BRAIN COMMUNICATIONS

Multi-model mapping of phonemic fluency

Lisa Cipolotti,[1,2] Tianbo Xu,[2] Bronson Harry,[3] Joe Mole,[1,2] Grace Lakey,[1] Tim Shallice,[4,5] Edgar Chan[1,2] and Parashkev Nachev[2]

The voluntary generation of non-overlearned responses is usually assessed with phonemic fluency. Like most frontal tasks, it draws upon different complex processes and systems whose precise nature is still incompletely understood. Many claimed aspects regarding the pattern of phonemic fluency performance and its underlying anatomy remain controversial. Major limitations of past investigations include small sample size, scant analysis of phonemic output and methodologically insufficient lesion analysis approaches. We investigated a large number of patients with focal unilateral right or left frontal ($n = 110$) or posterior ($n = 100$) or subcortical ($n = 65$) lesions imaged with magnetic resonance or computed tomography and compared their performance on the number of overall responses, words produced over time, extremely infrequent/unknown words and inappropriate words generated. We also employed, for the first time parcel-based lesion-symptom mapping, tract-wise statistical analysis as well as Bayesian multi-variate analysis based on meta-analytically defined functional region of interest, including their interactions. We found that left frontal damage was associated with greater impairment than right frontal or posterior damage on overall fluency performance, suggesting that phonemic fluency shows specificity to frontal lesions. We also found that subcorticals, similar to frontals, performed significantly worse than posteriors on overall performance suggesting that subcortical regions are also involved. However, only frontal effects were found for words produced over time, extremely infrequent/unknown and inappropriate words. Parcel-based lesion-symptom mapping analysis found that worse fluency performance was associated with damage to the posterior segment of the left frontal middle and superior gyrus, the left dorsal anterior cingulate gyrus and caudate nucleus. Tract-wise statistical analysis revealed that disconnections of left frontal tracts are critical. Bayesian multi-variate models of lesions and disconnectome maps implicated left middle and inferior frontal and left dorsomedial frontal regions. Our study suggests that a set of well localized left frontal areas together with subcortical regions and several left frontal tracts are critical for word generation. We speculate that a left lateralized network exists. It involves medial, frontal regions supporting the process of 'energization' , which sustains activation for the duration of the task and middle and inferior frontal regions concerned with 'selection', required due to the competition produced by associated stored words, respectively. The methodology adopted represents a promising and empirically robust approach in furthering our understanding of the neurocognitive architecture underpinning executive processes.

1 Department of Neuropsychology, National Hospital for Neurology and Neurosurgery, London WC1N 3BG, UK
2 Department of Brain Repair & Rehabilitation, Institute of Neurology, University College London, London WC1N 3BG, UK
3 The MARCS Institute, University of Western Sydney, Penrith South, NSW NSW 2747, Australia
4 Institute of Cognitive Neuroscience, University College London, London WC1N 3AZ, UK
5 International School for Advanced Studies (SISSA-ISAS), Trieste 34136, Italy

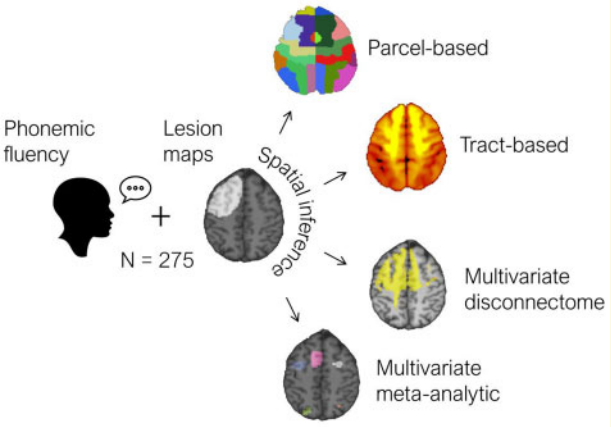
Correspondence to: Lisa Cipolotti PhD, Department of Neuropsychology,
National Hospital for Neurology and Neurosurgery, Queen Square, London WC1 3BG, UK.
E-mail: l.cipolotti@ucl.ac.uk

Keywords: frontal lobes; executive functions; fluency; focal lesion; lesion-symptom mapping

Received June 14, 2021. Revised September 10, 2021. Accepted September 13, 2021. Advance Access publication October 7, 2021

Abbreviations: GNT = Graded Naming Test; HCs = healthy controls; LFs = left frontals; LIFG = left inferior frontal gyrus; LMFG = left middle frontal gyrus; LSFG = left superior frontal gyrus; NART = National Adult Reading Test; PLSM = parcel-based lesion-symptom mapping; RAPM = Raven's Advanced Progressive Matrices; RF = right frontal; ROI = region of interest; SM = superior medial; TSA = tract-wise statistical analysis; VLSM = voxel-based lesion-symptom mapping

Graphical Abstract



Introduction

Fluency tasks have been used to assess the voluntary generation of novel responses. Like most tasks sensitive to frontal lesions, it draws upon different complex systems thought to be critical to active thinking.[1] Phonemic fluency is an example of a fluency task where the content is organized following a reliably unfamiliar rule—a sequence of words beginning with the same phoneme to be generated within a minute. In common with other fluency tasks, it is thought to require the ability to sustain activation for its duration, a process thought to depend on '.energization (cognitive effort) [as] necessary to activate operations not directly triggered in an overlearned fashion by perceptual and motivational inputs'.[2] It also requires specific processes, involved in the conceptualization of language production, thought to be linked to the greater selection demands due to the competition produced by associated stored words that are inappropriately generated by the task rules.[3]

Phonemic fluency is probably the most widely used test for the detection of frontal lobe dysfunction. Many lesion studies have reported reduced phonemic fluency following frontal lobe lesions compared to healthy controls (HCs) (e.g. Refs.[3–7] and Supplementary Table 1a), although some have not.[8,9] The very few studies comparing frontals to posteriors reported inconsistent results. Some studies reported impairment when frontals were compared to posteriors[3,7,10] whilst others did not.[11] Even fewer have contrasted performance of left frontals (LFs) with left posteriors and again contrasting results—significant impairment[12,13] or no impairment[14]—have been reported. Impairments following posterior lesions have also been reported.[7,11,15–17]

Several lesion studies have reported reduced phonemic fluency following LF lesions specifically, some reporting superior medial (SM) and left lateral lesions, including left inferior frontal gyrus (LIFG) when compared to right frontal (RF) (Refs.[3,7] and Refs.[9,15] only for left dorsolateral). However, some studies have also reported no difference between LF or RF patients (Refs.[5,9] only for the left versus right ventromedial comparison and Refs.[11,14,15]; only for the left versus right medial comparison) or no difference between RF and right posterior[18] or RF impairment.[14,19]

