

KEEPING TRACK OF A CONVERSATION: IMPAIRMENTS IN ALZHEIMER'S DISEASE

MARGHERITA ALBERONI*

Neurology Department, University of Milan, San Paolo Hospital, Milan, Italy

ALAN BADDELEY

Medical Research Council Applied Psychology Unit, Cambridge, UK

SERGIO DELLA SALA

Neuropsychology Unit, Dept of Neurology, Clinica del Lavoro Foundation, Medical Centre of Veruno (NO), IRCCS, Italy

ROBERT LOGIE

Department of Psychology, University of Aberdeen, UK

HANS SPINNLER

Neurology Department, University of Milan, San Paolo Hospital, Milan, Italy

SUMMARY

Earlier research suggests that AD patients tend to be particularly impaired in the central executive component of working memory, leading to problems in coordinating information from different sources. This suggests that they may be particularly handicapped in keeping track of conversations involving several people. This was studied in 19 AD and 19 matched control subjects. The patients were screened to minimize problems of face recognition and language comprehension and were then shown videotapes of conversations involving from two to five characters. After each tape, a statement made by one of the characters was presented and the subject required to point to the person who had made that statement. Performance was at ceiling for normals, except when the speakers had changed location, when some errors occurred. AD patients showed a clear tendency for performance to deteriorate as number of speakers increased, and to show higher error rates when participants changed location. Implications are discussed for AD patients attempting to cope with everyday social situations.

KEY WORDS—Alzheimer's disease, dementia, language comprehension, working memory.

Keeping track of 'who said what' in a conversation among several speakers is a common requirement of everyday cognition that requires moment-to-moment updating of information in working memory. It is accomplished with ease by most normal adults, but can present considerable difficulties to individuals suffering from deficits of working memory. Even healthy elderly people sometimes show a certain degree of impairment when attempting to follow a conversation among several speakers. Once other factors have been eliminated

such as impaired hearing and vision, problems in following a conversation may reflect cognitive ageing. Specifically, there may be increasing difficulties with age in allocating the resources needed to accomplish simultaneously both language comprehension and spatial and facial identification of the speakers, and link together these various sources of information. From clinical experience and during informal conversations with Alzheimer patients (AD/pts) this deficit appears to be paramount and one can readily gain the impression of an extreme difficulty in following a simple conversation, even when very few speakers are involved. The problem seems particularly acute when the patients are not addressed directly. They are not dramatically impaired in their grasp of the general meaning of the conversation, provided that the topics of the conversation are not too complex, that the cogni-

* Now at Department of Neurology, San Raffaele Hospital, Milan, Italy.

Address for correspondence: Dr Sergio Della Sala, Neuropsychology Unit, Centro Medico di Riabilitazione, 28010 Veruno (NO), Italy. Names of authors are listed in alphabetical order and do not imply different contributions by different authors.

tive deficits are not too severe, or that they do not have a profound language impairment. What appears to be impaired is their encoding of the 'what-who' association, or 'who said what'.

We would argue that this task shares the features of divided attention and, if performance is being tested after short delays, relies on the efficiency and capacity of working memory (eg Baddeley *et al.*, 1986). The task of following a conversation can be conceived of as comprising the ability to grasp, mostly from verbal interchange, the general meaning, while at the same time being aware of the person (who) and the specific (what) contributions to a conversation by each of the speakers. Rabbitt (1988) has discussed how the cognitive changes that occur in normal ageing can affect these and other forms of social interaction.

In the experiment reported here, we assessed in both normal elderly and in Alzheimer's patients the 'who' and 'what' aspects of the task, while ensuring that the verbal comprehension load was kept to a minimum. There is empirical evidence of an impairment of working memory ability in even mildly demented AD/pts (Baddeley *et al.*, 1986, 1992). In particular, such patients are greatly impaired relative to matched elderly controls when required to carry out two tasks simultaneously such as perceptual-motor tracking combined with auditory digit span. Baddeley *et al.* interpret their data in terms of a specific model of working memory. This comprises a central executive that is responsible for reasoning, decision-making and for allocating attention. It is also responsible for coordinating the operation of two subsidiary systems of working memory, namely the articulatory loop and the visuo-spatial scratch-pad. The phonological loop provides temporary verbal storage, while the visuo-spatial scratch-pad provides temporary storage of visuo-spatial material. Baddeley *et al.* argue that the dual-task impairment in AD/pts reflects a deficit in the coordinating function of the central executive component of working memory. This is coupled with a relative sparing of the articulatory system (Morris, 1984, 1986).

