with >10 occurrences in the text; [‡]based on the first 41 817 words in the text.

vocabulary; (ii) syntax; and (iii) lexical characteristics. Analyses that could be fully automated were conducted on complete concordances, while those requiring manual input (e.g. coding of lexical variables or assignment of grammatical category) were performed on the smaller random word lists. In the generation of data for the syntactic and lexical level analyses, Concordance was configured to ignore text that had been designated as dialogue, because of the possibility that such passages may have been written to reflect a style other than the author's own.

Results
Overall differences between books

Table 2 summarizes the general characteristics of the novels and illustrates that, although differences are apparent at this very broad level, there is no consistent pattern: the first book is subdivided into more chapters than either of the two later works, while the middle work is far and away the longest of the three; the proportion of text that is devoted to dialogue rather than narrative is largest in the final novel and smallest in the mid-career work; more characters (defined as unique person names occurring >10 times the text) appear in *Jackson's Dilemma* (19) than in *Under the Net* (11, including Jake, the narrator, and Mars, a film-star dog), though in this respect *The Sea, The Sea* shows greater similarity to the final work than to the first. This lack of consistency is likely to be due to a wide range of variation in these measures over all the published works (though as the leftmost column in the table illustrates, it was only possible to determine this range for two of the four comparisons).

More detailed comparative information can be obtained from an analysis of the range of vocabulary used in the three books, for as well as containing varying numbers of words overall (number of word tokens), the three books differ in the number of distinct words (word types). The number of word types occurring within a section of text provides a measure of the variety of vocabulary used and, when examined at successive points in a text, this figure reflects the rate at which the author introduces new words. This information is displayed in Fig. 2 for incremental portions (first 10 000, first 20 000 words, etc.) of each book. As might be expected, an increasing trend, reflecting the continuous introduction of previously unused words, is revealed in all three texts. In all samples up to and including the first 40 000 words, however, the word type counts are consistently ordered, the smallest number associated with *Jackson's Dilemma* and the largest with *The Sea, The Sea* (all differences significant at $P < 0.001$ by the χ^2 statistic). Moreover, the rates of accretion
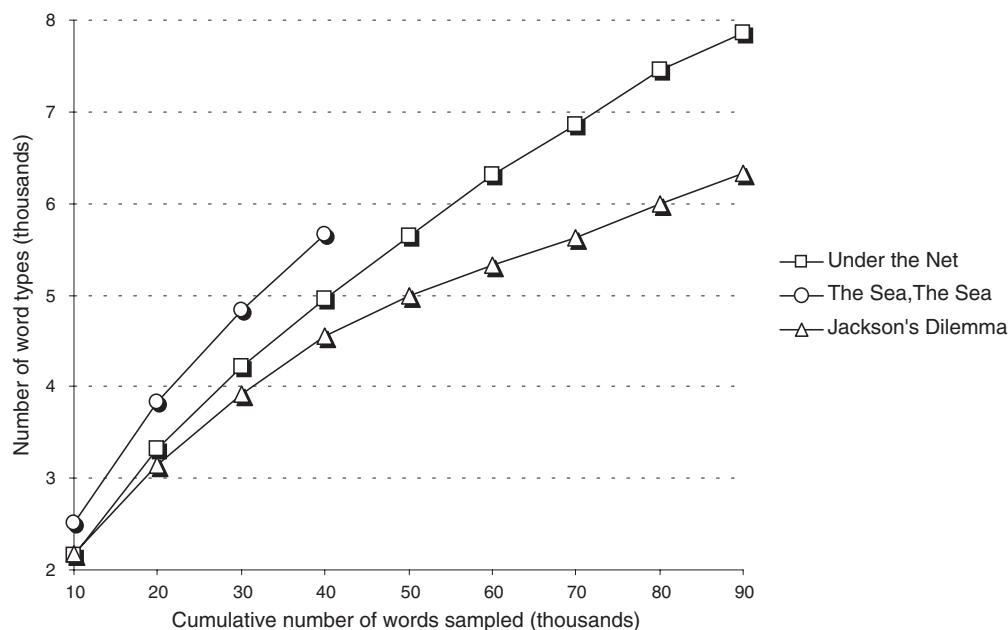
Fig. 2 Proportions of word types (*y*-axis) to word tokens (*x*-axis) at 10^3 word intervals in the texts of *Under the Net, The Sea, The Sea* and *Jackson's Dilemma*.