Reduced phonemic fluency has also been reported following subcortical lesions compared to HC.[20–23] The occasional studies comparing subcortical and cortical patients reported no significant difference between subcortical and cortical lesions[24] or subcortical and left temporal lesions.[25] To the best of our knowledge only one study reported phonemic fluency performance of frontals, posteriors and subcorticals in the context of a rehabilitation study. No formal comparisons among the three patient groups were reported.[26]

As discussed in an earlier paper, many claimed aspects of phonemic fluency performance and underlying anatomy remain controversial.[27] A major reason for the inconsistency reported, given the great variability in patients' premorbid abilities, must be the small size of many studies published in even excellent journals. Thus, while a few studies with moderately large samples ($n > 32$) exist,[3,7,10,13,14] the majority of the papers, some highly cited, have fewer than 30 frontal patients.[4–6,8,9,11,12,15,18,19] This includes 67% of

those cited where the size of the frontal groups is specified. Moreover, hardly any study has investigated in detail the words and errors produced by patients during the task. Now, nearly 60 years following the first investigations,[28] greater patient numbers are vital for definitive scientific conclusions.

Studies adopting a more refined lesion investigation are also needed. So far, investigations have been largely dominated by traditional lesion-mapping methods. These methods typically use arbitrary cut-off scores, clinical diagnoses or pre-specified regions of interest (ROI). Only a handful of studies have adopted techniques, such as voxel-based lesion-symptom mapping (VLSM), allowing quantitative statistical analysis of the relationship between lesion location and phonemic fluency performance (Refs.[25,27,29–32] and Supplementary Table 1b). These studies reported patients with hemispheric lesions not limited to specific cortical areas and mostly only stroke patients. Hence, lesions to some key frontal areas, such as the left middle (LMFG) and superior frontal gyrus (LSFG), may have been underrepresented.[31] Two of these studies reported only patients with left hemisphere lesions, precluding the assessment of the potential contribution of right hemisphere structures.[25,29] Similar to the results from traditional lesion approaches, VLSM findings have been somewhat diverse and have implicated the involvement of different brain regions. For example, they reported that phonemic fluency was associated with lesions of LF (BA 4,6,44), parietal, post central gyrus, anterio-temporal, insula and putamen regions[29]; or LIFG, insula, left rolandic operculum and LMFG[30] or inferior frontal gyrus, insula, middle frontal gyrus, precentral gyrus, lateral fronto-orbital gyrus as well as other subcortical areas[25] or left subcortical areas and left dorsal temporal regions[31]; or fronto-parietal cortices, anterior prefrontal cortex and insula.[32]

To the best of our knowledge only one study has investigated patients with focal LF and RF patients using parcel-based lesion-symptom mapping (PLSM) and tract-wise statistical analysis (TSA[27]). Lower scores on phonemic fluency were significantly associated with posterior LMFG damage, that according to the JHU atlas also include part of the LIFG, and disconnection in the superior longitudinal fasciculus I and II, frontal aslant tract, frontal orbitopolar tract, left anterior thalamic projections and pons.[27]

Of course, one must be cautious when drawing conclusions concerning the exact localization of critical sites for phonemic fluency using techniques, such as VLSM or PLSM. Indeed, in the last few years, the effectiveness and validity of the mass-univariate approaches to human brain mapping has been increasingly criticized.[33–35] Several studies have promoted the use of multi-variate decoding and computational modelling of data.[36–40] It has been argued that the problem lies in the complex structural architecture of lesions.[34,35] For example, Mah et al.[33] pointed out that methods, such as VLSM make the simplifying assumption that damage to any voxel exhibits a simple pattern of local dependence, which is flawed when applied to brain lesions caused by stroke. In this case, the arterial tree structure of the vascular system results in a high correlation between damage to functionally critical and non-critical regions fed by the same artery. Related criticisms may apply to brain tumours. However, if these criticisms do apply, the associated non-critical regions will not be the same (for further discussion see Ref.41).

Different neuroimaging techniques investigating the contribution of both cortical, subcortical areas and white matters tracts appear to be a fruitful approach in furthering our understanding of the neural substrates of phonemic fluency. Multi-variate analysis offers theoretically higher fidelity for modelling lesion–behaviour relations by enabling the principled removal of spatial biases arising from structured patterns of damage. However, its use is hindered by the imbalance between commonly achievable sample sizes and the number of modelled neuroimaging features,[42–45] compelling the acquisition of more data or the use of fewer features.

Where the rarity of a disorder makes the former infeasible, the latter is our only option. Here, we introduce a new approach to multi-variate lesion-deficit mapping that restricts the range of modelled areas to those functional imaging suggests as plausible candidates, over-inclusively framed. We derive a sparse, behaviourally specific parcellation of the brain from meta-analytic data filtered by textual references to fluency (https://neuroquery.org/), generating a set of functionally defined ROI compact enough to allow flexible multi-variate lesion-deficit modelling based on Bayesian inference. Projection of the white matter components of lesions to the grey matter regions they disconnect enables impact on the candidate regions to be comprehensively surveyed.[46]

We report detailed analyses of both behaviour and underlying anatomy of phonemic fluency in a large sample of 110 LF or RF, 100 posteriors and 65 subcorticals. We examined the behavioural and neural correlates across the three patient groups. As well as overall performance, we examined the qualitative aspects of phonemic output and errors produced. For the neuroimaging analyses, we used PSLM and TSA, allowing comparison with previous studies. For the first time, we also used Bayesian multi-variate analysis of the comparative dependence of fluency on meta-analytically defined functional ROI, including their interactions, modelling both lesions and the grey matter regions they are estimated to disconnect. This allowed us to reduce the potential impact of spatial biases induced by characteristic patterns of damage and to model the interactions between functionally distinct regions.

Materials and methods

Participants

Data from 350 patients with unilateral, focal lesions who attended the Neuropsychology Department of the National

Table 1 Demographics and cognitive test scores

	n	Frontal mean	n	Posterior mean	n	Subcortical mean
Age (years)	110	**46.5**[a,***]	100	51.0	65	55.4
(SD, minimum—maximum)		**(15.80, 19–80)**		(13.40, 20–80)		(14.50, 21–77)
Gender (male/female)		65/47		61/38		41/29
Tumour (left/right)	85	38/47	63	20/43	9	5/4
Stroke (left/right)	25	9/16	37	14/23	56	27/29
Education (years)	105	13.9	86	13.6	57	13.2
(SD)		(3.40)		(3.27)		(2.55)
Lesion volume (mm^3)	87	49.83	81	48.74	54	34.47
Premorbid NART IQ	105	107.9	88	108.4	48	107.8
(SD)		(11.20)		(12.51)		(12.92)
Fluid intelligence SS	95	10.3	64	9.5	48	9.8
(SD)		(3.30)		(3.46)		(3.80)
GNT (Correct/30)	91	20.6	71	21.8	48	21.0
(SD)		(3.70)		(4.05)		(4.20)
RMT Words (Percentile)	34	69.85	68	59.43	40	57.61
(SD)		(27.37)		(34.14)		(34.14)

Scores with significant *P*-values are in bold.
[a]Significant difference between frontal and subcortical patients.
***$P < 0.001$.
n = number; SD = standard deviation; NART = National Adult Reading Test; SS = scaled score; GNT = Graded Naming Test; RMT = Recognition Memory Test.