These results yield some insight into the deficits suffered by AD/pts. However, they arose from the use of rather artificial laboratory tasks and are not explicit as to how the deficits observed in the laboratory might manifest themselves in the day-to-day life of the patients. Our intention in this article is not to test the working memory deficit hypothesis. Rather, accepting the fact that AD/pts do suffer cognitive deficits, what aspects of *everyday* cogni-

tive activity are impaired? Thus for the purposes of the study reported here, we designed an experimental procedure that retained as far as possible some of the ecological features of everyday conversations. The AD/pts were required to watch and listen to these conversations. Our intention was that these attempts to retain some ecological validity in our procedures would provide insight into the everyday cognitive deficits of Alzheimer's patients and be useful for caregivers involved in their nursing. The performance measures were assumed to reflect the operation of working memory.

Thus the aims of this study were first to assess whether AD/pts with mild deterioration are more impaired than comparable healthy elderly in keeping track of conversation and, second, whether and how AD/pts employ the spatial cues linked to speakers taking part in the conversation.

METHOD

Subjects

There were two experimental groups as follows.

AD/patients: The demented group was made up of 19 right-handed AD/pts (11 female, 8 male), mean age 66.84 years, SD = 6.41, range 59–81. Their mean time spent in full-time education was eight years, SD = 4.52, range 3–17. The mean length of illness from first diagnosis was 2.1 years, SD = 1.01, range 0.5–4.0.

For 13 of these patients, the clinical deterioration had started before the age of 65, with deterioration for the remaining six patients commencing after the age of 65. The patients were selected from among all patients clinically diagnosed as having Alzheimer's disease and attending as outpatients at the First Neurological Clinic of the University of Milan (Della Sala *et al.*, 1986). Diagnosis of probable AD (McKhann *et al.*, 1984) was achieved on the basis of clinical history, neurological examination and a six-month longitudinal, neuropsychological assessment. This was combined with CT scans and laboratory data in order to exclude other possible dementing diseases (Spinnler and Della Sala, 1988).

Selection was based on the following criteria: the disease was in its early stages (ie less than four years); there was a mild degree of cognitive deterioration, broadly in line with those of DSM-III-R

(American Psychiatry Association, 1987) (criteria are detailed in Baddeley *et al.*, 1986); there was preserved autonomy in everyday life (exclusion of patients not currently in a family setting and requiring special care); there was no illiteracy; there was no sign of chronic, progressive liver, kidney, thyroid, lung or heart failure; there was no alcoholism and no use of drugs active on the CNS during the 24 hours prior to testing. Finally, the subject had to be right-handed and willing to take part in the study.

These criteria resulted in the selection of 30 patients, four of whom subsequently had to be excluded because they were unable to come to the laboratory. Three patients consistently failed to attend to the experimental stimuli, and a further patient consistently perseverated, producing incoherent confabulations triggered by one of the characters in the conversation test. From the remaining 26 AD/pts a further seven were excluded because they proved to be aphasic as measured by their performance on the Token Test (De Renzi and Faglioni, 1978), and/or scored below the chosen cutoff on a Face Recognition Test devised specifically for this study (see below).

The aim of this last selection was to reduce as far as possible the influence of aphasia (Token Test) and prosopagnosia (Face Recognition Test) acting as independent variables in the study.

