Qualitative Assessment of Verbal Fluency Performance in Frontotemporal Dementia

Esther van den Berg a, Lize C. Jiskoot a, b, Mariëlle J.H. Grosveld a, John C. van Swieten a, c, Janne M. Papma a

a Alzheimer Center and Department of Neurology, Erasmus University Medical Center, Rotterdam, b Department of Radiology, Leiden University Medical Center, Leiden, and c Department of Clinical Genetics, VU University Medical Center, Amsterdam, The Netherlands

Keywords
Primary progressive aphasia · Executive control · Semantic memory · Dementia

Abstract
Background/Aims: Verbal fluency is impaired in patients with frontotemporal dementia (FTD) and primary progressive aphasia (PPA). This study explored qualitative differences in verbal fluency (clustering of words, switching between strategies) between FTD and PPA variants. Methods: Twenty-nine patients with behavioral variant FTD (bvFTD) and 50 with PPA (13 nonfluent/agrammatic, 14 semantic, and 23 logopenic) performed a semantic and letter fluency task. Clustering (number of multiword strings) and switching (number of transitions between clustered and nonclustered words) were recorded by two independent raters. Between-group differences, associations with memory, language, and executive functioning, and longitudinal change (subsample) in clustering and switching were examined. Results: Interrater reliability was high (median 0.98). PPA patients generated (a) smaller (number of) clusters on semantic and letter fluency than bvFTD patients ($p < 0.05$). Semantic variant patients used more switches than nonfluent/agrammatic or logopenic variant patients ($p < 0.05$). Clustering in semantic fluency was significantly associated with memory and language (range standardized regression coefficients 0.24–0.38). Switching in letter fluency was associated with executive functioning (0.32–0.35). Conclusion: Clustering and switching in verbal fluency differed between patients with subtypes of FTD and PPA. Qualitative aspects of verbal fluency provide additional information on verbal ability and executive control which can be used for clinically diagnostic purposes.
Introduction

Frontotemporal dementia (FTD) is an early-onset dementia characterized by a heterogeneous clinical presentation including behavioral changes, fronto-executive deficits, and/or selective language disorders [1] caused by pathophysiological changes in the frontal and temporal lobes [2, 3]. Patients with the behavioral variant FTD (bvFTD) typically present with personality and behavioral changes such as inappropriate social conduct, apathy, or disinhibition [4, 5]. The language variants are collectively referred to as primary progressive aphasia (PPA) and include a semantic variant (characterized by multimodal loss of semantic knowledge), a nonfluent/agrammatic variant (characterized by agrammatism and apraxia of speech), and the more recently defined logopenic variant (characterized by a slowed rate of speech and impaired sentence repetition), which is also associated with Alzheimer-type pathology [1, 6].

The clinical diagnosis of FTD and PPA is aided by detailed neuropsychological examination including measures of language, executive functioning, and memory. Executive control is often evaluated by tests of verbal fluency [7, 8]. Verbal fluency refers to the ability to generate words from a semantic category (e.g., animals, groceries, fruits) or a specified letter of the alphabet within a limited time. Verbal fluency tasks are easy to administer, and the number of correct responses provides a rich source of information on semantic, lexical, and executive cognitive processes [9]. Based on the differences in clinical subtypes of FTD and PPA, a disparate pattern of impairment in verbal fluency tasks is expected. Indeed, a previous study showed that patients with semantic variant PPA were more impaired on semantic than letter fluency, whereas patients with bvFTD and nonfluent/agrammatic variant PPA were equally impaired in both types of tasks [8].

In this study, we not only aim to examine quantitative differences in verbal fluency in FTD and PPA, but also focus on potential qualitative differences. Qualitative aspects of lexical retrieval, such as the clustering of related words and switching between categories, can offer additional insights into the cognitive processes of fluency task performance in specific patient groups, especially into the decay of semantic knowledge [10] and executive control [11]. The diagnostic utility of clustering and switching in verbal fluency has been shown for different patient groups such as those with traumatic brain injury [12] and schizophrenia [13], but it is unknown for FTD. Also, whether measures of clustering and switching provide additional information on related cognitive functions such as memory, language, processing speed, or executive functioning in these patients is unclear. Qualitative verbal fluency measures may have prognostic value, already demonstrated in the prediction of conversion from mild cognitive impairment (MCI) to Alzheimer disease (AD) [14], but whether disease progression in FTD could be predicted by clustering and switching measures remains to be evaluated.