Hospital for Neurology and Neurosurgery, Queen Square, London were retrospectively screened. Inclusion criteria were: (i) presence of a stroke or tumoural lesion; (ii) ≥70% of the total lesion in the frontal or posterior or subcortical areas; (iii) ability to complete >75% of neuropsychological tasks and consent; (iv) age between 18 and 80 years; (v) no gross language impairments [no agrammatism, greater than fifth percentile on the Graded Naming Test (GNT)],[47] nor memory impairments [as judged clinically and, for the patients for whom quantitative data were available, on the verbal version of either the short or long Recognition Memory Test (RMT)[48,49]]; (vi) absence of psychiatric disorders, history of alcohol or substance abuse or previous neurological disorders; and (vii) native English speaking.

Application of these inclusion criteria resulted in 275 patients with unilateral, single, focal lesions, 110 frontals (LF 47; RF 63), of which only 26 were previously reported[27]; 100 posteriors (Left 34; Right 66) and 65 subcorticals (Left 32, Right 33; see Table 1). Of the patients, 22 were left handed. The median time between stroke/tumour resection and assessment was 19 days (IQR = 5–107) and median time between scan date and assessment was 8 days (IQR = 1–64). Patients were classified based on magnetic resonance imaging (MRI) or computed tomography (CT) scans obtained as part of their clinical investigation (see 'Neuroimaging investigations' section for further details). Information about age at assessment, gender and years of education was collected. Unfortunately, information on the quality of education was not available.

The study was approved by The National Hospital for Neurology and Neurosurgery & Institute of Neurology Joint Research Ethics Committee and conducted in accordance with the 'Declaration of Helsinki'.

Cognitive investigations

All patients were assessed with tests administered and scored in the published standard manner. Due to the retrospective nature of our study, certain data were unavailable for some participants.

Background tests

Premorbid optimal level of functioning was estimated using the National Adult Reading Test (NART).[50] Fluid intelligence (Fluid IQ) using Raven's Advanced Progressive Matrices (RAPM).[51] For patients not administered RAPM ($n = 66$), the Wechsler Adult Intelligence Scale III Performance IQ subtests,[52] known to load heavily on Fluid IQ, were used.[27]

Phonemic fluency S

The S fluency test required patients to generate as many words as possible starting with the letter S within 60 s. Patients were told not to produce proper nouns, change the ending of words (e.g. bake, baking) or repeat words. As the most frequent initial letter in English dictionaries, 'S' should give relatively high scores, compared with other letters. This may allow more opportunities for variables, such as words produced over time to show effects.

We analysed the following.

Overall performance.ó The total number of correct words generated.

Words produced over time.ó The difference in the percentage of correct words generated in the first versus last 15 s (number of words produced in the first 15 s minus

the number of words produced in the last 15 s/Total number of words produced × 100).[14]

'Extremely infrequent' and 'Unknown' words.ó Clinically we noted that frontal patients tended to produce infrequent words (e.g. 'salacious'). We initially investigated our patients' infrequent word production using the Celex database by identifying very low frequency words (Celex frequency count <1) and 'obscure words' not found in this database (e.g. supercalifragilisticexpialidocious). Notably, the Celex is based predominantly on written word-frequency counts and this may differ from spoken word-frequency. For example, 'socialise' has a Celex frequency count <1, despite been a relatively frequent spoken word. To refine our classification, we obtained subjective spoken word-frequency ratings from 10 HCs (see also Ref.53). HC re-classified each word identified as low frequency by the Celex Database according to their subjective spoken language frequency on a 6-points scale (1= 'unknown', 2 = 'extremely infrequent', up to 6 = 'frequent'). Intraclass correlations were significant ($P <$ 0.001) and indicated good levels of reliability (inter-rater reliability =0.810). For each word, we averaged the ratings and calculated the 'subjective spoken word frequency'. We focussed our analysis on words with a subjective frequency estimate <2. We grouped together words that HC judged as unknown or extremely infrequent because, if a subject judged a word as unknown, it is likely that this word is also extremely infrequent and as such share similar characteristics. For example, activation of fewer competing responses (e.g. Ref.54). Patients were classified on the basis of whether they produced at least one word with a subjective frequency estimate <2.

Inappropriate words.ó Although not explicitly stated as a rule, words socially inappropriate in the context of a neuropsychological assessment can be considered a type of error (e.g. *s**t*). Our clinical impression is that inappropriate words, although relatively infrequent, are more likely to be produced by frontal patients. For each patient, we recorded the number of inappropriate words produced as judged by two independent raters. Patients were classified on the basis of whether they produced at least one inappropriate word.

In a recent study, we have specifically investigated the effects of frontal lesions on total number of errors, rule break and perseverations.[27] Hence, on this study, the analyses of these latter types of errors are reported only in the Supplementary material.

Neuroimaging investigations

Patients underwent an MRI or CT scan as a part of their clinical investigation. MRI studies were obtained on Siemens scanners at 1.5 or 3T following a diversity of clinically determined protocols outside our control. CT studies were obtained on Toshiba or Siemens spiral scanners. Note that since the input to the imaging models is *not* raw image data but comparatively large, manually traced, binary lesion masks, the effect of variations in acquisition parameters is likely to be negligible, and—in keeping with other published studies in the field—is not explicitly modelled here. Imaging data were available for 226 patients ($n = 213$ MRI, $n = 13$ CT; $n = 88$ frontal, $n = 84$ posterior, $n = 54$ subcortical). Both MRI and CT scans were used for lesion mapping. Lesions were traced and independently classified using MIPAV (https://mipav.cit.nih.gov/) by J.M. and E.C. and checked by P.N., who was blind to the study results. The lesion masks were segmented and non-linearly normalized to Montreal Neurological Institute (MNI) stereotaxic space at $2 \times 2 \times 2$ mm resolution using SPM-12 software (Wellcome Department of Imaging Neuroscience, London, England: http://www.fil.ion.ucl.ac.uk see Cipolotti et al.[55] for details). The distribution of our patients' lesions is displayed in Fig. 1.

Behavioural analysis

Neuropsychological data were assessed for skewness and kurtosis and tested for normality using the Shapiro–Wilk test.

One-way univariate analysis of variance (ANOVA) or chi-square analyses were conducted for continuous and categorical data respectively to investigate differences between patient groups on demographic variables (age, gender, years of education, lesion volume and NART IQ) and performance on Fluid IQ and naming tasks.