As we mentioned in our introduction, we were not concerned with testing the working memory deficit hypothesis. However, we did feel that it was important to assess the overall level of intellectual performance of the patients. We accomplished this by using Raven's matrices (Raven, 1938) to measure overall non-verbal intelligence, the Street Completion Test as a measure of perceptual performance (Street, 1931), and a Digit Cancellation Test (Della Sala *et al.*, 1992; Spinnler and Tognoni, 1987) as a measure of selective attention. Normative data are available for all of these, based on a study involving a sample of 321 normal subjects (Spinnler and Tognoni, 1987). Performance of the AD/pts on Raven's matrices gave a mean of 12.6 (SD 8.2), whereas the median score for normals was 28.5, and fifth percentile cutoff 15.0. On the Token Test, the patients scored a mean of 26.6 (SD 3.6) compared with a normal median of 33.0 and a fifth percentile cutoff of 26.5. Mean performance on the Street Test was 4.8 (SD 2.8) for the patients as against a median for normal subjects of 7.0 and a fifth percentile cutoff of 2.25. Mean performance on the Digit Cancellation Test was 30.9 (SD 13.6)

for AD/pts, whereas the median for normals is 49, and the fifth percentile cutoff is 30.

Control group: This group consisted of 19 right-handed healthy elderly subjects (HE/s) (14 women and 5 men) whose mean age was 67.95 years (SD 5.43, range 58–76) and whose mean full-time education was 8.21 years (SD = 3.29, range = 4–13). They were approximately matched for age and educational level with the AD/pts, being mostly selected from the relatives of the AD/pts entering the study, and thus from the same sociocultural background. Criteria for inclusion in the experiment were: absence of any current or previous neurological, psychiatric or systemic illness which could interfere with the cognitive performance, and no use during the previous 48 hours of alcohol or any drug that might influence the CNS performance.

Both AD/pts and HE/s had normal or corrected to normal vision and no clinically apparent hearing difficulties.

Tests and procedures

Experimental material was videorecorded using a Sony Betamax camera and presented on a 20 inch colour monitor.

Face Recognition Test

This test was aimed at excluding patients with severe prosopagnosia, or impairments in memory for faces. Also it provided a measure of face recognition ability for those patients included in the study.

Testing procedure: The faces of two silent people (targets) were filmed separately in foreground one after the other, with each in view for about 15 seconds. Next, the two targets were shown sitting around a table with three other people who acted as silent distractors. Subjects were tested individually. Prior to testing, they were instructed to remember the faces of the two targets and to point to them on the screen when they appeared mixed with the distractors.

There were in total three recognition trials for each subject, involving two different targets on each trial. For any one trial, the two targets were of opposite sex, and were distinguished from the distractors through age and distinctive facial features.

None of the people appearing in this test appeared in the subsequent conversation test.

Scoring procedure: The possible responses were: correct, no response, or a false recognition. One point was given for each correct response, and the range of scores was 0–6. A score of less than 4 on this test was used as one of the exclusion criteria for the patient sample. There was no time limit for responding or correcting an initial response. This test required approximately seven minutes to complete.

Keeping Track of Conversations Test

Testing procedures: One practice conversation and eight test conversations were presented on the colour TV screen, with conversations held between two, three, four or five characters. At the beginning of each conversation the group of characters was filmed as a whole, as they sat in silence around a table. After a few seconds one of the characters started speaking and the camera moved forward and framed him or her in the foreground. The camera moved in this way from one character to another as they started speaking. At the end of the conversation, the camera moved back and the whole group was shown again for a few seconds.

The image of the group then faded away, and reappeared on the screen after a two–three second delay. In four out of eight scenes characters reappeared sitting in the same position in which they sat during the conversation. These were labelled the ‘consistent conversations’. In the remaining four scenes two of the characters who had been sitting close to one another swapped their positions, and these were labelled the ‘varied conversations’.

Following the presentation of the group as a whole, the experimenter attracted the patient’s attention and repeated, slowly and emphatically, one of the sentences that had been pronounced during the conversation. The subject was then asked to point on the screen to the person that he or she remembered as having spoken that sentence. There was no time limit for responding or correcting an initial response.

Subjects were tested individually. Prior to testing they were carefully instructed in what the task involved and an example conversation was shown.

