Neuropsychologia 44 (2006) 469–478

NEUROPSYCHOLOGIA

www.elsevier.com/locate/neuropsychologia

Qualitative features of semantic fluency performance in mesial and lateral frontal patients☆

Carlo Reverberi[a,*], Marcella Laiacona[b], Erminio Capitani[c]

[a] *International School for Advanced Studies (SISSA-ISAS), Trieste, Italy and Department of Psychology,*
Università Milano - Bicocca, Piazza Ateneo Nuovo, 1 20126 Milano, Italy
[b] *S. Maugeri Foundation, Scientific Institute of Veruno (NO), Italy*
[c] *Università Statale di Milano, Milano*

Received 4 November 2004; received in revised form 16 May 2005; accepted 19 May 2005

Abstract

Semantic verbal fluency is widely used in clinical and experimental studies. This task is highly sensitive to the presence of brain pathology and is frequently impaired after frontal lesions. Besides the total number of words generated, a qualitative analysis of their sequence can add valuable information about the impaired cognitive components. Thirty-four frontal patients and a group of matched controls were examined. Besides the number of words and subcategories retrieved by each group, we analysed two distinct aspects of the word sequence: the search strategy through a semantically organised store and the ability to switch from one subcategory to another. We checked whether the pattern of impairment changed according to the lesion site within the frontal lobe. Overall, patients produced fewer words than controls. However, only lateral frontal patients presented a reduced semantic relatedness between contiguously produced words and a specifically increased proportion of switches to different subcategories. The performance of lateral frontal patients was in line with the hypothesis of a search strategy impairment and cannot be attributed to a switching deficit. The performance of mesial frontal patients could be ascribed to a general deficit of activation.

Keywords: Verbal fluency; Semantic memory; Strategic retrieval; Switching; Frontal lobes; Executive functions

---

1. Introduction

The semantic fluency task (Bousfield & Sedgewick, 1944; Gruenewald & Lockhead, 1980; Rosen & Engle, 1997) requires the subject to produce as many words as possible from a given category (e.g. "animals") within a brief fixed time. Although semantic fluency is widely used in clinical and experimental studies, it is still debated how the produced words should be evaluated in order to optimise the information yielded by the task. The number of words produced per se is highly sensitive to the presence of brain pathology, and is useful for basic clinical purposes; however, the same measure does not give information about the specific cognitive steps of the task. For this, we should first consider how semantic fluency operates in normal subjects. Then we can attempt to identify which stage of the normal processing is impaired in a patient, also with reference to our knowledge of the functions of different parts of the brain. The following facts seem relevant for our analysis.

When requested to produce all the items of a category (e.g. "fruit"), healthy subjects tend to generate sequences of words from the same subcategory, e.g. "soft fruit", "dried fruit", etc. (Gruenewald & Lockhead, 1980; Wixted & Rohrer, 1994). After having retrieved typical and familiar items of one-subcategory subjects generally shift to another subcategory. We can presume that shifting to a different subcategory represents a more efficient strategy than continuing to retrieve less typical or familiar items from the former subcategory. Accordingly, a hypothetical sequence

---

☆ The study was carried out in the Santa Maria della Misericordia Hospital (Udine, Italy), in SISSA (Trieste, Italy), in the S. Maugeri Foundation, Veruno, Italy, in Università Statale di Milano and in the Università Milano - Bicocca (Milano, Italy). The study was approved by the ethical committee of SISSA.

* Corresponding author. Tel.: +39 0264483776; fax: +39 0264483706.
E-mail address: carlo.reverberi@unimib.it (C. Reverberi).

0028-3932/$ – see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.neuropsychologia.2005.05.011

of produced fruit exemplars might be: *apple–pear–orange–tangerine–lemon–apricot–peach–strawberry–blackberry–raspberry–chestnut–hazelnut–almond–peanut*. In symbolic notation, each word represented by a letter of the alphabet standing for a specific subcategory, the sequence can be written as AABBBCCDDDEEEE. Summing up, normal behaviour seems based on two distinct aspects: knowledge of the semantic relations within a category as the basis for forming subcategories, and the tendency to abandon a subcategory at a given point of the probe, i.e. to "switch".

Among neurological patients, cases of frontal damage are of special interest for this task. Frontal patients generate fewer words in a semantic fluency task (Baldo & Shimamura, 1998; Troyer, Moscovitch, Winocur, Alexander, & Stuss, 1998). From a recent meta-analysis, the sensitivity of semantic fluency to frontal lesions is comparable to that of letter fluency and is higher than that of the Wisconsin Card Sorting Test (Henry & Crawford, 2004), although different lesion sites, particularly the left temporal lobe, can also cause an impairment. Troyer, Moscovitch and Winocur (1997) suggested that semantic fluency implies an executive function devoted to strategic retrieval processes and monitoring (the "frontal lobe" component) while a different component would be responsible for automatic encoding and information retrieval (the "temporal lobe" component). Imaging data provide further evidence in favour of the crucial role of frontal areas in the neural network subserving semantic fluency (Frith, Friston, Liddle, & Frackowiak, 1991; Gurd et al., 2002; Warburton et al., 1996).

Given these findings, a consequent research task is to identify which of the cognitive functions normally working during semantic fluency can be ascribed to the frontal lobes, as their inactivation is presumably responsible for the impairment that frontal patients present in this task. The most likely candidates (Stuss & Levine, 2002; Troyer et al., 1998) are: (i) the generation and implementation of an effective search strategy; (ii) the ability to switch between different subcategories; (iii) the ability to initiate a task and keep it activated; (iv) the ability to monitor the task (useful, for instance, for avoiding repetitions of already retrieved words). These aspects deserve a closer analysis.