To investigate S fluency overall performance and words produced over time, separate 3×2 ANCOVAs were used, with Group (frontal, posterior and subcortical) and Laterality (left, right) as the between subjects factors and covaried for age. Significant main effects of Group were examined *post-hoc* adjusting for multiple comparisons with Bonferroni correction ($0.05/3 = P = 0.016$) and controlling for the effects of age. We also conducted simple effects analyses to investigate differences between left versus right lesions for the three patient groups separately, while controlling for the effect of age. Fisher's Exact Tests were used to investigate differences in the proportion of patients who produced at least one 'extremely infrequent' or 'unknown' word and the proportion of patients who produced at least one inappropriate word. Due to the relatively small numbers, comparisons were made between frontal and non-frontal (posterior and subcortical) patients.

Lesion-deficit mapping

These analyses were conducted only on overall performance.

PLSM ANALYSES were completed using the NiiStat toolbox for Matlab (http://www.nitrc.org/projects/niistat). To increase statistical power and remain faithful to the anatomical resolution relatively large lesions can

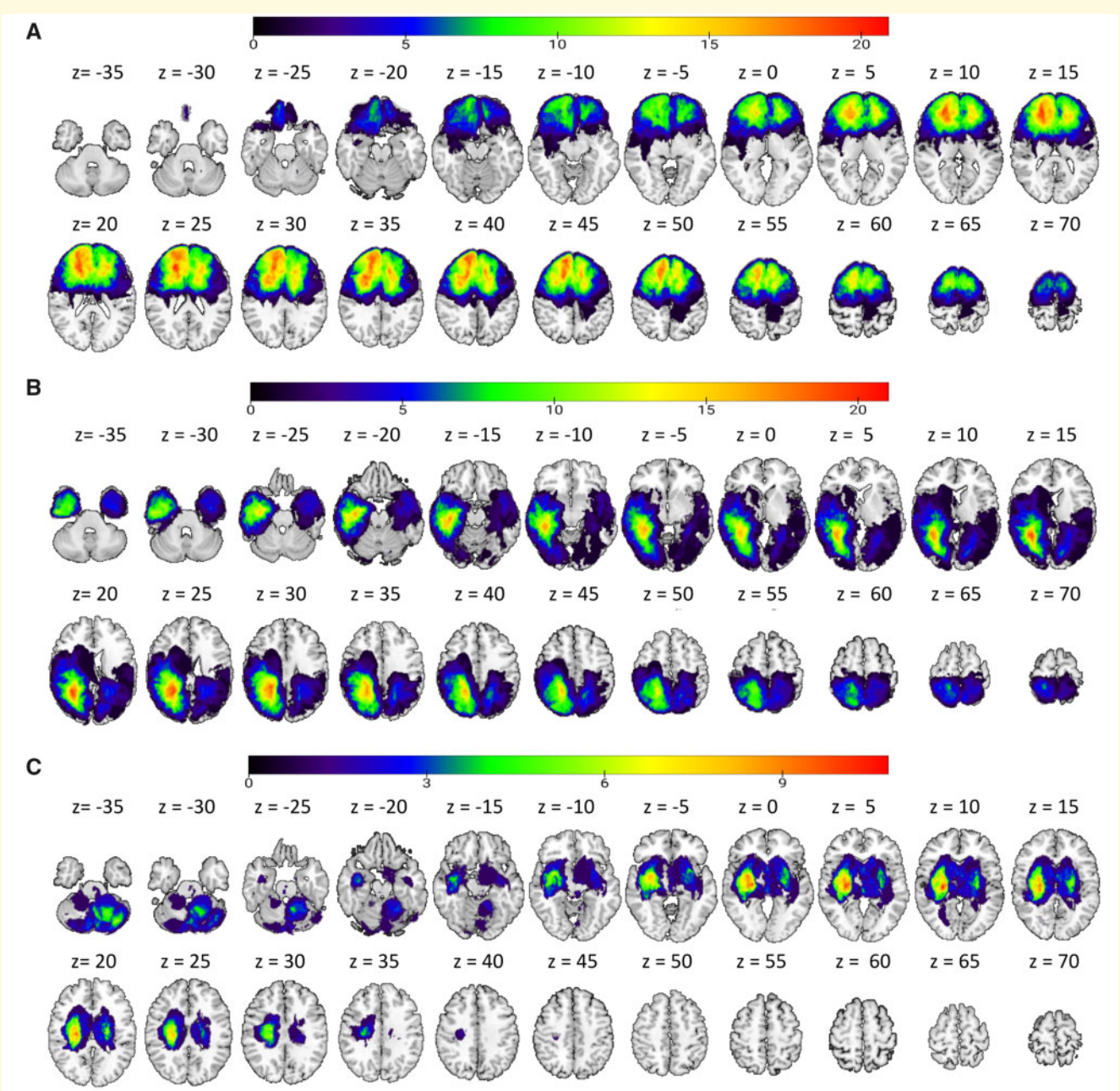


Figure 1 Lesion distribution volume map for A frontals, **B** posteriors and **C** subcorticals. Results are displayed on transversal slices (numbers indicate MNI coordinates) of the ch2better.nii.gz template in MRIcroGL (https://www.nitrc.org). The colour code indicates in how many patients a given voxel was lesioned.

conceivably achieve the brain was parcellated in different regions based on the JHU-MNI atlas,[56] rather than doing voxel-by-voxel analyses. This atlas contains 189 different ROIs, 20 frontals, covering both grey and white matter of the whole brain. As we used TSA analyses (see below), only grey matter regions were considered. ROIs that are infrequently damaged have low statistical power, when increasing the number of comparisons. To ensure enough statistical power to detect a difference, only ROIs where at least 10 patients had damage were included. Three Freedman–Lane permutations[57] were performed with age and lesion volume always entered as nuisance regressors. Permutation thresholding (5000 permutations) was used to correct for multiple comparisons and control the family-wise error rate. An alpha of 0.05 was the cut-off for significance.

TSA ANALYSIS was performed with the Tractotron software (part of the BCBtoolkit, http://www.brainconnectivitybehaviour.eu). This approach allowed us to map patients' lesions (normalized in the MNI 152 referential)

onto tractography reconstructions of white matter pathways obtained from a group of HC by Rojkova et al.[58] For a given lesion, Tractotron provides a probability of disconnection for 55 tracts. When a lesioned voxel overlaps on a white matter tract with a probability >50%, the tract is considered disconnected. Having identified the spared/disconnected white matter tracts for each patient, we then used univariate ANOVA to compare overall performance between patients for each tract separately (spared versus disconnected), controlling for age and lesion volume. To guard against departures from distributional assumptions, results are reported for bootstrapped regressions performed on the basis of 5000 permutations to derive bootstrap confidence intervals. A tract was included in the analysis only if disconnection was observed in 10 or more patients. Significance threshold was set at a *P*-value of 0.05, corrected for multiple comparisons using the false discovery rate.