The experimenter had to be satisfied that the subject fully understood the instructions. As in the formal test, subjects were shown a conversation among a group of characters. Each conversation consisted of four syntactically and semantically simple sentences, referring to everyday life. Two English translations of example conversations are shown in the Appendix.

Sentences never exceeded 15 words and their content concerned commonplace events or activities. This procedure was intended to overcome difficulties with memory load and language comprehension that might arise from long or very complex sentences. Characters taking part in the conversations were never the same. In each conversation single characters could be easily recognized through age, sex, clothes and distinctive paraphernalia such as beards and scarves. Each set of varied and consistent conversations consisted of four conversations respectively between two, three, four and five characters. In each conversation one of the characters did not speak (the ‘silent’ character). The silent character did not change places even in the varied conversations.

Possible answers in the attribution of a sentence to a character were correct, no response, or false attribution. In this last category two particular kinds of errors were considered. The first of these was attribution of a sentence to the silent character. The second error occurred only in the varied condition. This involved attributing the sentence to the character sitting in the position where the sentence had been spoken originally rather than to the original speaker who had subsequently moved.

Scoring procedure: Scores were based on information theory measures. Each correct answer was given a score of $-\log$ of the probability that the correct answer could be due to pure chance. This probability obviously varies with the number of alternatives, that is, the number of characters in each conversation. This procedure allowed us to compare performance in conversations that varied in difficulty. Each correct answer in the two-character conversations ($p = 0.5$) was given a score of $-\log(0.5) = 0.69$ (possible score range 0–1.38), for the three-character conversations ($p = 0.33$), $-\log(0.33) = 1.1$ (possible score range 0–2.2), for the four-character conversations ($p = 0.25$), $-\log(0.25) = 1.39$ (possible score range 0–4.17) and for the five-character conversations ($p = 0.2$), $-\log(0.2) = 1.61$ (possible score range 0–6.44).

RESULTS

Table 1 shows the mean correct scores for AD/pts and HE/s on the Face Recognition Test and for the varied and consistent conversation tests. These results are shown in Fig. 1 in terms of mean transformed per cent scores.

Table 1. Mean scores on the Face Recognition Test (FRT) and on the Keeping Track of Conversations Test (Consistent and Varied) from AD/pts ($N = 19$) and age-education matched healthy controls ($N = 19$)

	FRT (0-6)	Consistent (0-14.19)	Varied (0-14.19)
HE's			
Score	6	13.85	12.39
SD	0	0.86	1.99
Range	—	10.97-14.19	8.19-14.19
AD/pts			
Score	5.26	8.39	5.69
SD	0.8	2.66	2.76
Range	4-6	3.87-14.19	1.61-12.8

Nine out of 19 AD/pts produced at least one error on the Face Recognition Test. All of the HE/s performed without error on the Face Recognition Test.

We first examined the relevance of this apparent face recognition deficit in terms of overall performance on the conversation test. A two-way ANOVA was performed on AD/pts face recognition score (4, 5 or 6) by varied versus consistent versions of the conversation test. The analysis revealed there was no evidence for a difference in overall conversation performance as a function of face recognition score ($F_{2,16} = 0.35$). There was a main effect of varied versus consistent performance ($F_{1,16} = 14.05$; $p < 0.005$), but the variables did not interact ($F_{2,16} = 1.73$, NS).

Next, we conducted a three-way ANOVA between the two experimental groups (AD/pts and HE/s) and the two forms of the conversation test, for two, three, four or five characters. There was a highly significant group difference ($F_{1,36} = 100.30$; $p < 0.001$) between types of conversation test ($F_{1,36} = 32.66$; $p < 0.001$) and for different numbers of characters ($F_{3,108} = 58.19$; $p < 0.001$). Groups interacted with number of characters ($F_{3,108} = 22.05$) but only marginally with type of conversation test ($F_{1,36} = 3.03$;

$0.05 < p < 0.1$). Pooling the data across groups, the type of test interacted with the number of characters ($F_{3,108} = 3.19$; $p < 0.05$). Finally, there was three-way interaction between groups, type of test and number of characters ($F_{3,108} = 5.29$; $p < 0.005$). This last interaction is shown in Fig. 1. It is clear from the figure that the three-way interaction is due largely to the error-free performance of the healthy elderly on the consistent form of the conversation test. It is also interesting to note that even the healthy elderly make some errors on the varied version of the test when there are four or five characters taking part in the conversation.