The present study thus aimed to examine clustering and switching in semantic and letter fluency in patients with bvFTD and those with PPA. Additionally, associations with other cognitive functions, as well as longitudinal changes in fluency measures in a subsample of patients, were evaluated.

Subjects and Methods

Participants

This retrospective study included 79 patients who visited the Alzheimer Center of the Erasmus University Medical Center, Rotterdam, The Netherlands, between January 2012 and December 2015 and were diagnosed with bvFTD or PPA. The clinical diagnosis was made in a multidisciplinary consensus meeting using international diagnostic consensus criteria. For bvFTD (n = 29), the diagnostic criteria of Rascovsky et al. [5] were used. The diagnosis of patients with PPA (14 semantic, 13 nonfluent/agrammatic, and 23 logopenic) was made according to the International Working Group on Primary Progressive Aphasia.
Verbal Fluency Measures

All participants performed a semantic and letter fluency task as part of a standardized neuropsychological assessment (see below). For the semantic fluency task, the participants were asked to generate as many exemplars as possible from the category of animals in 60 s. For the letter fluency task, the participants were asked to generate as many different words as possible beginning with the letter D, then A, and then T, with 60 s allowed for each letter. The letters D, A, and T were considered the Dutch equivalent to the letters F, A, and S [16]. At follow-up, the letters K, O, and M or P, G, and R were used as parallel versions. Participants were instructed not to generate proper names or a previously generated word with only a different suffix. The total number of correct animals and the total number of correct words generated for the letters D, A, and T were recorded. Impaired performance of individual patients was defined as a T-score ≤27 compared to age-, sex-, and education-adjusted normative data (corresponding to ≤1st percentile). In addition, the semantic and letter fluency performance of each participant was reevaluated by means of the scoring system reported by Ledoux et al. [17]; adaptation of [11]. Table 1 shows a summary of the scoring guidelines; further details on the scoring procedure for clustering and switching are reported elsewhere [17]. Based on this scoring procedure, the following measures were scored for the semantic and letter fluency performance of each participant:

1. Total number of words: sum of all words produced, excluding repetitions and rule breaks; for letter fluency, the sum of the letters D, A, and T was used
2. Number of clusters: number of multiword strings; each cluster contained at least 2 successive words
3. Number of switches: number of transitions between clustered or nonclustered words (i.e., switches from one associative strategy to another or to none at all)
4. Total cluster size: sum of all clustered words
5. Mean cluster size: total cluster size divided by the number of clusters

Other Cognitive Functions

The standardized neuropsychological assessment included validated tests covering the major cognitive domains. Memory was assessed with the Dutch version of the Rey Auditory Verbal Learning Test (RAVLT [18]; immediate recall, delayed recall, delayed recognition), the story recall subtest from the Rivermead Behavioural Memory Test [19], and the visual association test [20]. Language was examined with the 60-item Boston Naming Test [21]. Information processing speed was assessed with the letter digit substitution test [22], the trail-making test part A [23], and the Stroop Color and Word Tests I and II [24]. Attention and executive functioning were assessed with the trail-making test B/A ratio, the Stroop Color and Word Test III/II ratio, the modified Wisconsin Card Sorting Test (mWCST [25]; number of concepts), the digit span subtest of the Wechsler Adult Intelligence Scale, ed 3 (WAIS-III [26]; total score), and the zoo map and key search subtests of the Behavioural Assessment of the Dysexecutive Syndrome (BADS [27]). Raw test scores were converted to standardized z-scores based on the mean and standard deviation of the study population. The z-scores were then averaged per cognitive domain into a composite domain score.

Statistical Analysis

Statistical analyses were performed using SPSS Statistics 21.0 (IBM Corp., Armonk, NY, USA). Between-group differences were analyzed with analysis of variance for continuous data, Mann-Whitney U tests for...
Results

Demographics

The characteristics of the patient sample are presented in Table 2. The subgroups differed significantly in age ($F(3, 75) = 3.6, p < 0.05, \eta^2 = 0.13$) and gender ($\chi^2(3) = 12.4, p < 0.01$), with higher age and more women in the logopenic group than in the other groups. The groups did not differ in educational level, disease duration, or FAB score (all $p > 0.05$), but the logopenic group had a lower MMSE score than the bvFTD and the nonfluent/agrammatic variant patients ($F(3, 71) = 6.5, p < 0.01, \eta^2 = 0.22$).