of new word types are also strikingly greater over the first 90 000 words of *Under the Net* than in the equivalent samples from *Jackson's Dilemma*: values of $\chi^2(1)$ for this comparison increase from 4.12 at the 20 000 word point to 152.2 at 90 000 words. These results suggest an enrichment of available vocabulary between the early and middle stages of I.M.'s writing career, followed by a relative impoverishment during the composition of the final work.

Syntactic differences
Complexity

A number of measures are available for characterizing written texts at the syntactic level, ranging from simple ratios such as mean number of words or clauses (Kemper *et al.*, 1989) per sentence, to complex metrics derived from the order of emergence of grammatical constructions during language development (Rosenberg and Abbeduto, 1987). (For a comprehensive review, and estimates of correlations among these various measures, see Cheung and Kemper, 1992.)

Table 3 displays the mean values of two syntactic measures derived from *Under the Net*, *The Sea, The Sea* and *Jackson's Dilemma*. Because there is no way of determining the beginning and end of a clause automatically, the mean number of subordinate clauses per sentence was calculated on a subset of sentences, namely the first 10 sentences from the first, middle and final chapters of each book. Mean word counts and number of clauses per sentence were computed for the same samples. As detailed in the table, separate one-way analyses of variance (ANOVAs) on the values of both variables for the

three books suggested significant differences. *Jackson's Dilemma* was associated with the smallest and *The Sea, The Sea* with the largest values, both of which differed significantly in *post hoc* comparisons ($P < 0.05$ for words per sentence, $P < 0.001$ for clauses per sentence). In contrast, none of the four pairwise comparisons involving *Under the Net* was statistically significant.

Although the comparisons just reported were made on similarly sized samples from equivalent points in the three books, the individual distributions within these consecutive series of values would have been influenced not only by the book's overall syntactic complexity, but also by the local thematic context. Because of this confound, two additional measures of syntactic complexity, based on automated analyses of the entire texts, were implemented: these considered (i) the total number of words divided by the total number of sentence-ending markers (full stops, exclamation marks and question marks); and (ii) the proportion of times the 10 most common words in each text were repeated within the space of five words ('auto-collocations'). Both these analyses were carried out using the Concordance software referred to in Materials and methods, with passages of dialogue and direct quotation excluded in both cases.

(i) As before, the difference between books using the expanded measure of sentence length was significant [$\chi^2(2) = 14.5$, $P < 0.01$]. In this analysis, however, there was a different rank order of values (see Table 3), and in this instance the only non-significant pairwise comparison was that between *The Sea, The Sea* and *Jackson's Dilemma* [$\chi^2(1) = 1.63$, $P > 0.05$].