We next examined the correlations between the performance variables and the specific ability measures. For the patient group, Raven's score correlated with face recognition performance ($r = 0.55$, $p < 0.01$), and the Street score correlated with performance on the consistent conversation test ($r = 0.51$, $p < 0.05$). None of the other specific ability measures correlated significantly with performance on face recognition, or with either version of the conversation task. Moreover, performance on the varied version of the conversation test did not correlate significantly with performance on the consistent version of the test. This was also true for the healthy elderly subjects.

DISCUSSION

From these analyses and from Fig. 1, it is clear that the AD/pts were significantly impaired on the conversation test relative to the healthy elderly. This difference between the groups was most marked with the large number of characters in the conversation. It is equally clear that the consistent task was performed better than the varied task for both groups, with a tendency for the patient group to be more disrupted by the characters moving around than were the control group. Again this effect was more noticeable with more characters.

Clearly, in the varied version of the task, the location of the speaker cannot act as a spatial cue. Thus the tendency for more errors to occur in this version of the task suggests that subjects may be using a spatial cue. Such a strategy may be adopted in the absence of other accurate information about the speaker, specifically the association between a speaker and their utterance.

There is a large literature which argues that degeneration in Alzheimer's disease is faster, or more frequent, in the left hemisphere than in the

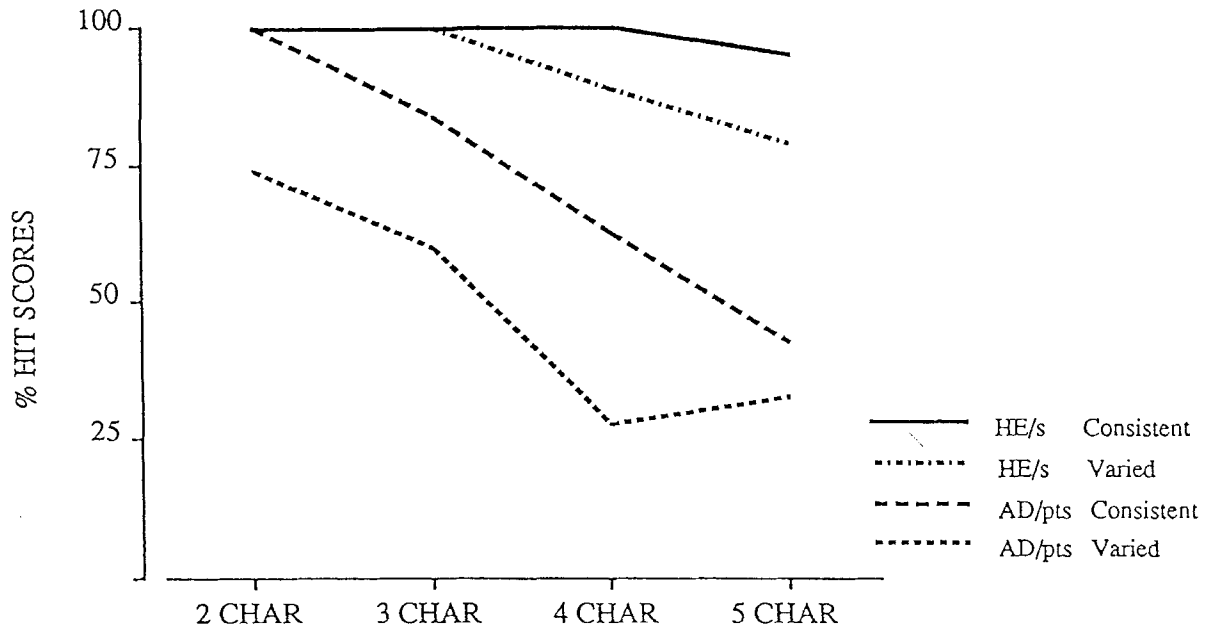


Fig. 1. Mean percentage of hit scores obtained by the two experimental groups (healthy elderly and Alzheimer patients) in the two tasks (consistent and varied) in the four experimental conditions (with two, three, four or five characters taking part in the conversations)