Interrater Reliability

Two independent raters, blinded to patient status, scored the fluency clustering and switching parameters (see Subjects and Methods) based on the scoring system described in Ledoux et al. [17]. Table 3 shows high levels of interrater reliability for both the semantic (range 0.90–1.0) and the letter fluency variables (range 0.80–0.99).
Baseline Verbal Fluency Performance

In semantic fluency, 46% of the patients showed impaired performance compared with the normative data (i.e., a T-score ≤27). Patients with PPA produced a significantly lower total number of words than patients with bvFTD \(F(1, 75) = 5.8, p < 0.05, \eta^2 = 0.07\). In addition, PPA patients produced smaller semantic clusters than patients with bvFTD (Table 2; total cluster size: \(F(1, 74) = 6.3, \eta^2 = 0.08\); mean cluster size: \(F(1, 74) = 6.3, p < 0.05, \eta^2 = 0.08\)). Post hoc comparisons within the PPA group showed no significant differences in the fluency variables between patients with semantic, patients with nonfluent/agrammatic, and patients with logopenic variants (all \(p > 0.05\)).
In letter fluency, 29% of the patients showed impaired performance compared with the normative data (i.e., a T-score ≤ 27). The total number of words generated did not differ between patients with PPA and those with bvFTD ($F(1, 74) = 1.5, p = 0.22, \eta^2 = 0.02$). However, patients with PPA produced a lower number of clusters and total cluster size than patients with bvFTD (Table 2; $F(1, 70) = 3.9, p < 0.05, \eta^2 = 0.05; F(1, 70) = 4.1, p < 0.05, \eta^2 = 0.06$). Post hoc comparisons within the PPA groups showed that patients with semantic variant PPA produced significantly more words ($F(2, 43) = 3.7, p < 0.05, \eta^2 = 0.15$) but also more switches ($F(2, 43) = 8.3, p < 0.01, \eta^2 = 0.28$, adjusted for total number of words) than patients with nonfluent/agrammatic and logopenic variants (Fig. 1).

**Change over Time in Verbal Fluency Performance**

Nineteen patients (3 bvFTD, 6 semantic, 4 nonfluent/agrammatic, and 6 logopenic) performed the semantic and letter fluency task at follow-up. For semantic fluency, the patient group as a whole showed a significant decline in total number of words ($F(1, 18) = 9.8, p < 0.01, \eta^2 = 0.35$) and cluster size ($F(1, 18) = 5.0, p < 0.05, \eta^2 = 0.32$). For letter fluency, the patient group as a whole showed a significant decline in total number of words ($F(1, 18) = 5.7, p < 0.05, \eta^2 = 0.24$), number of clusters ($F(1, 18) = 8.4, p < 0.05, \eta^2 = 0.32$), number of switches ($F(1, 18) = 10.3, p < 0.01, \eta^2 = 0.37$), and cluster size ($F(1, 18) = 7.3, p < 0.05, \eta^2 = 0.29$).

Examination of the time × group (bvFTD vs. PPA) interaction for semantic fluency showed that patients with PPA had a greater decline in total number of words ($F(1, 17) = 5.3, p < 0.05, \eta^2 = 0.24$) and number of clusters ($F(1, 17) = 4.2, p < 0.05, \eta^2 = 0.20$) than patients with bvFTD. For letter fluency, no significant time × group interactions were observed.

**Association with Other Cognitive Functions**

The results of the regression analysis (Table 4) showed that all but one measure of semantic and letter fluency correlated significantly with information processing speed (range standardized regression coefficient $\beta$: 0.14–0.54). Furthermore, for semantic fluency the total number of words, the number of clusters, and the cluster size showed significant associations with the domain scores for language and memory ($\beta$ 0.24–0.38). For letter fluency, the total
number of words and the number of switches correlated significantly with attention and executive functioning \((\beta 0.32–0.35)\).