Bayesian multi-variate modelling of lesions and disconnectome maps with functional meta-analytically derived ROI

Parcellation

To focus attention on regions of the brain implicated in fluency within the functional imaging literature, we created a novel functional sparse parcellation of the brain based on the NeuroQuery meta-analytic framework.[59] The keyword 'fluency'—left unqualified to assure inclusivity—was used to generate a predicted distribution of activations based on the functional imaging literature indexed by NeuroQuery in which the term occurs ($N = 71$). This continuous probabilistic map was thresholded at a Bonferroni-corrected threshold of $P < 0.05$, yielding nine distinct clusters depicted in Supplementary Fig. 1.

Damage quantification

The integrity of each ROI was quantified in two ways: by the degree of overlap between the lesion and the area—an index of focal grey matter damage—and the extent of disruption of the white matter tracts terminating at the area—an index of focal grey matter disconnection. The latter was estimated with the aid of the Disconnectome map module of the BCB Toolkit (http://toolkit.bcblab.com/). Formally, the pure lesion index for each ROI was its summed intersection with the lesion divided by the ROI volume, yielding a measure of proportional focal damage. The disconnectome index

for each ROI was given by its summed intersection with either the lesion or voxels estimated to be disconnected with a probability of 0.5 or higher, again normalized by ROI volume. The signal within each voxel of this intersection ranged from 0.5 to 1, capturing the aggregate effect of both direct damage and disconnection.

Bayesian multi-variate modelling

To enable graceful modelling of structured patterns of damage related to the underlying pathological process that have been shown to invalidate mass-univariate maps,[33,60] we employed Bayesian multi-variate regression implemented in BayesReg,[61] with each ROI treated as a predictor variable, including all pair-wise interactions. Each index of regional damage—pure lesion and disconnectome—was evaluated in separate models with a Laplace noise model and a *g* shrinkage prior, including all ROIs within the chosen parcellation scheme and age as predictors. A Laplace noise model was preferred over Gaussian to allow better modelling of outliers. The posterior distributions of the regression coefficients were estimated with Markov chain Monte Carlo sampling over 100 000 samples with a 100 000 burn-in interval and thinning set at 10, reporting the means and standard deviations of the regression coefficients that survive a 95% Bayesian credibility interval. The effective sample size was >98 for all models. The results were visualized by multiplying each credible ROI by the mean posterior *t*-statistic for the estimated coefficient, overlaying the image on a template cortical surface transformed into MNI space.

Data availability

The data that support the findings of this study are available from the corresponding author, LC, upon reasonable request.

Results

Demographic and baseline variables

Patients were well-matched for gender, years of education and lesion volume across patient groups ($P > 0.05$). There was a significant overall difference between groups for age [$F(2, 272) = 8.01$, $P < 0.001$; see Table 1], with frontal patients being younger than subcortical patients ($P < 0.001$).

There were no significant differences between patient groups for NART, Fluid IQ, GNT or long or short RMT Words (all $P > 0.05$; see Table 1).

Overall S performance

There was no significant difference between tumour and stroke patients for overall s fluency performance or mean time between resection/stroke and neuropsychological assessment [$t(267) = -1.24$, $P = 0.216$; $t(210) = -0.39$, $P = 0.70$, respectively; see Table 1]. A 3×2 ANCOVA controlling for age, found a highly significant difference between patient groups [$F(2, 268) = 7.14$, $P = 0.001$; see Table 2]. *Post-hoc* pair-wise comparisons showed that frontal and subcortical patients performed significantly worse than posterior patients ($P = 0.002$; $P = 0.010$,

Table 2 Overall performance on S fluency

	Frontals $n = 110$	Posteriors $n = 100$	Subcorticals $n = 65$	Frontals		Posteriors		Subcorticals	
				Left $n = 47$	Right $n = 63$	Left $n = 34$	Right $n = 66$	Left $n = 32$	Right $n = 33$
Mean number of correct words generated (SD)	**13.3**[a,**]	15.5	**12.5**[b,*]	**11.8**[c,***]	14.3	14.8	15.9	12.4	12.6
	(5.80)	(4.70)	**(5.40)**	**(5.80)**	(5.40)	(5.00)	(4.80)	(6.00)	(4.10)

Scores with significant *P*-values are in bold.
[a]Dignificant difference between frontals and posteriors.
[b]Significant difference between subcorticals and posteriors.
[c]Significant difference between left frontals and right frontals.
*$P < 0.05$,
**$P < 0.01$,
***$P < 0.001$.

Table 3 Words produced over time on S fluency

	Frontals $n = 51$	Posteriors $n = 53$	Subcorticals $n = 45$	Frontals		Posteriors		Subcorticals	
				Left $n = 24$	Right $n = 27$	Left $n = 18$	Right $n = 35$	Left $n = 22$	Right $n = 23$
Difference in the % of words produced in the first 15 and last 15 s	**28.24**[a,*,b,**]	18.68	16.22	**31%**[c,*]	26%	19%	18%	15%	17%
(SD)	**(17.97)**	(15.20)	(15.85)	**(22)**	(13)	(18)	(14)	(22)	(11)

Scores with significant *P*-values are in bold.
[a]Significant difference between frontals and posteriors.
[b]Significant difference between frontals and subcorticals.
[c]Significant difference between left frontals and left subcorticals.
*$P < 0.05$,
**$P < 0.01$.
n = number; SD = standard deviation.

respectively). We found no significant difference between frontals and subcorticals.

Left hemisphere patients performed significantly worse than right hemisphere patients [$F(1, 268) = 4.18$, $P = 0.042$]. There was no significant interaction between group and laterality [$F(2, 268) = 1.41$, $P = 0.247$).

Simple effects analyses, comparing left versus right patients for each patient group separately revealed a significant difference only in the frontal group (left < right, $P = 0.009$). Importantly, all results for overall s fluency performance remained unchanged when aetiology was included as a covariate (see Supplementary material).

Notably, performance fells ≤2 SD below the performance of the group of 50 HC in 36.2% of LF patients but only in 17.5% of RF patients (see Supplementary material).

Words produced over time

A 3×2 ANCOVA controlling for age, found a highly significant difference between patient groups in the percentage of words generated in the first versus last 15 s [$F(2, 142) = 7.14$, $P = 0.001$; see Table 3]. *Post-hoc* pair-wise comparisons showed that this difference was significantly greater in frontals than in posteriors and subcorticals ($P = 0.014$; $P = 0.002$, respectively). There was no significant difference between posteriors and subcorticals nor was there a significant main effect of laterality [$F(1, 142) = 0.13$, $P = 0.717$] or interaction between group and laterality [$F(2, 142) = 0.49$, $P = 0.615$].