Table 3 *Comparative measures of syntactic complexity in the three books*

	Under the Net	*The Sea, The Sea*	*Jackson's Dilemma*		*P* value
30-sentence samples				*F*(2,87)	
Words per sentence	19.7 (11.62)	22.7 (14.1)	15.3 (10.8)	3.5	<0.05
Clauses per sentence	1.5 (1.8)	2.2 (1.0)	0.8 (1.3)	7.4	<0.01
Automated analysis				$\chi^2(2)$	
Words per sentence-ending mark	13.9	15.1	14.6	14.5	<.01
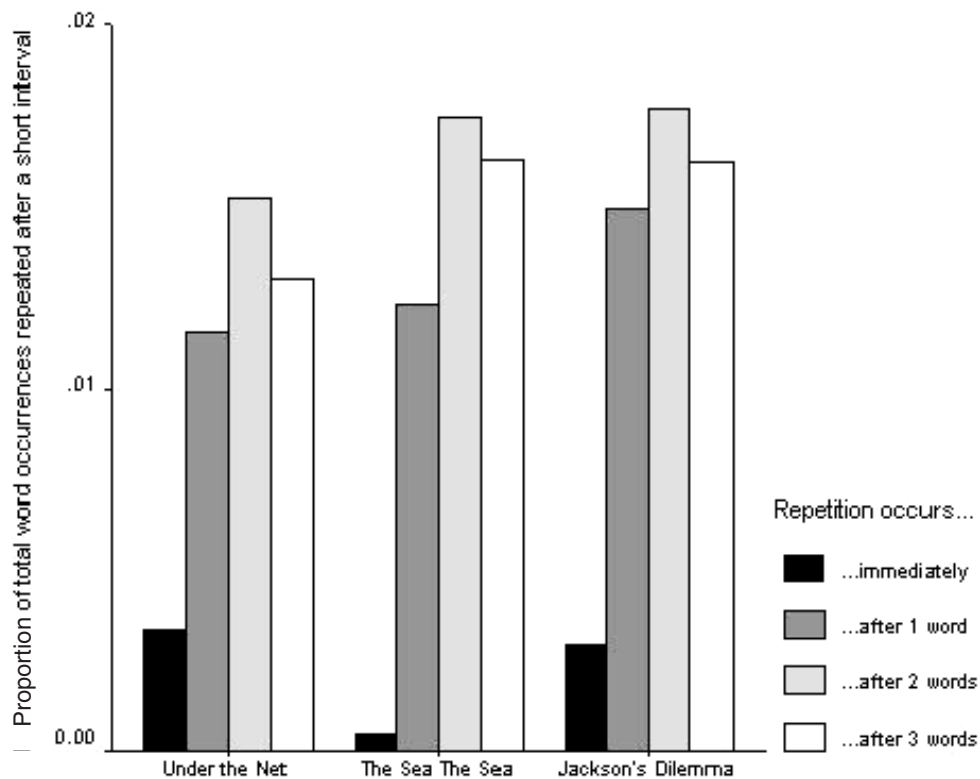
Fig. 3 Mean proportions of usages of the 10 most frequently occurring words in each book that appear twice within a series of short intervals, ranging from consecutive positions in the text to a separation of three intervening words.

(ii) The mean proportions of repetitions, at intervals of 1–4 words, of the 10 most common words in each book (all, unsurprisingly, function words such as 'the', 'a', 'and', 'of', etc.) are shown in Fig. 3. These data were analysed using a repeated-measures ANOVA with the number of intervening words as four-level within-subject factor, and book as between-subjects factor. The analysis revealed that the overall variation in the values associated with different intervals was highly significant [$F(3,57) = 12.1$, $P < 0.001$], but that the pattern of differences was similar across the three texts [$F(2,57) < 1$ for the main effect of book and $F(6,57) < 1$ for the book by interval interaction].

Grammatical class
The proportions of word types belonging to each of the defined grammatical categories (nouns, verbs, adjectives/

adverbs and function words) in each of the three random samples are displayed in Table 4. Function words clearly outnumber other grammatical classes by a sizeable margin in all three lists, but words of all classes are equally distributed among the samples [$\chi^2(6) = 4.9$, $P = 0.6$]. When the analysis was conducted on unique words (i.e. word types), the proportions of function words in all three books were predictably diminished (Table 4), but the null result persisted [$\chi^2(6) = 5.1$, $P > 0.1$].

Comment
The syntactic analyses reported above have produced conflicting results, which reflect a number of limitations of the methods used. The selection of sentences for manual analysis was imperfect because the sentences chosen represent consecutive elements of larger passages, with their own higher order structure, of which the length of the constituent sentences may have been a reflection. On the other hand, although the designation of a sentence as consisting of

Table 4 *Proportions of words in each book belonging to four grammatical categories*

	Grammatical class	*Under the Net*	*The Sea, The Sea*	*Jackson's Dilemma*
All words	Nouns	0.21	0.25	0.22
	Verbs	0.15	0.15	0.19
	Descriptors	0.15	0.14	0.15
	Function words	0.49	0.46	0.44
Unique word types only	Nouns	0.39	0.46	0.37
	Verbs	0.25	0.23	0.32
	Descriptors	0.29	0.23	0.24
	Function words	0.08	0.08	0.07

all the words between successive sentence-ending markers may overcome this confound, it obscures the obvious fact that some very long sentences are grammatically very simple, e.g.