**Discussion**

The present study examined clustering and switching in verbal fluency in patients with bvFTD and PPA. We found both quantitative and qualitative differences in verbal fluency between these patient groups. Patients with PPA not only generated fewer words on semantic fluency than patients with bvFTD, but they also produced fewer and smaller clusters on semantic and letter fluency. These findings fit well with the theoretical notion that clustering relies on the generation of semantic, associative, or letter-related words and is thus mainly supported by the integrity of the semantic system. Damage to the left temporal lobe, as is present in PPA, evidently impairs access to lexical information and leads to deterioration of semantic storage \([28]\).

Our results also showed differences between subtypes of PPA. Patients with the semantic variant produced more words than patients with the nonfluent/agrammatic and logopenic variants. The number of words generated in letter fluency in semantic variant PPA patients was comparable to that generated in the bvFTD group, reflecting the uncompromised fluency of speech in semantic variant PPA. However, semantic variant patients showed significantly more switches than patients with the nonfluent/agrammatic or logopenic variant. This relative increase in the number of switches in semantic variant PPA may reflect the combination of both intact fluency of speech and deterioration of semantic knowledge. Switching is thought to invoke a deliberate, controlled search and is associated with executive functioning \([17]\). It is mainly supported by frontal lobe functioning, which is relatively spared in semantic variant PPA, at least in the early stages of the disease.

As would be expected, patients with nonfluent/agrammatic PPA showed the lowest output in both semantic and letter fluency, which is thought to result primarily from speech apraxia \([6, 8, 29–31]\). The cluster sizes in patients with nonfluent/agrammatic PPA were, however, comparable to those found in patients with semantic and logopenic variant PPA. Whether the PPA subtypes might rely on different strategies to counter degradation of

---

**Table 4. Relation between measures of verbal fluency and cognitive domain scores**

<table>
<thead>
<tr>
<th></th>
<th>Language</th>
<th>Memory</th>
<th>Processing speed</th>
<th>Attention and executive functioning</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Semantic fluency</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of words</td>
<td>0.35**</td>
<td>0.24*</td>
<td>0.48**</td>
<td>0.24*</td>
</tr>
<tr>
<td>Number of clusters</td>
<td>0.25*</td>
<td>0.25*</td>
<td>0.50**</td>
<td>0.22</td>
</tr>
<tr>
<td>Number of switches</td>
<td>0.21</td>
<td>-0.02</td>
<td>0.35*</td>
<td>0.22</td>
</tr>
<tr>
<td>Total cluster size</td>
<td>0.26*</td>
<td>0.38**</td>
<td>0.47**</td>
<td>0.17</td>
</tr>
<tr>
<td>Mean cluster size</td>
<td>0.07</td>
<td>0.26</td>
<td>0.14</td>
<td>-0.04</td>
</tr>
<tr>
<td><strong>Letter fluency</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of words</td>
<td>0.14</td>
<td>0.03</td>
<td>0.55**</td>
<td>0.35**</td>
</tr>
<tr>
<td>Number of clusters</td>
<td>0.20</td>
<td>0.07</td>
<td>0.53**</td>
<td>0.20</td>
</tr>
<tr>
<td>Number of switches</td>
<td>0.05</td>
<td>-0.02</td>
<td>0.50**</td>
<td>0.32*</td>
</tr>
<tr>
<td>Total cluster size</td>
<td>0.23</td>
<td>0.08</td>
<td>0.54**</td>
<td>0.20</td>
</tr>
<tr>
<td>Mean cluster size</td>
<td>0.33*</td>
<td>-0.04</td>
<td>0.50**</td>
<td>0.11</td>
</tr>
</tbody>
</table>

Data are age- and sex-adjusted standardized regression coefficient \(\beta\) values. * \(p < 0.05\), ** \(p < 0.01\).
semantic knowledge (e.g., spelling or sound similarities vs. semantic similarities) could not be derived from the present analysis. There were no clear differences in clustering and switching between patients with nonfluent/agrammatic and those with logopenic variant PPA.

The diagnostic utility of verbal fluency in identifying subtypes of dementia has been examined by several previous studies [7, 14, 30, 32]. Although patients with FTD overall showed a worse fluency performance than patients with AD, the disparity between letter and semantic fluency is particularly effective in differentiating the two patient groups [7]. Moreover, AD can be distinguished from nonfluent/agrammatic and semantic variant PPA based on their profile of verbal fluency (number of words, word frequency, and type of errors) [30] and a recent study indicated that measures of lexical similarity in verbal fluency predict conversion of MCI into AD [14]. In line with these findings, the results of the present study indicate that subtypes of FTD can also be distinguished by qualitative inspection of clustering and switching measures. More specifically, analysis of the clustering and switching variables supports the semantic deficit in semantic variant PPA in contrast to the agrammatism and motor speech errors in nonfluent/agrammatic PPA, thereby aiding timely and accurate diagnosis.