Simple effects analyses, comparing left versus right patients for each patient group separately revealed no significant differences in words produced over time. However, qualitatively, this difference appeared to be greatest in the LF (see Table 3). *Post-hoc* comparison revealed a significant difference between LF and left subcortical patients ($P = 0.019$).

As for overall s fluency performance, all results for the words produced over time remained unchanged when aetiology was included as a covariate (see Supplementary material).

'Extremely infrequent' and 'Unknown' words

HC did not significantly differ from patients in terms of age or years of education [HC mean age =44.60 years, SD = 19.07, $t(283) = -1.17$, $P = 0.376$; HC mean years of education =15.20, SD = 2.25, $t(256) = 1.58$, $P = 0.116$).

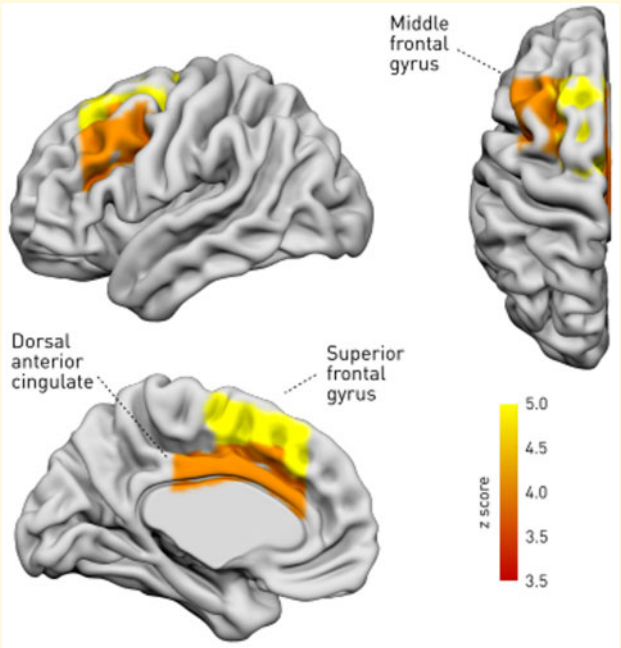


Figure 2 PLSM significant cortical regions associated with overall S performance. Results are displayed on sagittal, coronal and transversal slices of the ch2better.nii.gz template in MRIcron (https://www.nitrc.org/projects/mricron).

Fisher's Exact Test revealed a significant difference in the proportion of frontal (10/110) versus non-frontal (4/165) patients who produced at least one 'extremely infrequent' or 'unknown' word ($P = 0.022$).

Inappropriate words

Fisher's Exact Test revealed a significant difference in the proportion of frontal (28/110) versus non-frontal (20/165) patients who produced at least one inappropriate word ($P = 0.006$).

Lesion analyses

PLSM ANALYSES revealed that worse overall S performance was associated with damage to the posterior segment of the LMFG and LSFG, the left dorsal anterior cingulate gyrus and caudate nucleus (see Fig. 2).

TSA ANALYSIS revealed that patients with disconnection in the left anterior thalamic projections, frontal aslant tract, pons, superior longitudinal fasciculus I and II, cingulum, anterior cingulum, corticospinal tract, frontal commissural, inferior and superior frontal longitudinal tracts performed significantly worse than patients without disconnection in these tracts (FDR corrected, $P < 0.05$). No difference was found between the performance of patients with disconnected or spared RF tracts (FDR corrected, $P > 0.05$).

Bayesian modelling

Pure lesion models

Models of fluency implicated the LMFG and LIFG meta-analytic regions in interaction with age (beta =−1.19953, CI = −2.10637 to −0.38909, *t*-statistic =−2.815), and the left dorsomedial frontal region (LDMF; beta =−0.55634, CI = −1.04656 to −0.08359, *t*-statistic = −2.249; see Fig. 3).

Disconnectome models

Models of fluency produced a similar pattern: LMFG and LIFG in interaction with age (beta =−1.31381, CI = −2.18453 to −0.38359, *t*-statistic =−2.880), and the LDMF (beta =−0.82397, CI = −1.28051 to −0.34017, *t*-statistic =−3.448), but now including an interaction between the LMFG and LIFG and the left inferior temporal region (beta =−2.27361, CI = −4.26994 to −0.30100, *t*-statistic =−2.247, Fig. 3).

Discussion

We investigated the behavioural and neural correlates of phonemic fluency. This task, like most frontal executive tasks, is complex and requires several different processes and systems to come into play. Given its complexity, thorough investigation of phonemic fluency output, errors and fine-grained anatomical investigations require a large sample of patients. We investigated the largest number ($n = 275$) of patients with single, focal, unilateral, right or left, frontal or posterior or subcortical lesions reported so far in the literature. We compared their performance on four variables: overall performance, words produced over time, extremely infrequent/unknown words and inappropriate words. To our knowledge, no previous study has reported such detailed analyses. We also employed, for the first time, in the same patient population, four different neuroimaging techniques: PLSM, TSA as well as Bayesian multi-variate analyses on meta-analytically defined functional ROI, including their interactions. We modelled both the lesions and the grey matter regions they are estimated to disconnect.

We adopted a mixed aetiology approach, similarly to other groups.[32,62–67] We have previously shown that the most serious dangers of using the mixed aetiology approach do not in fact hold. We reported that, in 100 frontals and 91 posteriors with stroke or different types of tumour pathology, performance on phonemic fluency and other cognitive tasks, did not differ with respect to aetiology.[41,68] This supports the view that certain aetiologies do not result in more severe impairments than others and combining across vascular and tumour pathologies is unlikely to produce a major distortion in neuropsychological performance (see Ref.27 for further discussion). One has no reason to assume that the same would not hold for subcortical lesions. Hence, we argue that the