> He feared the currents, the wind, the grim force of the waves, more savage now, larger, louder, taller, curling over in great white arches, hurling themselves in deafening impact against the slithering wall of stones, and in destroying themselves, each wave in its demise receding, dragging clattering down a grinding mass of sand and stones. (*Jackson's Dilemma*, p. 103)

and that the skilful deployment of embedded clauses and imaginative metaphor can produce shorter sentences with remarkable density of meaning:

> The great brown eyes, which once opened so blandly upon the world, seemed narrowed, and where Anna had used to draw a dark line upward at their corners the years had sketched in a little sheaf of wrinkles. (*Under the Net*, p. 37)

The collocation analysis was relatively free from either of these difficulties, and the fact that all the words used in these analyses were short, closed-class function words of uniformly high frequency implies that syntactic complexity is a major source of constraint on the density of their use within a single sentence. The analysis as conducted, however, did not distinguish between auto-collocations within the same sentence and those that crossed sentence boundaries, making it harder to distinguish between syntactic and stylistic or other higher order constraints on this variable.

Finally, no syntactic differences were found between the relative proportions of words belonging to different grammatical classes in the three books. Taken as a whole, therefore, this series of syntactic analyses provides limited evidence for any consistent differences among the three books at this level. It is anticipated that lexical differences will be somewhat simpler to quantify and interpret.

Lexical differences

Differences at the level of individual lexical variables were sought by comparing the distribution of values of these variables associated with the word samples that were drawn randomly from each of the three books.

Values of frequency (Kucera and Francis norms), familiarity, age of acquisition and concreteness were assigned to as many words as possible using the MRC online psycholinguistic database. Length was computed automatically for every word. Not surprisingly, the resulting data set contained large numbers of missing values: familiarity and concreteness ratings could only be found for 68 and 63% of words, respectively, while age of acquisition scores were available for only 9%. Values for lexical frequency, on the other hand, were associated with 94%, and word length was (obviously) obtainable for 100% of words.

Because of the large number of missing values, we decided to report only comparisons for content words (i.e. nouns, verbs and descriptors) in relation to lexical frequency and word length. The existing literature suggests contrasting predictions for these two variables. The large body of work documenting the increased vulnerability of low frequency words in production tasks, such as object naming or picture description, in Alzheimer's disease and semantic dementia (Lambon Ralph *et al.*, 1998; Thompson-Schill *et al.*, 1999; Bird *et al.*, 2000), predicts that the words used in *Jackson's Dilemma* should be associated with higher frequency values than those in *Under the Net* and *The Sea, The Sea*. In contrast, the absence of any influence of word length on naming in Alzheimer's disease reported in the study of Kremin *et al.* (2001) predicts that the mean number of letters per word in the three books should not differ.