Compliance with a strategy is essential to conduct an efficient search through a semantically organised store such as the semantic memory. The hallmark of a disorganised search through semantic memory is the production of a disorganised word sequence. An example of such a sequence for the category "fruit" may be: *apple–orange–cherry–blackberry–pear–tangerine–banana–raspberry–lemon–apricot*; in symbolic notation, the sequence is ABCCABDCBE. The same effect on sequence organization could depend on a semantic memory damage able to scramble the common associative and categorical links among the items; however, the literature does not report reliable cases of primary semantic deficits due to frontal dysfunction. Therefore, we should presume that, although the semantic store per se is spared in frontal patients (Sylvester & Shimamura, 2002), it cannot be acted upon by a suitable search strategy during the word retrieval process: we term this "loss of strategy". In this case, the sequence of produced words would be no longer grouped in subcategorical clusters. As a consequence, we should observe a lower average relatedness between successive words and a corresponding higher frequency of relative switching between different subcategories.

A further type of deficit might affect the ability to abandon an old subcategory for a new one, i.e. to switch after the production of the main exemplars of a given subcategory in order to avoid excessive slowing down of the retrieval speed. An example of a sequence that may be produced by a patient suffering from switching deficit is: *orange–tangerine–lemon–grapefruit–clementine–strawberry–blackberry–gooseberry–blueberry–raspberry,* in symbolic notation AAAAABBBBB. In the case of an inability to abandon a perused subcategory, defined here as "loss of switching", we would expect a lower switch rate and a higher semantic relatedness between successive words. Moreover, this class of patients is also expected to explore less subcategories than healthy participants.

Frontal patients could also suffer from an initiation and activation deficit and be simply slower while retrieving each item of the sequence. The sequence produced in this case would be similar to the normal one but the number of words would be lower due to the time constraint. Also the number of subcategories and absolute number of switches would be less, while the relative number of switches and average semantic relatedness between successive words would remain unaffected. In symbolic notation, a performance of this type might be AABBBCC (compare with the sequence for healthy controls: AABBBCCDDDEEEE).

Finally, a failed monitoring of already retrieved words or already used cues would increase the repetitions. The influence of this deficit on semantic relatedness or on the number of switches would depend on whether or not the repetitions are included in the computations.

The literature provides evidence that the cognitive functions cited above are related to the frontal lobes or even to more circumscribed parts of these structures. Frontal patients present deficits in the generation and implementation of strategies (Alexander, Stuss, & Fansabedian, 2003; Burgess & Shallice, 1996b; Fletcher, Shallice, & Dolan, 2000; Gershberg & Shimamura, 1995; Incisa della Rocchetta & Milner, 1993; Owen, Downes, Sahakian, Polkey, & Robbins, 1990; Stuss, Alexander, Palumbo, & Buckle, 1994), in switching between cognitive sets (Owen, Roberts, Polkey, Sahakian, & Robbins, 1991; Troyer et al., 1998; Warrington, 2000), in the initiation of responses (Burgess & Shallice, 1996b; Godefroy, Lhullier-Lamy, & Rousseaux, 2002) and in the monitoring and checking of an ongoing task (Henson, Shallice, & Dolan, 1999; Luu, Flaisch, & Tucker, 2000; Reverberi, Lavaroni, Gigli, Skrap, & Shallice, 2005).

Coming back to semantic fluency, what we need is a way to evaluate the status of the different cognitive components of the task starting from a raw series of words. In an influential work, Troyer and collaborators (1997) devised a principled solution based on a two-component model of memory search (Moscovitch, 1992), which assumes the existence of an executive component devoted to strategic retrieval processes and monitoring and another component responsible for automatic encoding and information retrieval. Accordingly, Troyer and collaborators introduced two indices for scoring semantic fluency, that they termed "switching" and "clustering". Switching was defined as the *absolute* number of transitions between subcategories and clustering was defined as the average number of words of semantically related items in a sequence ("cluster"). In the study by Troyer and colleagues, frontal patients produced a lower number of words and, interestingly, a lower absolute number of switches between subcategories. However, in our opinion, the *absolute* count of the switches between different subcategories does not help to discriminate among the above-mentioned possible components of frontal impairment. In fact, all the factors that decrease the number of produced words can also affect the *absolute* switch index. In particular, a mere initiation deficit would decrease the absolute number of switches, as a potentially normal sequence would be cut only because of its slow production (see online supplementary material for an extended discussion of these points). Strong criticism based on the numerical implications of the proposal of Troyer and colleagues was made by Mayr (2002), while Reverberi, Capitani and Laiacona (2004) pointed out the danger connected to a too-loose definition of subcategories and, consequently, of switches.

Considering the different criticisms put forward regarding the "clustering" and "switching" indices, Reverberi and colleagues (2004) suggested a different approach that avoids counting switches but introduces instead an index for describing the structure of the word sequence. Studying the category "fruit", the authors preliminarily collected semantic proximity ratings for each pair of the 32 items most frequently produced in a semantic fluency task. Using these data, they computed the average semantic proximity between the pairs of successively produced words, a quantitative index that reflects the degree of semantic organization of a sequence.

The average semantic proximity seems a promising index for rating the semantic structure of a series of words[1] and the approach for analysing semantic fluency proposed by Reverberi et al. (2004) closely fits the aims of the present study. Moreover, proximity data also grants a sound and empirically validated way to identify the taxonomical tree of the explored category, providing a principled definition of the subcategories. Indeed, only through a stringent definition of the subcategories and a list of their components the number of switches can be reliably quantified. Since the absolute number of switches is ambiguous, it seems wiser to calculate the *relative* number of switches with respect to the sequence length.

The aim of the present report is to analyse the cognitive dimensions that characterise the semantic fluency impairment of frontal patients. We studied a sample of patients affected by frontal lesions by means of some original indices derived from previous studies on healthy participants carried out by our group. As the literature suggests that the mesial aspects of the frontal lobe are related to general activation while the lateral aspects are crucial for strategy compliance, we aimed to verify whether the type of performance of our patients differed according to the locus of the frontal lesion.