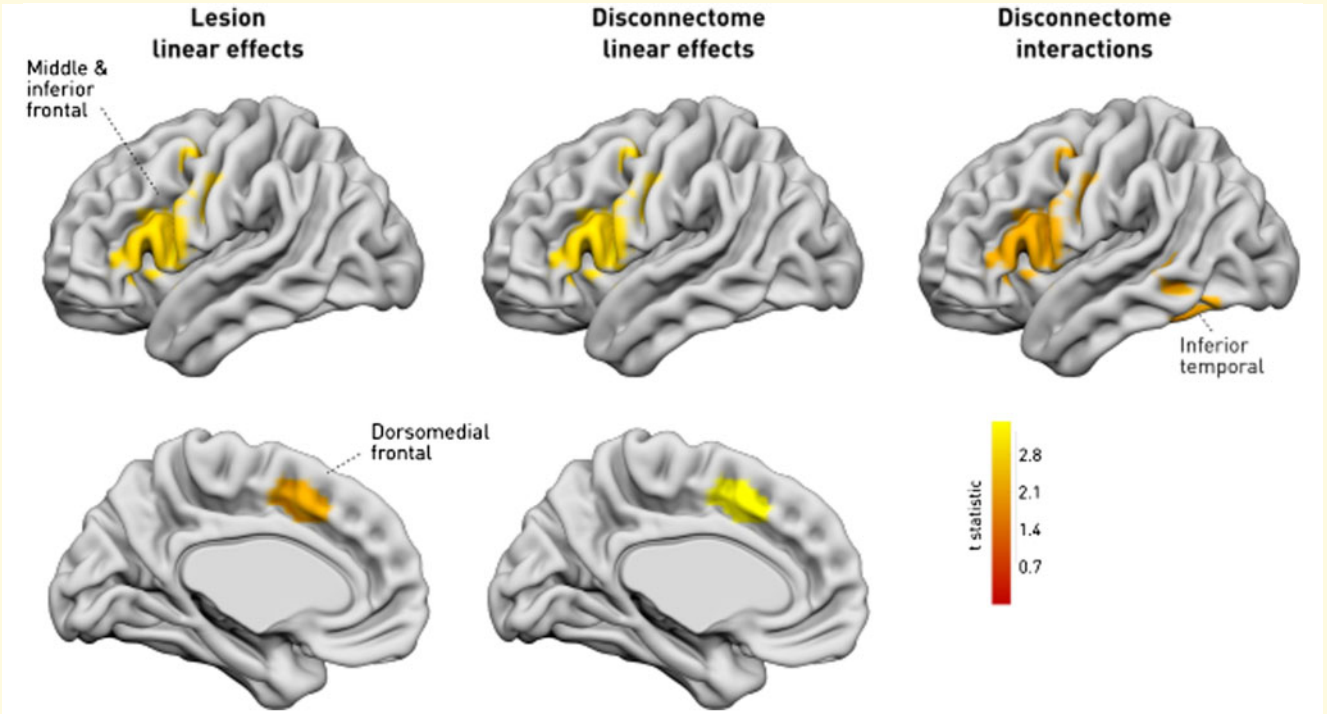


Figure 3 Meta-analytically defined ROI. Depiction of the comparative dependence of phonemic fluency on meta-analytically defined ROI identified as credibly associated within a Bayesian multi-variate model of all regions and their interactions. The colour map shows the *t*-statistic for the estimated coefficient of each region surviving a 95% Bayesian credibility interval. The first column shows linear effects derived with pure lesion models; the second and third columns show linear and interactional effects, respectively, for disconnectome models.

grouping together of patients with stroke or tumoural lesions for the purposes of examining phonemic fluency performance is methodologically justifiable.

We found evidence suggesting that frontals are significantly more impaired than posteriors on overall performance. Although left hemisphere patients performed significantly worse than right hemisphere patients, we found no interaction between group and laterality. The very few studies comparing frontals to posteriors reported inconsistent results (Refs.[4,5,8] impairment–no impairment[9]). Similarly, the handful of studies contrasting performance of LF with left posteriors also reported inconsistent results.[7,13,14] Our finding allows us to conclude that phonemic fluency is a task that shows specificity to frontal lesions, even when overall performance is compared with left posterior patients. Hence, it is a useful clinical tool for the diagnosis of focal LF dysfunction.

We also found evidence suggesting that not only frontals but also subcortical structures play a key role. Thus, subcorticals, similar to frontals, performed significantly worse than posteriors on overall performance. Impaired performance in various executive tasks has been reported in patients with focal or degenerative subcortical lesions[69] (see Ref.70 for a review). To the best of our knowledge, this is the first documentation of the strong effect that focal subcortical lesions play in the executive processes involved in phonemic fluency, traditionally thought to rely on cortical areas (e.g. Refs.[2,13]). There are many possible causes and, indeed, these may differ across subcortical regions. Notably, Chouiter et al.'s[31] VLSM study reported a number of subcortical (putamen, caudate nucleus and pallidum) as well as cortical areas. Subcortical areas may support generalized arousal and initiation abilities involved in speech generation (see Ref.31 for further discussion). The impaired performance of our subcorticals may also arise because of the high connectivity between subcortical areas, in particular the striatum and frontal lobes.[71–73] It might be thought that subcortical damage may just result in dysfunctional brain circuitry disrupting frontal lobe activity. According to this view, damage anywhere in the circuit—subcortical to frontal areas—should produce similar behavioural effects.[74] In line with this, we documented no significant difference in overall performance between subcortical and frontals. However, we found that frontals, not subcorticals, showed a significant difference in the percentage of words generated in the first versus last 15 s. Moreover, frontals produced a significantly higher proportion of extremely infrequent/unknown and inappropriate words than non-frontals. Only frontals showed evidence of lateralized functional organization. These important differences suggest that subcortical and frontal areas may contribute differentially to phonemic fluency. Moreover, different subcortical structures may also contribute differentially. For example,

basal ganglia and thalamus are thought to play different roles in spoken language generation.[75] In addition, certain subcortical areas may have direct links with each other. Thus, the cerebellum modulates the activity of the striatum via a disynaptic pathway in mice.[76] A larger sample of subcorticals is needed so that different subcortical regions can be compared separately to investigate these interesting, and, so far, largely neglected, issues further.

We now turn to discuss our findings in the frontals. We speculate that the significant difference of words produced over time may reflect impairment in the process of 'energization', which sustains responding over time.[2] Stuss et al.[63,77] suggested that, when it is necessary to perform tasks at a fast pace, contention scheduling operating alone would be suboptimal for performance, because selected schema would gradually lose activation over several seconds. They argued that the dominant impairment of patients with SM frontal lesions (notably many of their patients classed as SM had in fact SM and inferior medial lesions, involving anterior cingulate, AC) is one of energization. Consistent with this Stuss et al.[14] reported that patients with SM and left dorso-lateral lesions had proportionately greater reduced production over time. This is broadly in keeping with the performance reported in our frontal patients.

We would like to argue that the significantly higher number of frontals producing at least one extremely infrequent/unknown word may reflect impairment in the process of 'selection'. Phonemic fluency requires selection demands as the task produces competition amongst associated stored words that are inappropriate.[10] In our previous studies, we documented that patients with LIFG impairments were relatively unimpaired in generating sentences from low frequency words.[54] In contrast, they were impaired in generating sentences from high frequency words. Low frequency words have fewer referents than high frequency words and, hence, activate fewer competing responses. Instead, high frequency words activate many conceptual propositions that compete with each other.[54] Our findings that only frontals produced higher proportions of extremely infrequent/unknown words converge with our previous studies with LIFG patients. Of course, future studies will be necessary to determine the reliability of our categorization of extremely infrequent/unknown words.