The five word samples drawn from each book were merged to give three lists of words, one for each book, with counts summed across all samples in which the word had appeared. The mean values of the two variables of interest, weighted by their frequency of occurrence in their book of origin, are displayed in Fig. 4. It is apparent from Fig. 4A that the words drawn from *Jackson's Dilemma* have an overall higher mean frequency than those drawn from either of the other two works. Formal comparison of these values using one-way ANOVA revealed a difference of borderline significance [$F(2,548) = 2.5$, $P = 0.08$], with *post hoc* pairwise comparisons significant for the contrast between *Jackson's Dilemma* and *The Sea, The Sea* [$P < 0.05$ (corrected)] and marginal for that between *Jackson's Dilemma* and *Under The Net* [$P = 0.07$ (corrected)]. Comparison using non-parametric statistics confirmed the overall trend [Kruskal Wallis test: $\chi^2(2) = 5.2$, $P = 0.08$]. In contrast,
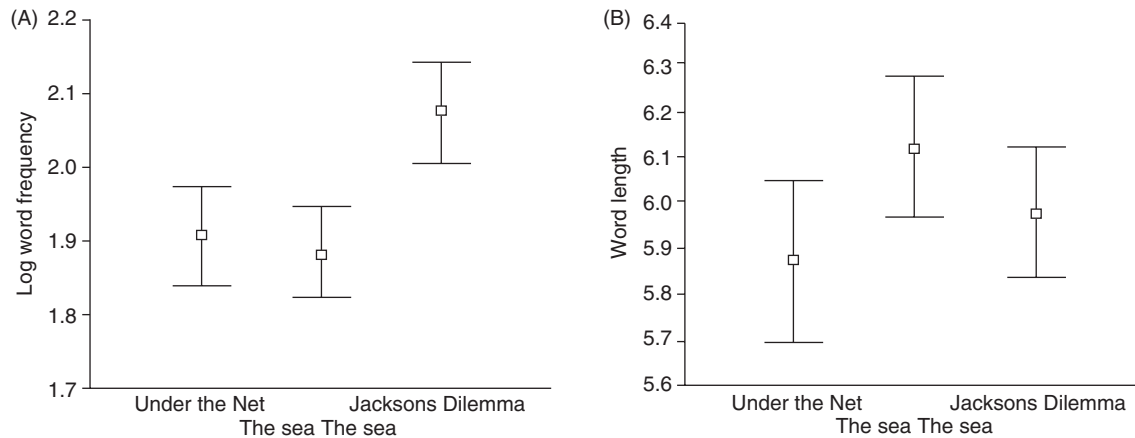
Fig. 4 Comparative distributions of values of: (**A**) frequency and (**B**) word length in the three books.

the much smaller differences in the distributions of word length seen in Fig. 4B failed to achieve statistical significance using either parametric $[F(2,580) < 1]$ or non-parametric $[\chi^2(2) = 3.3, P > 0.1]$ tests.

Discussion

Through a systematic comparison of three texts, we have attempted to establish, at a number of different levels of analysis, the ways in which a prominent writer's literary output might have been affected by the earliest stages of post-mortem-proven Alzheimer's disease. The first of these texts—*Under the Net*—was composed while the subject (I.M.) was in her early thirties, on the threshold of a distinguished career as writer and academic. The second—*The Sea, The Sea*—dates from a later period, but one in which she was still at the height of her creative powers. After four decades of prolific literary output, the third text—*Jackson's Dilemma*—was produced. At around the same time, the earliest symptoms of Alzheimer's disease were just beginning to emerge, albeit still unrecognized, above the surface of her cognitive performance.

A published literary text usually represents an end product that has evolved through multiple versions, often with the intervention of an editor. There is, however, good reason to believe that none of the three texts analysed herein was drafted for later amendment; I.M. always wrote out and submitted her books in longhand after lengthy periods of 'working out' the plot and characters in note form (Wilson, 2003). It is also highly unlikely that any editorial interference was tolerated, as this would have been entirely out of character (Bayley, 1998; Conradi, 2001).

I.M.'s clinical course and neuropsychological profile after the diagnosis had been established were both consistent with a temporal lobe predominance of Alzheimer's disease pathology: early naming problems attributable to disruption of the semantic system, together with the characteristic breakdown of episodic memory and variably severe involvement of frontal/executive and visuospatial functions. This putative neuroanatomical distribution was later confirmed at autopsy, and suggests that the composition of *Jackson's Dilemma* may have coincided with the spread of the pathological load from the earliest (transentorhinal) stage to involve the temporal neocortex proper (Braak and Braak, 1991). We hypothesized that this early temporal neocortical involvement was responsible for an alteration in her linguistic abilities, which in turn led so many critics and commentators to express their unexpected disappointment with the strangely altered quality of her final novel. An important question was whether this linguistic change would be manifested in a global fashion or confined to specific types of comparisons among the three texts.