2. Material and methods

2.1. Participants

Thirty-four patients with a single frontal focal lesion detected after either CT or MRI scan were recruited from the Neurological and Neurosurgical ward of the Ospedale Civile of Udine (Italy): patients were classified as suffering from mesial or lateral lesions by two radiologists, and there was a complete agreement between their classifications. All patients gave their consent to participate in the study. The aetiology was mixed (Table 1). Exclusion criteria were a clinical history of psychiatric disorders, substance abuse or a previous neurological disease, neuro-radiological evidence of diffuse brain damage, and age under 18 or over 70 years.

Length of illness ranged from 15 to 1579 days, and did not significantly differ between the two subgroups with lateral and mesial lesions (Mann–Whitney test, $P > 0.10$). For the neoplasia cases time from surgery was taken as the length of illness. Thirty-six normal volunteers, recruited from the patients affected by slipped disc pathology at the Udine hospi-

---

[1] As an example of how this index works consider again the above fragments of sequences, "normal" versus "disorganised", and the sequences representative of "switch" and "initiation" deficits (the proximity of each item to the following in the sequence is reported):

(i) Example of a hypothetical normal production: the average proximity is **6.77** (apple-8.41-pear-3.81-orange-9.47-tangerine-8.41-lemon-2.33-apricot-7.50-peach-5.00-strawberry-8.31-blackberry-9.68-raspberry-2.10-chestnut-6.50-hazelnut-9.00-almond-7.55-peanut).
(ii) Example of a clearly disorganised production: the average proximity is **3.59** (apple-5.47-orange-2.81-cherry-6.68-blackberry-2.64-pear-3.55-tangerine-3.10-banana-2.20-raspberry-3.55-lemon-2.33-apricot).
(iii) Example showing a switch deficit: the average proximity is **7.76** (orange-9.47-tangerine-8.41-lemon-8.75-grapefruit-8.05-clementine-2.22-strawberry-8.31-blackberry-8.17-gooseberry-8.00-blueberry-8.50-raspberry).
(iv) Example showing an initiation deficit: the average proximity is **6.66** (apple-8.41-pear-3.81-orange-9.47-tangerine-8.41-lemon-2.33-apricot-7.50-peach).

The proximity index should decrease in the case of a strategic deficit, increase in the case of a switch deficit but should not be sensitive to initiation problems since a fairly "organised" sequence can be produced with just a few words.

Table 1
Number of patients and aetiology for each lesion group

	Mesial frontal lesions	Lateral frontal lesions	Frontal patients overall
Glioma	2	2	4
Meningioma	15	8	23
Ependymoma	1	–	1
Stroke	2	3	5
Histiocytosis	–	1	1

tal and from patients' relatives, constituted the control group. Controls and patients were matched for age and education (Table 2).

2.2. Examination procedure

Participants were asked to produce as many words as possible within 3 minutes from the category "*fruit*". They were explicitly instructed to say aloud any fruit name that "popped into mind" during the search, even if they were aware of repeating themselves. The *pop in mind* procedure (Rosen & Engle, 1997) was preferred over the standard one (in the common test setting participants are asked to avoid repetition) since we were interested in analysing retrieval processes, not the post-retrieval stages such as monitoring and checking. If participants had been instructed to avoid uttering repetitions, an unpredictable distortion could have been introduced in the underlying real sequence.

Before beginning the fluency task with the category "fruit", participants were given an example of a correct sequence for the category "*colours*". A "pop in mind" fluency trial lasting 45 seconds with the category "animals" was also carried out to ascertain task comprehension and the compliance with instructions.

In contrast with most studies from the literature where the allotted time was 1 minute, we collected fluency data over a 3 minutes period. This was so in order to collect a larger series for analysis and better evaluate the strategic aspects of the task (which can be better discerned after the most frequent or familiar words of a category have been uttered).

2.3. Scoring and experimental measures

We registered for each subject:

(i) The *number of New Words* produced in 3 minutes.

(ii) The *number of Repeated Words*. This variable was introduced in the analysis as a percentage of the total number of produced words (new plus repeated), and was submitted to angular transformation.

(iii) The *number of Subcategories* represented in the list of produced words. We used a non arbitrary classification of fruit into discrete subcategories, based on the authors' previous study (Reverberi and collaborators, 2004). In that paper we estimated the semantic proximity on a 1–10 scale (where 10 indicates the top similarity) between all the pairs derived from the set of 32 most frequently produced fruit names. Using these data, we identified five fruit clusters by means of a cluster analysis of the above 32 items. By means of discriminant analysis we fixed the threshold of semantic proximity that best discriminated pairs of fruit exemplars belonging to the same cluster from pairs belonging to different clusters (5.84). In the present study, 84 fruit names were produced by patients and controls. The semantic proximity for each of the 420 additional pairs of fruit items was estimated by eight healthy participants (see later). All items whose mutual semantic proximity (with the corresponding item of the pair) was higher than 5.84 were included in the same subcategory. In this way we identified 15 groups of fruit, eight of which included more than two items.

The subcategory analysis was carried out considering both the raw number of produced subcategories and the relative number of subcategories with respect to the total number of produced words. The usefulness of these different ways of scoring subcategories is that we have different empirical expectations regarding these measures according to at least two of the hypotheses about the nature of the cognitive deficit of frontal patients: (i) a switching deficit and (ii) an initiation deficit. In the case of a pure switching deficit we would expect a decrease in both the raw number of subcategories and the proportion of subcategories over the total number of uttered words. In the case of an isolated initiation deficit we would expect to find a *decrease* in the number of produced subcategories, because the sequence production is slowed down, but a normal proportion of subcategories with respect to the total number of words. Finally, in the case of a strategic deficit the predictions regarding subcategories are less straightforward, but neither a decrease of the raw score nor a decrease of the proportional score would be expected.