right (for a review see eg Capitani *et al.*, 1990). This would suggest that right hemisphere functioning would be less impaired. The literature on hemispheric differences generally argues that the right hemisphere is more efficient at processing spatial information. With a relative sparing of function in the right hemisphere, the AD/pts should then have available cognitive resources to allow them to try to take advantage of spatial cues when other, possibly more accurate information is not available. This view is entirely consistent with our interpretation as to the use of spatial cues by the AD/pts in our study.

What finally are the practical implications of our results? First, they support the clinical impression that AD/pts may have difficulty in keeping track of a conversation, even when it is very simple, as in the present study. The problem increases with the number of participants and is further exacerbated when people move from one location to another. While one might hope that the task might be slightly easier if all participants are familiar, nevertheless it seems likely that problems in keeping track will occur. As already pointed out by other authors (eg Bayles and Kaszniak, 1987), the best way of dealing with this problem is probably

to make relatives aware of the difficulty, and to point out that the fewer the conversational participants, the better the chance of the patient understanding the situation, regardless of whether this is a true conversation among visitors or a passively observed event such as a TV programme. Our findings can also have practical clinical implications. For instance, several commonly used remedial therapies tend to be group based (see eg Holden and Woods, 1988 for a review of the widely applied reality orientation) and our results seem to confirm the impression that 'being part of a group imposes additional processing demands in that subjects must repeatedly shift and refocus attention and cope with the memory load involved in keeping track of who says what' (Hart and Semple, 1990, p. 242). It could be worth considering that AD/pts probably communicate better in face-to-face interaction (Bayles and Kaszniak, 1987; Hart and Semple, 1990).

ACKNOWLEDGEMENTS

We are grateful to Ms Mirella Crevacore for her editorial help. The research reported was partly

supported by grant number 87.00233.04.115.12235 of the Consiglio Nazionale delle Ricerche, Italy and grant number 339 to H. Spinnler from the Regione Lombardia, Italy.