In the analyses reported here, the most compelling findings emerged at the level of vocabulary and lexical selection. Evidence for the availability of a more restricted vocabulary during the writing of the final work was provided by: (i) the smaller number of unique word types relative to the overall word count; and (ii) the lower rate of increase of this proportion over successive incremental samples, in *Jackson's Dilemma* compared with the two earlier works. Both observations imply a greater rate of repetition of already used words in the final book, and a greater rate of introduction of new words in the two earlier works. The fact that these differences were more marked when the comparison was between *Jackson's Dilemma* and the mid-career book makes it unlikely that they represent no more than the evolution of I.M.'s literary style over time. The same is true of the differences that were found in the lexical frequency values of random selections of words from the three novels: although these did not differ by large magnitudes, they were nonetheless higher in the sample taken from *Jackson's Dilemma* than in those taken from either of the earlier works (significantly so in the case of *The Sea, The Sea*). That these results represent genuine rather than randomly occurring differences among the three samples is supported by the additional finding of no difference within the same samples with respect to the mean value of another lexical variable (word length).

In contrast to these relatively clear distinctions in vocabulary, syntactic differences between books from the beginning, middle and end of I.M.'s writing career were more difficult to demonstrate by any of four different methods.

(i) When manually selected groups of sentences from each book were compared using a range of measures that have been found to correlate with syntactic complexity, *Jackson's Dilemma* was associated with the lowest values (i.e. was least syntactically complex), and *The Sea, The Sea* with the highest. Although this pattern mirrors those found in the lexical analyses, the method of selecting sentences consecutively from larger continuous passages is susceptible to unseen biases relating to the higher order structure of the passages from which they are drawn.

(ii) A second method, based on the ratios of sentence-ending markers to word tokens, was free from these selection biases, but did not distinguish between genuinely complex long sentences with multiple embedded clauses, and those that consisted of concatenated lists or conjoined simple clauses. The fact that the contrasts obtained using this automated method differed from those identified using the manual method added to the uncertainty about the validity of this technique as a surrogate for syntactic complexity.

(iii) A third method, which looked at the repetition rate of closed-class function words at various intervals, showed no greater tendency for short-range repetition in any of the texts.

(iv) Finally, the proportion of word tokens belonging to each of four common grammatical classes revealed no differences in the proportions of any word types among the three books.

This pattern of findings, suggesting the use of a smaller, higher frequency vocabulary but relatively unchanged syntactic structure, is consistent with the predominantly temporal lobe distribution of pathology that was eventually demonstrated, and as such probably represents an early subclinical manifestation of the disease. This conclusion has both clinical and theoretical implications. From a clinical point of view, the results support the idea that the occurrence in the brain of Alzheimer's disease pathology may pre-date the onset of the earliest overt symptoms by years, or even decades (Ohm *et al.*, 1995). This in turn raises the possibility that an intellect of exceptional pre-morbid quality, and/or a lifetime's engagement with intellectual work, may either protect against cognitive deterioration or enable it to be masked (Hultsch *et al.*, 1999; Bennett *et al.*, 2003). The availability of multiple time points in the course of the present subject's lifetime at which similar analyses might be conducted suggests a novel means of testing the second of these possibilities.

Although the syntactic analyses reported here were of an exploratory nature, they are also germane to theoretical models of language organization, insofar as they provide some support for the separability of semantic and syntactic aspects of lexical information (Breedin and Saffran, 1999).