(iv) The *Relative Number of Switches*, i.e. the ratio of the raw number of observed switches divided by the total num-

Table 2
Demographic variables of patients and controls

	Mesial frontal patients	Lateral frontal patients	Frontal patients overall	Control participants
Number	20	14	34	36
Mean age (S.D.)	54 (14)	50 (12)	52 (13)	47 (9)
Mean education (S.D.)	9.0 (3.2)	8.8 (2.9)	8.9 (3.0)	10.3 (3.6)
Gender (female rate)	0.56	0.50	0.53	0.53
Length of illness in days (range)	352 (22–1555)	566 (15–1579)	455.5 (15–1579)	–

ber of words generated minus 1, including repetitions. On the basis of our proximity data, we calculated the absolute number of switches between subcategories. A switch was defined as any transition between two items that had a proximity index lower than 5.84. In the analyses we considered the Relative Number of Switches, not the absolute number, as we judged it more appropriate for the interpretation of the patients' performance in the present theoretical framework (see online supplementary material for an extended discussion of this point). However, also the absolute switch values will be reported in the results.

(v) The *Order Index*. The main reason for introducing this index was the observation that the number of switches, number of subcategories and number of produced words are not independent. Because the range of possible number of switches depends on both the number of subcategories and number of produced words, the comparability of the number of switches of different subjects could be undermined if these subjects present different amounts of produced words and subcategories. A way to make the number of switches from different participants comparable is to calculate the discrepancy between the theoretical maximum number of switches and the actually observed number of switches, and to divide this discrepancy by the difference calculated as follows: theoretically admitted maximum *minus* theoretically admitted minimum number of switches. The conceptual meaning of the Order Index is analogous to the Relative Number of Switches, but the Order Index has the advantage of avoiding the bias due to interdependence of the variables at issue. Moreover, the use of this index could be mandatory if one wishes to compare semantic fluency tasks based on categories that encompass a different number of items or subcategories (e.g. "animals" and "fruit"). See online supplementary material for a detailed explanation on how Order Index was calculated. We have also made available online a spreadsheet file (both in Microsoft Excel and Open XML standard) devised to automatically compute the angular and linear Order Index from N, SC and a SW.

(vi) The mean *Semantic Proximity* between each successive pair of produced words. This index was calculated by Reverberi et al. (2004) on a sample of 32 fruit most frequently reported in a semantic fluency task. Seventy-eight healthy subjects were requested to rate the similarity between each pair of fruit. In addition, the semantic proximity for each of 420 additional pairs of fruit items was estimated by eight healthy participants in order to have an exhaustive and objective similarity rating between all the pairs of successive fruit items produced by the patients included in our experiment. Theoretically this index ranges on a continuous scale from 1 to 10, and, accordingly, the observed values range from a minimum of 1 (e.g. pine-kernel and watermelon) to a maximum of 10 (e.g. fig and early fig).

2.4. *Statistical analysis*

2.4.1. *Our experiment had three dimensions:*

(A) The number of produced words, partitioned into the subtypes of New Words and Repeated Words.

(B) The number of Subcategories produced by each subject and the percentage of Subcategories with respect to the total number of produced words (also the latter underwent the angular transformation).

(C) The "order" according to which the words were produced. This dimension can be assessed by each of the three variables listed above, i.e. the mean Semantic Proximity between each successive pair of produced words, the Order Index, and the Relative Number of Switches. In principle, the first of the above measures, that reflects the whole sequence, should provide distinct information with respect to the quantification of a discontinuous phenomenon like switching. Notwithstanding its theoretical limits, we still included relative switching as a variable because the concept of switching has been used very often in the literature on frontal lobe disorders. We adopted an overall protection for Semantic Proximity, Order Index and the Relative Number of Switches.

According to the dimensions of the experiment, we decided to run three families of statistical analyses, and in each case to adopt an overall 0.95 protection against type I error. All statistical comparisons within each of the three dimensions (for all dependent variables and all subject groups) were carried out with a multiple comparisons approach. As such, the overall type I error risk *per experiment* was at most 0.15. Whenever a given comparison was not significant according to the adopted protection but would have become significant if viewed through the most powerful (and less protected) approach, i.e. with an error risk of 0.05 *per comparison*, we report this finding as a *trend*.

The analyses were based on MANOVA and were adjusted for age and education. Multiple comparisons were carried out according to Roy and Bose (1953). The difference between the two frontal subgroups was evaluated within a general analysis including also controls, adjusted for age and education, and in a separate analysis including only frontal patients, where also the length of illness (transformed into its reciprocal) was introduced. Whenever a variable was a percentage, its analysis was performed after an angular transformation.

3. Results

Table 3 shows the means of the experimental variables for each group.

Tables 4a and 4b report the correlation matrices between the most relevant variables displayed in Table 3.

As anticipated in the Methods, a preliminary MANOVA was carried out for each of the three variable sets corresponding to the dimensions of our experiment. The results are

Table 3
Means of the experimental variables (S.D. in parenthesis)

Variable	Controls ($n = 35$)	Frontal mesial ($n = 20$)	Frontal lateral ($n = 14$)	Frontal patients ($n = 34$)
Number of produced words				
New words	23.09 (5.48)	17.30 (4.62)	18.29 (4.53)	17.71 (4.54)
Repeated words	3.97 (4.63)	5.50 (4.12)	4.64 (3.78)	5.15 (3.95)
Repeated words/total words	0.13 (0.12)	0.22 (0.14)	0.19 (0.14)	0.21 (0.14)
Total Words	27.06 (7.09)	22.80 (6.60)	22.93 (5.00)	22.85 (5.91)
Subcategories				
Subcategories	7.83 (1.52)	7.20 (1.37)	7.71 (1.86)	7.41 (1.58)
Subcategories/total words	0.30 (0.07)	0.34 (0.12)	0.35 (0.09)	0.34 (0.10)
Order				
Order index	0.35 (0.12)	0.29 (0.08)	0.24 (0.10)	0.27 (0.09)
Relative number of switches	0.67 (0.09)	0.72 (0.07)	0.77 (0.07)	0.74 (0.07)
Semantic proximity	4.81 (0.53)	4.59 (0.42)	4.35 (0.53)	4.49 (0.48)
Absolute number of switches	17.69 (4.92)	15.95 (3.93)	17.50 (4.80)	16.59 (4.31)

Variables that were actually included in the statistical analyses are printed in italics.