REFERENCES

- American Psychiatry Association (1987) *Diagnostic and Statistical Manual of Mental Disorders*, 3rd Edition revised. APA, Washington, pp. 103–123.
- Baddeley, A. D., Bressi, S., Della Sala, S., Logie, R. and Spinnler, H. (1992) The decline of working memory in Alzheimer's disease. *Brain*, **114**, 2521–2542.
- Baddeley, A. D., Logie, R., Bressi, S., Della Sala, S. and Spinnler, H. (1986) Dementia and working memory. *Quart. J. Exp. Psychol.* **38A**, 603–618.
- Baddeley, A. D., Hitch, G. (1974) Working memory. In *The Psychology of Learning and Motivation* Vol. VIII (G. A. Bower, Ed.), Academic Press, New York, pp. 47–90.
- Bayles, K. A. and Kaszniak, A. W. (1987) *Communication and Cognition in Normal Ageing and Dementia*. Little, Brown, Boston.
- Capitani, E., Della Sala, S. and Spinnler, H. (1990) Controversial neuropsychological issues in Alzheimer's disease: Influence of onset-age and hemispheric asymmetry of impairment. *Cortex* **26**, 133–145.
- Della Sala, S., Laiacina, M., Spinnler, H. and Ubezio, M. C. (1992) A Cancellation Test: Its reliability in assessing attentional deficits in Alzheimer's disease. *Psychol. Med.*
- Della Sala, S., Nichelli, P. and Spinnler, H. (1986) An Italian series of patients with organic dementia, *Ital. J. Neurol. Sci.* **1**, 27–41.
- De Renzi, E. and Faglioni, P. (1978) Normative data and screening of a shortened version of the Token Test. *Cortex* **14**, 41–49.
- De Renzi, E. and Spinnler, H. (1966) Visual recognition in patients with unilateral cerebral disease. *J. Nerv. Ment. Dis.* **6**, 515–524.
- Hart, S. and Semple, J. M. (1990) *Neuropsychology and the Dementias*. Taylor and Francis, London.
- Holden, U. P. and Woods, R. T. (1988) *Reality Orientation: Psychological Approaches to the 'Confused' Elderly*, 2nd edn. Churchill Livingstone, Edinburgh.
- McKhann, G., Drachmann, D., Folstein, M., Katzman, R., Price, D. and Stadlan, E. M. (1984) Clinical diagnosis of Alzheimer's Disease. *Neurology*, **34**, 939–944.
- Morris, R. G. (1984) Dementia and the functioning of the articulatory loop system. *Cog. Neuropsychol.* **1**, 143–157.
- Morris, R. G. (1986) Short term forgetting in senile dementia of Alzheimer type, *Cog. Neuropsychol.* **3**, 77–97.
- Rabbitt, P.M.A. (1988) Social psychology, neurosciences and cognitive psychology need each other; (and gerontology needs all three of them) *Psychologist*, **12**, 500–506.
- Raven, J. C. (1938) *Standardized Progressive Matrices. Sets A, B, C, D, and E*. Lewis, London.
- Spinnler, H. and Della Sala, S. (1988) The role of clinical neuropsychology in the neurological diagnosis of Alzheimer's disease. *J. Neurol.* **235**, 258–271.
- Spinnler, H., Della Sala, S., Bandera, R. and Baddeley, A. D. (1988) Dementia, ageing and the structure of human memory. *Cog. Neuropsychol.* **5**, 193–211.
- Spinnler, H. and Tognoni, G. (Eds) (1987) Standardizzazione e taratura italiana di test neuropsicologici. *Ital. J. Neurol. Sci.*, Suppl. 8 to No. 6.
- Street, R. F. (1931) A gestalt completion test. Bureau of Publications, Teachers College, New York.

APPENDIX

Translated text of the dialogues of the Keeping Track of Conversations Test. *indicates the sentences that are repeated by the examiner in order to have the subject try to attribute them to a character.

Two characters (A + silent character)

- A Yesterday I went fishing
I went with two friends of mine that you have met
We travelled over 1000 kilometres
*And caught not a single fish
- A2 *Yesterday I went around the town looking for a birthday present for a friend of mine
I knew that he would like an English tie
I walked along half the streets in the town to try and find one
Later I discovered that he already had one just like it

Three characters (A + B + silent character)

- A *This morning I made the mistake of going to work by car
- B Did it take you longer than by bus and underground?
- A Much longer: the whole town centre was jammed with traffic
- B *A better way to travel around the town is by motor-bike
- A2 *There's one thing I very much regret: not being a good cook!

- B Being a good cook is a matter of enthusiasm
*Mostly what you need is time and patience
- A2 I believe that one should learn such things
at school, as school subjects

Four characters (A + B + C + silent character)

- A I came here just to say good bye to you
*I'm leaving tomorrow morning on the 8
o'clock train
- B *Send our best regards to your family
- C *And come back to see us soon
- A2 Markets in Milan are very crowded just now
- B2 *There's one market in my neighbourhood
where you can find good fruit at very good
prices
- A2 *They cause, nevertheless, great problems for
traffic

- C2 *There was a time when markets were held
only for village festivals

*Five characters (A + B + C + D + silent
character)*

- A *I'll have a look at the newspaper to see if
there's something good on the TV tonight
- B *With the mass of independent TV channels,
it's hard to choose
- C *Yes, but the independent channels interrupt
programmes with advertising too often
- D *We are getting used to it now
- A2 *Have you ever noticed that people write let-
ters much less often than they used to?
- B2 *And when can you find time to write letters?
- C2 *Well, it's obvious the telephone is so much
handier!
- D2 *Yes, but since it's so expensive, you can't
use the telephone for everything