The ability of most Alzheimer's disease patients to produce well-formed sentences well into the later stages of the disease has been documented extensively (Kempler *et al.*, 1987; Bates *et al.*, 1995). Together with the predominance of lexical over grammatical errors in experimental production tasks, this finding has been interpreted as evidence for modularity of organization within the language processor. Bates *et al.* (1995) took issue with this view on a number of grounds, including differences in production frequency between over-learned grammatical structures on the one hand, and all but the most common lexical units on the other. Part of the basis for this claim was that production of grammatically correct speech is more automatic than lexical selection, which relies more on controlled retrieval processes. Although this distinction may be applicable to spoken language, it is hard to see how, in writing, the production of a grammatically correct sentence requires a lesser degree of controlled processing than the selection of an individual word. Our methods were far from exhaustive, and any conclusions must therefore be drawn with caution; but the notion that 'grammatical processes are carried out within a large and heterogeneous lexicon of distributed representations' (Bates and MacWhinney, 1989) seems at odds with our findings of clear-cut lexical changes in I.M.'s final book without obvious effects on the grammatical properties of the text.

The problem of how best to operationalize grammatical complexity was not the only shortcoming of this study. Another issue is that, because of the time-consuming procedures involved in the acquisition and analysis of texts, the comparisons drew on only three out of I.M.'s 26 novels and numerous works of non-fiction. It must, therefore, be considered possible that the observed differences may have fallen within normal bounds had more data points been available. Although this is certainly a valid criticism, three aspects of the findings reported here suggest that our conclusions are not challenged by this limitation. First, our sampling points were selected to reflect three contrasting periods in I.M.'s writing career: one of her earliest attempts at creative writing, the prime of her literary life and the early stages of a neurodegenerative illness. Secondly, the cognitive neuropsychology of Alzheimer's disease allowed clear predictions to be made about what differences might be expected between books (lexical frequency but not word length; semantically but not syntactically related measures), and our findings were largely consistent with these predictions. Finally, the expertise developed during a lifetime of professional writing might be expected to exert a cumulative effect on the characteristics of any author's output, and thus to produce a consistent and monotonic trend throughout that writer's career. Whilst the lexical level comparisons between the early and mid-career works were in keeping with this assumption, the reverse was the case in equivalent comparisons between the final book and either of the earlier texts. This pattern was most obvious in the serial measurements of type-to-token ratios (see Fig. 2), which reflect variety in vocabulary and the rate of introduction of new words. In this analysis, *Jackson's Dilemma* was

consistently associated with the lowest, and *The Sea, The Sea* the highest values.

The extent to which such quantifiable aspects of a literary *oeuvre* do evolve over the course of a writer's career is, of course, an empirical question, to which we hope future studies will provide an answer. The emergence of discernible trends, using additional automated techniques including latent semantic analysis (Landauer and Dumais, 1997) to probe different properties of I.M.'s writing, open up the intriguing possibility of demonstrating deviations from normality before anyone had the remotest suspicion of any untoward intellectual decline.

Acknowledgements

References

Astell AJ, Harley TA. Tip of the tongue states in lexical access in dementia. Brain Lang 1996; 54: 196–215.

Bates E, MacWhinney B. Functionalism and the competition model. In: MacWhinney B, Bates E, editors. The crosslinguistic study of senence processing. Cambridge: Cambridge University Press; 1989. p. 3–73.

Bates E, Harris C, Marchman V, Wulfeck B, Krutchevsky M. Production of complex syntax in normal aging and Alzheimer's disease. Lang Cognit Process 1995; 10: 487–544.

Bayley J. Iris: a memoir of Iris Murdoch. London: Duckworth; 1998.

Bennett DA, Wilson RS, Schneider JA, Evans DA, Mendes de Leon CF, Arnold SE, et al. Education modifies the relation of AD pathology to level of cognitive function in older persons. Neurology 2003; 60: 1909–15.

Bird H, Lambon Ralph MA, Patterson K, Hodges JR. The rise and fall of frequency and imageability: noun and verb production in semantic dementia. Brain Lang 2000; 73: 17–49.

Bozeat S, Lambon Ralph MA, Patterson K, Garrard P, Hodges JR. Nonverbal semantic impairment in semantic dementia. Neuropsychologia 2000; 38: 1207–15.

Braak H, Braak E. Neuropathological staging of Alzheimer-related changes. Acta Neuropathol 1991; 82: 239–59.

Breedin SD, Saffran EM. Sentence processing in the face of semantic loss: a case study. J Exp Psychol Gen 1999; 128: 547–62.