Table 4a
Correlation matrix between the variables referring to the type of produced words that were submitted to the statistical analysis

	Repeated words/total words	Subcategories	Subcategories/total words
New words	−0.341	0.604	−0.415
Repeated words/total words		−0.281	−0.468
Subcategories			0.252

Table 4b
Correlation matrix between the variables referring to the order dimension that were submitted to statistical analysis

	Relative switching	Semantic proximity
Order index	−0.937	0.796
Relative switching		−0.892

reported as the F approximation of the Wilks' Criterion. For the first domain (New Words and Repetitions/Total Words) the Group factor yielded an F-value of 3.627 with df = 4.126 ($P = 0.0078$). For the second domain (Subcategories and Subcategories/Total Words), the F-value with df = 4.126 was <1 (n.s.). For the third domain (Order Index, Semantic Proximity and Relative number of Switches) the F-value (df = 6.124) was 2.663 ($P = 0.0183$). This means that the number of new words and the ratio of repeated words to total number of words differed between groups, and the same conclusion applies to the order indices, whereas the number of subcategories and their ratio to the total number of words did not. Accordingly, the multiple comparisons are only warranted for the dimensions "number of produced words" and "order".

Tables 5a and 5b show the outcome of the multiple comparisons concerning the variables of the two domains for which the group effect was significant.

The outcome of the analysis is quite clear. Both frontal groups produced fewer New Words than Controls, without any difference between lateral and mesial lesion groups. The percentage of repetitions was not significantly different between groups, showing only a trend towards being higher in the mesial frontal group. The other domain of variables, which evaluates the semantic relatedness between the successively produced words, showed a consistent and robust

Table 5a
Multiple comparisons on the number of new words and the ratio of repeated words to total words

	Controls vs. frontal patients (overall)	Controls vs. frontal mesial patients	Controls vs. frontal lateral patients	Frontal patients: mesial vs. lateral
New words	**0.174** (0.105)	**0.147** (0.105)	**0.105** (0.105)	0.001 (0.127), 0.002 (0.229)
Repeated words/total words	0.083 (0.105)	0.079 (0.105)	0.040 (0.105)	0.003 (0.127), 0.004 (0.229)

Note: For each comparison we report the observed point of the Generalised Beta Distribution yielded by the analysis and, in parenthesis, the significance threshold. The comparison between frontal subgroups is reported also after a separate analysis that did not include controls, and could therefore be adjusted also for the length of illness (second value after the comma). The type I error risk was 0.05 for the whole table. The hypothesis was considered unidirectional for the comparisons between Controls and Frontal patients. Significant comparisons are printed in bold. Trends (see statistical methods) are underlined.

Table 5b
Multiple comparisons concerning the variables pertaining to the "order" dimension: order index, semantic proximity and relative number of Switches

	Controls vs. frontal patients (overall)	Controls vs. frontal mesial patients	Controls vs. frontal lateral patients	Frontal patients: mesial vs. lateral
Order index	0.145 (0.155)	0.046 (0.155)	**0.163** (0.155)	0.047 (0.155), 0.115 (0.247)
Semantic proximity	0.094 (0.155)	0.018 (0.155)	0.126 (0.155)	0.052 (0.155), 0.148 (0.247)
Relative number of switches	0.151 (0.155)	0.035 (0.155)	**0.191** (0.155)	0.052 (0.155), 0.178 (0.247)

See Table 5a for details on the reported statistics

Table 6
Effect size of the significant comparisons and for the trends. The values reported in the table refer to Cohen's d index. The effect size is considered large for $d > 0.80$, medium for $0.50 < d < 0.80$, and small for $0.20 < d < 0.50$. n.s. denotes the comparisons that were originally not significant

	Controls vs. frontal patients (overall)	Controls vs. frontal mesial patients	Controls vs. frontal lateral patients	Frontal patients: mesial vs. lateral
New words	1.07	1.12	0.52	n.s.
Repeated words/total words	0.62	0.71	n.s.	n.s.
Order index	0.75	n.s.	0.96	n.s.
Semantic proximity	0.63	n.s.	0.87	0.51
Relative number of switches	0.86	n.s.	1.18	0.71

pattern. The semantic "chain" was less structured for the frontal lateral group, which significantly differed from controls on two variables concerning this dimension: Order Index was decreased compared to controls, and Relative Switching was increased. Moreover, this group showed a robust analogous trend for the third variable, average Semantic Proximity, which was decreased. Finally, as can be seen from Table 3, the lateral frontal group did not have a lower absolute number of switches. Mesial frontal patients were intermediate between controls and lateral frontal patients, and never significantly differed from either group.

Beyond the significance values, the above findings were confirmed by a study of the effect size according to Cohen (1988) *d* index: effect size is considered large for $d > 0.80$, medium for $0.50 < d < 0.80$, and small for $0.20 < d < 0.50$. This analysis confirmed that the effect size was large for almost all the comparisons that were fully significant on multiple comparisons (Tables 5a and 5b) with the exception of the contrast between controls and frontal lateral patients on New Words (medium effect). Whenever we found a trend, the effect was at least medium, and in some cases it was even large. The actual *d* values are reported on Table 6.

4. Discussion

This experiment focused on some variables aimed at describing the *qualitative* structure of semantic fluency: (a) three variables sensitive to the order of produced words and to the semantic relatedness of the pairs of successively produced words, and (b) the number of produced subcategories (based on a stringent definition of subcategory).