The scoring method used in the present study, as described by Ledoux et al. [17], is based on the original scoring by Troyer et al. [11]. The interrater reliability of this method was adequate in the present study (median 0.98) and highly similar to the interrater reliability reported in the original study by Ledoux et al. [17]. Clustering and switching are considered robust theory-based fluency measures, a notion that is supported by data from patients with schizophrenia, HIV, Huntington disease, Parkinson disease, multiple sclerosis, and traumatic brain injury [12, 13, 33–36]. A recent study showed that machine learning in verbal fluency outperformed traditional structural MRI measures in predicting conversion from MCI into AD (area under the curve for MRI measures 0.76, vs. 0.87 for machine learning predictors) [14]. Although this machine learning approach may not be easily adopted in standard clinical care, our findings provide further support for the diagnostic value of verbal fluency tests in FTD. Moreover, it might be useful to examine the value of different types of relationships (i.e., phonological, orthographic, or semantic) as these are not separately recorded in the Ledoux scoring method. In this regard, inclusion of an AD comparison group could potentially yield additional information on these processes. Longitudinal data on verbal fluency in relation to onset or progression of dementia are scarce. Raoux et al. [37] and Auriaimbe et al. [38] showed that patients with AD already produce fewer words and switches 5 years before their clinical diagnosis. In addition, Pakhomov and Hemmy [39], using an automated approach, showed that in 239 cognitively intact subjects a larger semantic cluster size was associated with a 38 and 26% reduced dementia risk after 6 and 17 years, respectively. In the present study, clustering and switching in letter fluency showed a particularly strong decline in PPA patients, regardless of the subtype of PPA. Letter fluency thus appears to be a sensitive measure of cognitive decline in PPA.

Verbal fluency is supported by both common and distinct cognitive processes. Our results show a general association with measures of processing speed, which is in line with previous studies [40, 41] and probably reflects the timed nature of the task. Clustering in semantic fluency was significantly associated with memory and language, whereas in letter fluency switching was associated with executive functioning. These findings fit well within the theoretical model of verbal fluency proposed by Troyer et al. [11], in which clustering is thought to rely upon temporal lobe processes such as verbal memory and word storage, whereas switching relies upon frontal lobe processes such as strategic search processes, cognitive flexibility, and shifting.

Strengths of the present study include the relatively large patient sample and the detailed verbal fluency scoring method that showed a high interrater reliability. A limitation of this
study is the fact that the fluency data in this study were part of the diagnostic neuropsychological assessment. Patients with worse performance in semantic fluency are more likely to be classified as having semantic variant PPA than nonfluent/agrammatic PPA. The clustering and switching measures were, however, specifically calculated for this study, and were therefore not used for the diagnostic workup. This study included patients with a clinical diagnosis of bvFTD or PPA; autopsy-confirmed cases were unfortunately unavailable. This means that no pathological confirmation of the dementia subtype was possible. In addition, some patients with a clinical diagnosis of logopenic variant PPA may have underlying Alzheimer-type pathology. Disease duration is a difficult measure to assess in FTD. We estimated disease duration as the time between the first report of symptoms and the clinical assessment. It is possible that bvFTD is diagnosed in a later stage than PPA as the presenting symptoms (behavioral changes vs. language impairment) are more easily recognized in PPA. Given that the patients with bvFTD tended to outperform the PPA patients on most fluency measures, this potential bias in the estimation of disease duration is unlikely to have significantly influenced our results.

In sum, the present study showed differences in clustering and switching in verbal fluency between patients with PPA and those with bvFTD and a significant decline in verbal fluency over time. Clustering was specifically associated with memory and language, and switching with executive functioning. These results show that qualitative differences in verbal fluency provide additional information on different cognitive functions in FTD and PPA that can be used in clinical practice to improve diagnostic accuracy.

Disclosure Statement

The authors declare that they have no conflict of interest.

References