Cheung H, Kemper S, Competing complexity metrics and adults' production of complex sentences. Appl Pscyolinguistics 1992; 13: 53–76.

Conradi PJ. Iris Murdoch: a life. London: Harper Collins; 2001.

Crosile B, Ska B, Brabant M-J, Duchenne A, Lepage Y, Gilbert A, et al. Comparative study of oral and written picture description in patients with Alzheimer's disease. Brain Lang 1996; 53: 1–19.

Galton CJ, Patterson K, Xuereb JH, Hodges JR. Atypical and typical presentations of Alzheimer's disease: a clinical, neuropsychological, neuroimaging and pathological study of 13 cases. Brain 2000; 123: 484–98.

Garrard P, Lambon Ralph MA, Patterson K, Hodges JR. Semantic feature knowledge and picture naming in dementia of Alzheimer's type: a new approach. Brain Lang. In press 2004.

Grossman M, Rhee J. Cognitive resources during sentence processing in Alzheimer's disease. Neuropsychologia 2001; 39: 1419–31.

Hodges JR, Salmon DP, Butters N. Semantic memory impairment in Alzheimer's disease: failure of access or degraded knowledge? Neuropsychologia 1992; 30: 301–14.

Hultsch DF, Hertzog C, Small BJ, Dixon RA. Use it or lose it: engaged lifestyle as a buffer of cognitive decline in aging? Psychol Aging 1999; 14: 245–63.

Kemper S, Kynette D, Rash S, Sprott R, O'Brien K. Life-span changes to adults' language: effects of memory and genre. Appl Psycholing 1989; 10: 49–66.

Kempler D, Curtiss S, Jackson C. Syntactic preservation in Alzheimer's disease. J Speech Hear Res 1987; 30: 343–50.

Kremin H, Perrier D, De Wilde M, Dordain M, Le Bayon A, Gatignol P, et al. Factors predicting success in picture naming in Alzheimer's disease and primary progressive aphasia. Brain Cogn 2001; 46: 180–254.

Lambon Ralph MA, Graham KS, Ellis AW, Hodges JR. Naming in semantic dementia—what matters? Neuropsychologia 1998; 36: 775–84.

Landauer TK, Dumais ST. A solution to Plato's problem: the latent semantic analysis theory of the acquisition, induction, and representation of knowledge. Psychol Rev 1997; 104: 211–40.

Murdoch I. Under the net. London: Chatto and Windus; 1954.

Murdoch I. The sea, the sea. London: Chatto and Windus; 1978.

Murdoch I. Jackson's dilemma. London: Chatto and Windus; 1995.

Nebes RD. Semantic memory in Alzheimer's disease. Psychol Bull 1989; 106: 377–94.

Ohm TG, Muller H, Braak H, Bohl J. Close-meshed prevalence rates of different stages as a tool to uncover the rate of Alzheimer's disease-related neurofibrillary changes. Neuroscience 1995; 64: 209–17.

Porlock, H. Why can't critics reveal whether a book is worth reading? Sunday Times, October 22, 1995.

Rosenberg S, Abbeduto L. Indicators of linguistic competence in the peer group conversational behavioiur of mildly retarded adults. Appl Psycholing 1987; 19–32.

Snowdon DA, Greiner LH, Markesbery WR. Linguistic ability in early life and the neuropathology of Alzheimer's disease and cerebrovascular disease. Findings from the Nun Study. Ann NY Acad Sci 2000; 903: 34–8.

Thompson-Schill SL, Gabrieli JD, Fleischman DA. Effects of structural similarity and name frequency on picture naming in Alzheimer's disease. J Int Neuropsychol Soc 1999; 5: 659–67.

Warrington EK, James M. Visual object and space perception test. Bury St Edmunds (UK): Thames Valley Test Company; 1991.

Watt RJC. Concordance (v3.0). 2002. http://www.rjcw.freeserve.co.uk

Wilson AN. Iris Murdoch as I knew her. London: Hutchinson 2003.