The study aimed to contrast three possible causes for the reduced semantic fluency of frontal patients: (i) a strategy loss, (ii) a switching deficit and (iii) an initiation deficit. In the case of a strategy deficit we would have expected a *decrease* of the order of the word sequence (thus, a decrease of Order Index and average Semantic Proximity and an increase of Relative Switching) but a *normal* (or even increased) number of subcategories. In the case of a pure switching deficit we would have expected an *increase* of the sequence order and a *decrease* of both the number and the proportion of subcategories explored. Finally, in the case of an isolated initiation deficit we would have expected to find a *decrease* of both the number of produced words and the raw number of subcategories, but a normal proportion of subcategories with respect to the total words. In the latter case, the sequence production slows down, but the sequence order should not be affected. Given the presumed functional heterogeneity of different frontal regions, we also checked whether distinct patterns of production were detectable in the groups with damage restricted either to the mesial or lateral aspects of the frontal lobes.

The results of our investigations provide some answers to these questions, and the pattern resulting from our analysis indicates that, as presumed at the outset, the outcome of frontal lesions changes according to which frontal region is affected. Therefore, we will first comment on the results separately for each frontal lesion subgroup and will subsequently address more general questions regarding the meaning of switching and strategy deficits.

4.1. Frontal mesial subgroup

Mesial frontal patients produced fewer words than the control group, but the three order measures showed that the overall structure of the word sequence was not significantly different from that of controls. Nor was the number of subcategories explored by mesial frontal patients significantly lower. This outcome does not conform to any of the patterns predicted for our three possible causes. Regarding the order measures, the mesial frontal group means are intermediate between controls and the lateral frontal group. However, this group shows a trend towards producing sequences *more ordered* than the lateral frontal patients, but no trend towards being *less ordered* with respect to normal subjects. By itself, normally ordered sequences are compatible with an initiation deficit hypothesis, but we cannot safely endorse this conclusion because the mesial frontal group did not produce a significantly lower number of subcategories: the comparison between control and mesial frontal groups for the raw number of subcategories measure only approached significance ($P = 0.066$, one tailed) when judged with the minimal protection ("hypothesis-wise") against type I error.

On the whole, we think that a loss of activation is much more likely than a disruption of the search strategy to be the right explanation for the fluency impairment of this group, but for reaching a definite conclusion a more powerful experiment is in order.

4.2. *Lateral frontal subgroup*

The frontal lateral subgroup generated a lower number of words while the number of subcategories did not significantly differ from controls. However, at variance with the mesial subgroup, their word sequences were significantly less organized than the sequences of the controls: Order Index was decreased and Relative number of Switches was increased with respect to the Control group, and Semantic Proximity showed a trend toward decrease. Finally, they did not produce significantly more repetitions than the Control Group. This pattern clearly conforms to the predictions of the strategic deficit hypothesis. The number of produced subcategories was very similar to that of control subjects, and this makes it very unlikely that we are in the presence of a failed rejection of the null hypothesis.

4.3. *Semantic fluency and switching deficit*

One of the better known and documented phenomena presented by patients with frontal lesions is perseveration behaviour, i.e. the repetition of former responses and the inability to switch a response pattern to meet changes in task demands (Milner, 1964; Owen et al., 1991; Stuss et al., 2000; Warrington, 2000). On this basis, a switching deficit has been hypothesised to follow lesion to frontal lobes, particularly to their lateral aspect (Stuss & Levine, 2002). In agreement with this theoretical expectation, a study on semantic fluency by Troyer and collaborators (1998) argued that switching deficit is one of the leading causes of the frontal impairment on semantic fluency. In the present study, however, lateral frontal patients did not repeat significantly more words than controls, and the qualitative aspects of semantic fluency in lateral frontal patients did not reveal the presence of any switching deficit. Should this outcome be considered in conflict with the theory and previous findings?

First, we need to discuss the possible claim that if we had asked participants to avoid uttering repetitions we might have observed a difference in the repetition index between frontal patients and controls. We agree that this is a possibility. However, the latter policy would have blurred the interpretation of the hypothetical increase of repetitions. In fact, when a participant repeatedly utters the same word, two serial events are necessary: (i) an already retrieved word continues to “pop into mind” at a rate higher than normal, and this component is active in the phase of semantic-lexical search; (ii) the participant is unaware that the actually retrieved word is a repetition and the utterance is delivered as if it were a new word. The latter phenomenon can be interpreted as a monitoring deficit. The very contribution of the “pop in mind” procedure is to discriminate the origin of repetitions, since the monitoring deficit component would be excluded by the explicit requirement to “say aloud repetitions as well”. As such, given that we were interested in the repetitions that arise at the semantic-lexical level (i.e. at the level of the search process), the “pop in mind” procedure cannot be considered responsible for the missed detection of experimental effects at this level. On the contrary, the lack of significant differences in word repetition between Controls and the Lateral Frontal Group suggests that the latter group does not show perseverations at the semantic-lexical level.

Secondly, the majority of tests used to induce and document perseveration errors (not simply an increased switching time cost) are rule attainment tasks such as the Wisconsin Card Sorting Test (Stuss & Levine, 2002; Stuss et al., 2000) and other instruments modelled on it (Burgess & Shallice, 1996a; Delis, Squire, Bihrle, & Massman, 1992; Owen et al., 1993; Owen et al., 1991). In this family of tests participants are required to sort the presented stimuli according to one of several possible criteria (e.g. for the WCST, colour, shape and number). At a given point, participants must abandon the former criterion and shift to a new one according to a change in the feedback provided by the examiner. Critically, the criterion or cognitive set to be abandoned in favour of a new one would probably have received, before the switching request, an amount of positive reinforcement that could be substantial (e.g. in the WCST version in the paper by Stuss et al., 2000 the criterion changes after ten consecutive correct responses). It might be that a high activation of a particular cognitive set is *mandatory* in order to induce perseveration errors. An empirical finding consistent with our results and with the above discussed hypothesis is the absence of perseveration errors in the Brixton task (Burgess & Shallice, 1996a; Reverberi et al., 2005). The Brixton task is a rule attainment test modelled on the WCST but which uses a less strong reinforcement schedule and more abstract rules. The settings in the semantic fluency task are even less prone than those of the Brixton task to induce perseveration behaviour, at least if the category “fruit” is used. In a fluency test, a reasonable measure of the activation of a given subcategory at any step (s) of the test is the number of times the subcategory active at (s-1) has been consecutively accessed or used. In our experiment the average cluster size for the control group – not reported in the Results – was 2.37 (starting the count from the *first* word of a cluster), and therefore the average subcategory activation during the fluency test might not have been strong enough to trigger a burst of perseverative responses. It could be that in a fluency task based on semantic categories endowed with larger semantic clusters and stronger subcategory internal correlations, frontal patients might produce a larger number of perseveration responses.

Third, Troyer et al. (1998) found that frontal patients had a significantly lower *absolute* switching score than controls, which they explained as a sign of a cognitive-shifting deficit during the task. There are both interpretation and empirical differences between the present study and the one by Troyer and collaborators:

(i) *Interpretation differences*. The measure of switching deficit used by Troyer and collaborators – the *absolute* number of times a participant departs from a particular subcategory – is sensitive also to initiation deficit, another cognitive deficit that can affect patients with frontal lobe damage. Therefore, the pattern found by Troyer and collaborators could also be explained by reference purely to an initiation deficit (see also Mayr, 2002), particularly if we consider that only one minute was granted to participants to accomplish the task. The interpretation uncertainty between a deficit of switching or initiation, however, does not apply to the decrease of the *relative* switch index used in the present study.

Moreover, it can be easily verified that the relative number of switches is closely related to the reciprocal of the average cluster size (see online supplementary material for further detail), another variable used by Troyer et al. (1998). On the basis of this correspondence, we can say that either our frontal lateral patients presented an increase of relative switches or they presented a decrease of average cluster size, and these conclusions would have the same empirical content. Theoretically, we have interpreted an increase of relative switching (i.e. a decrease of average cluster size) in a frontal damaged sample as a sign of strategy deficit, while Troyer and collaborators considered the decrease of average cluster size only a hallmark of temporal lobe damage.

(ii) *Empirical differences*. Troyer and collaborators (1998) found no significant differences in either frontal or temporal groups from the Control Group for size of semantic clusters. This is not consistent with the findings of our study, where lateral frontal patients had a *higher* relative switching index (i.e. a *lower* average cluster size) than controls.

The inconsistencies between the experiments could depend in part on the different procedures followed in our study and in the work by Troyer and colleagues: the longer duration of the tests (3 minutes versus 1) and the use of a smaller category (“fruit” instead of “animals”) probably increased, in our case, the need for an efficient strategy during the fluency task. Actually, an efficient strategy may be more crucial in a later stage of word production when highly familiar or prototypical items are no longer available. Thus, the empirical difference could be partly due to an increased sensitivity of our fluency setting to strategy deficit.

4.4. Conclusions

Our study suggests that frontal lateral patients are affected by a deficit of strategy and search organisation while exploring the lexical/semantic system. The neuropsychological literature has reported strategic deficits of frontal patients in several experiments on episodic memory (e.g. Gershberg & Shimamura, 1995; Stuss et al., 1994), on working memory (e.g. Owen, Morris, Sahakian, Polkey, & Robbins, 1996), and also on tasks devoid of an overt memory component (e.g. Shallice & Burgess, 1991). The present study extends the available evidence in two ways: (i) it documents a strategic deficit during a pure retrieval task such as semantic fluency, and (ii) it shows that the lateral frontal lobes are one of the crucial anatomical substrates for the creation and implementation of an efficient retrieval strategy.

Acknowledgements

Research partly supported by a MIUR Grant to E.C. Rosemary Allpress revised the English text.

Appendix A. Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.neuropsychologia.2005.05.011.

References

Alexander, M. P., Stuss, D. T., & Fansabedian, N. (2003). California verbal learning Test: Performance by patients with focal frontal and non-frontal lesions. *Brain*, *126*(Pt 6), 1493–1503.

Baldo, J. V., & Shimamura, A. P. (1998). Letter and category fluency in patients with frontal lobe lesions. *Neuropsychology*, *12*(2), 259–267.

Bousfield, W. A., & Sedgewick, H. (1944). An analysis of sequences of restricted associative responses. *Journal of General Psychology*, *52*, 83–95.

Burgess, P. W., & Shallice, T. (1996a). Bizarre responses, rule detection and frontal lobe lesions. *Cortex*, *32*(2), 241–259.

Burgess, P. W., & Shallice, T. (1996b). Response suppression, initiation and strategy use following frontal lobe lesions. *Neuropsychologia*, *34*(4), 263–272.

Cohen, J. (1988). *Statistical power analysis far the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Delis, D. C., Squire, L. R., Bihrle, A., & Massman, P. (1992). Componential analysis of problem-solving ability: Performance of patients with frontal lobe damage and amnesic patients on a new sorting test. *Neuropsychologia*, *30*(8), 683–697.

Fletcher, P. C., Shallice, T., & Dolan, R. J. (2000). Sculpting the response space—an account of left prefrontal activation at encoding. *Neuroimage*, *12*(4), 404–417.

Frith, C. D., Friston, K. J., Liddle, P. F., & Frackowiak, R. S. (1991). A PET study of word finding. *Neuropsychologia*, *29*(12), 1137–1148.

Gershberg, F. B., & Shimamura, A. P. (1995). Impaired use of organizational strategies in free recall following frontal lobe damage. *Neuropsychologia*, *33*(10), 1305–1333.

Godefroy, O., Lhullier-Lamy, C., & Rousseaux, M. (2002). SRT lengthening: Role of an alertness deficit in frontal damaged patients. *Neuropsychologia*, *40*(13), 2234–2241.

Gruenewald, P. J., & Lockhead, G. R. (1980). The free recall of category examples. *Journal of Experimental Psychology: Human Learning and Memory*, *6*(3), 225–240.

Gurd, J. M., Amunts, K., Weiss, P. H., Zafiris, O., Zilles, K., Marshall, J. C., et al. (2002). Posterior parietal cortex is implicated in continuous switching between verbal fluency tasks: An fMRI study with clinical implications. *Brain*, *125*(Pt 5), 1024–1038.

Henry, J. D., & Crawford, J. R. (2004). A meta-analytic review of verbal fluency performance following focal cortical lesions. *Neuropsychology*, *18*(2), 284–295.

Henson, R. N., Shallice, T., & Dolan, R. J. (1999). Right prefrontal cortex and episodic memory retrieval: A functional MRI test of the monitoring hypothesis. *Brain*, *122*(7), 1367–1381.

Incisa della Rocchetta, A., & Milner, B. (1993). Strategic search and retrieval inhibition: The role of the frontal lobes. *Neuropsychologia*, *31*(6), 503–524.

Luu, P., Flaisch, T., & Tucker, D. M. (2000). Medial frontal cortex in action monitoring. *Journal of Neuroscience*, *20*(1), 464–469.

Mayr, U. (2002). On the dissociation between clustering and switching in verbal fluency: Comment on Troyer, Moscovitch, Winocur, Alexander and Stuss. *Neuropsychologia*, *40*(5), 562–566.

Milner, B. (1964). Effects of different brain lesions on card sorting: The role of the frontal lobes. *Archives of Neurology*, *9*, 100–110.

Moscovitch, M. (1992). Memory and working-with-memory: A component process model based on modules and central systems. *Journal of Cognitive Neuroscience*, *4*(3), 257–267.

Owen, A. M., Downes, J. J., Sahakian, B. J., Polkey, C. E., & Robbins, T. W. (1990). Planning and spatial working memory following frontal lobe lesions in man. *Neuropsychologia*, *28*(10), 1021–1034.

Owen, A. M., Morris, R. G., Sahakian, B. J., Polkey, C. E., & Robbins, T. W. (1996). Double dissociations of memory and executive functions in working memory tasks following frontal lobe excisions, temporal lobe excisions or amygdalo-hippocampectomy in man. *Brain*, *119*(Pt 5), 1597–1615.

Owen, A. M., Roberts, A. C., Hodges, J. R., Summers, B. A., Polkey, C. E., & Robbins, T. W. (1993). Contrasting mechanisms of impaired attentional set-shifting in patients with frontal lobe damage or Parkinson's disease. *Brain*, *116*(Pt 5), 1159–1175.

Owen, A. M., Roberts, A. C., Polkey, C. E., Sahakian, B. J., & Robbins, T. W. (1991). Extra-dimensional versus intra-dimensional set shifting performance following frontal lobe excisions, temporal lobe excisions or amygdalo-hippocampectomy in man. *Neuropsychologia*, *29*(10), 993–1006.

Reverberi, C., Laiacona, M., & Capitani, E. (2004). Variabili semantico-lessicali relative a tutti gli elementi di una categoria semantica: Indagine su soggetti normali italiani per la categoria "frutta". *Giornale Italiano di Psicologia*, *31*(3), 497–522.

Reverberi, C., Lavaroni, A., Gigli, G. L., Skrap, M., & Shallice, T. (2005). Specific impairments of rule induction in different frontal lobe subgroups. *Neuropsychologia*, *43*(3), 460–472.

Rosen, V. M., & Engle, R. W. (1997). The role of working memory capacity in retrieval. *Journal of Experimental Psychology: General*, *126*(3), 211–227.

Roy, S. N., & Bose, R. C. (1953). Simultaneous confidence interval estimation. *Annals of Mathematical Statistics*, *24*, 513–536.

Shallice, T., & Burgess, P. W. (1991). Deficits in strategy application following frontal lobe damage in man. *Brain*, *114*(Pt 2), 727–741.

Stuss, D. T., Alexander, M. P., Hamer, L., Palumbo, C., Dempster, R., Binns, M., et al. (1998). The effects of focal anterior and posterior brain lesions on verbal fluency. *Journal of the International Neuropsychological Society*, *4*(3), 265–278.

Stuss, D. T., Alexander, M. P., Palumbo, C. L., Buckle, L., et al. (1994). Organizational strategies with unilateral or bilateral frontal lobe injury in word learning tasks. *Neuropsychology*, *8*(3), 355–373.

Stuss, D. T., & Levine, B. (2002). Adult clinical neuropsychology: Lessons from studies of the frontal lobes. *Annual Review of Psychology*, *53*, 401–433.

Stuss, D. T., Levine, B., Alexander, M. P., Hong, J., Palumbo, C., Hamer, L., et al. (2000). Wisconsin Card Sorting Test performance in patients with focal frontal and posterior brain damage: Effects of lesion location and test structure on separable cognitive processes. *Neuropsychologia*, *38*(4), 388–402.

Sylvester, C. Y., & Shimamura, A. P. (2002). Evidence for intact semantic representations in patients with frontal lobe lesions. *Neuropsychology*, *16*(2), 197–207.

Troyer, A. K., Moscovitch, M., & Winocur, G. (1997). Clustering and switching as two components of verbal fluency: Evidence from younger and older healthy adults. *Neuropsychology*, *11*(1), 138–146.

Troyer, A. K., Moscovitch, M., Winocur, G., Alexander, M. P., & Stuss, D. (1998). Clustering and switching on verbal fluency: The effects of focal frontal- and temporal-lobe lesions. *Neuropsychologia*, *36*(6), 499–504.

Warburton, E., Wise, R. J., Price, C. J., Weiller, C., Hadar, U., Ramsay, S., et al. (1996). Noun and verb retrieval by normal subjects. Studies with PET. *Brain*, *119*(Pt 1), 159–179.

Warrington, E. K. (2000). Homophone meaning generation: A new test of verbal switching for the detection of frontal lobe dysfunction. *Journal of the International Neuropsychological Society*, *6*(6), 643–648.

Wixted, J. T., & Rohrer, D. (1994). Analyzing the dinamic of free recall: An integrative review of the empirical literarure. *Psychonomic Bulletin and Review*, *1*, 89–106.