We suggest that the significantly higher proportion of frontals producing at least one inappropriate word may reflect impairment in frontally located self-monitoring processes. These processes allow monitoring the appropriateness of the phonemic output to the social circumstances. In keeping with this it has been suggested that frontal lobe damage increase swearing. The LF lobe has been thought as a likely candidate for control and inhibitory operations in inappropriate words production.[78] Generation of inappropriate words during phonemic fluency in frontotemporal dementia over Alzheimer disease has been reported.[79] A recent fMRI study reported greater activation of a distributed thalamo-cortical network including medial frontal cortex with its peak in AC, LIFG, left posterior middle temporal gyrus and right thalamus in a picture word interference paradigm where inappropriate words significantly slowed picture-naming latencies compared to neutral words.[80] Our findings that frontals produced inappropriate words more than posteriors and subcorticals is broadly in keeping with these studies and the well-known behavioural disinhibition associated with frontal lobe damage.

Notably, in our previous work, we did not find a significantly higher number of low frequency words (Celex <1) or inappropriate words in a much smaller sample of frontal patients ($n = 30$) when compared with Parkinson and Progressive Supranuclear Palsy patients.[81] This suggests once more that any meaningful investigation of quantitative and qualitative aspects of phonemic output in frontal patients necessitates a larger patient sample.

Lastly, we found evidence of lateralized functional organization for overall S fluency with LFs significantly more impaired than RFs. This finding is in broad agreement with some previous studies[3,25,27,28,82] and does not support the suggestion that reduced phonemic fluency may also be associated to RF lesions.[10,19] Our LF lateralization finding was complemented and extended by the results of our four neuroimaging investigations. PLSM analysis found that worse overall performance was associated with damage to the posterior segment of the LMFG and LSFG, the left dorsal anterior cingulate gyrus and caudate nucleus. TSA analysis revealed that patients with disconnection in the left anterior thalamic projections, frontal aslant tract, pons, superior longitudinal fasciculus I and II, cingulum, anterior cingulus, corticospinal tract, frontal commissural, inferior and superior frontal longitudinal tracts performed significantly worse than patients without disconnection in these tracts. No difference was found between the performances of patients with disconnected or spared RF tracts. These findings are broadly in keeping with those we previously reported using the same two analyses in a sample of frontals.[27] The additional new areas reported in our current study are likely to be due to the inclusion of subcorticals and posteriors, not present in our previous study and the higher number of frontals. In our view, the results, we obtained in our current study with PLSM and TSA represent an important instance of independent replication in an area where traditionally there has been a paucity of replication and an abundance of contradictory findings.[83]

We are acutely aware that the effectiveness and validity of techniques, such as PLSM and VLSM to human brain mapping, have been criticized.[34,35] However, the emphasis of our study is to map a symptom to a large anatomical or functional unit rather than to a fine-grained location. Hence, PLSM and TSA may be viewed as coarser approaches, expected to be more robust to spatial biases, as parcels and tracts are more likely, given their

size, to encompass the displacement of the critical locus. Ideally, if one was able to collect a much larger number of patients with focal lesions, possibly in the order of several hundreds, it may be possible to switch from parcel/tract-wise to multi-variate voxel-wise approach. However, obtaining such a large number of patients to examine symptoms and localization in detail is a well-recognized problem in neuropsychology. For example, Van der Linden[84] noted that to obtain 13 frontals it '…took four years and involved five large hospitals…'.

Given these constraints, we have developed a novel Bayesian approach to multi-variate lesion-deficit mapping that focuses modelling on areas meta-analytic functional imaging suggests as plausible candidates. We deliberately chose the broader term 'fluency' and left it unqualified to assure inclusivity. Of course, this may have resulted in including findings from other fluency modalities, such as semantic fluency, which some authors consider to have a different cortical underpinning from phonemic fluency. Nevertheless, we felt that this was a safer approach to ensure that our analysis included all potentially relevant areas. Our pure lesion models implicated LMFG and LIFG regions in interaction with age, as well as LDMF. Our disconnectome models produced a similar pattern of results, with the addition of an interaction between LMFG and LIFG regions and the left inferior temporal gyrus. These results provide convergent support for a critical role for left middle and inferior frontal cortex in phonemic fluency, but also demonstrate their interaction with other areas that purely linear lesion-deficit mapping has no power to reveal. Previous lesion-deficit dissociations between frontal and temporal contributions to phonemic and semantic fluency may conceivably be explained by failure to model anatomical interactions rather than truly dissociable neural substrates.[29] In any event, increasing appreciation of the value of network models of the brain[85] makes the modelling interactions between lesioned areas imperative. In essence, our four neuroimaging techniques converge in indicating the critical role of several LF grey and white matter areas, in interaction with inferior temporal cortex. The left lateralized network involves medial frontal regions supporting energization and middle and inferior regions supporting selection in tandem with left temporal cortex.

One important limitation of our study is that we used only one measure (S) of phonemic fluency. It would have been helpful to examine fluency characteristics across additional measures of phonemic fluency, such as, for example, letters F and A. This would have allowed us to collect a larger sample of phonemic output production that may have rendered possible investigation, with our neuroimaging techniques, of the anatomical correlates of performance of words produced over time and extremely infrequent/unknown and inappropriate words. In addition, in focussing our study on one specific executive task in the largest cohort reported so far, it was not possible to consider in detail the performance of our patients across other domains that may have influenced their performance on the phonemic fluency task (e.g. processing speed and other executive processes). There was too much variability in other tests carried out. The issue will remain a question for future research.

In conclusion, our study suggests that a set of well localized LF areas together with subcortical regions and disconnection of LF tracts are critical for generation, as measured by fluency performance. The investigation of a large number of patients with single, focal, unilateral lesions together with detailed quantitative and qualitative analyses of phonemic fluency output and different neuroimaging techniques is a fruitful approach in furthering our understanding of the neurocognitive architecture underpinning the executive processes involved in phonemic fluency. Our study also highlights that a major problem in using just the standard clinical measure of the overall number of words generated does not allow distinguishing between frontal and subcortical patients. This makes clinically more relevant the use of additional measures, especially words over time, as well as qualitative measures, such as the production of extremely infrequent/unknown and inappropriate words, and, as previously suggested, rule break errors,[27] when assessing executive functions in brain damaged patients.

We would like to suggest that the methodology adopted in our study represents a promising and empirically robust approach in further investigation of the frontal lobes' contributions to executive functions, whose precise nature is still incompletely understood.[86]

Supplementary material

Supplementary material is available at *Brain Communications* online.

Acknowledgements

The authors are grateful to Lauren Moore for her help with the manuscript preparation.

Funding

This work was supported by the Welcome Trust Grant (089231/A/09/Z). This work was undertaken at University College London Hospitals/University College London, which received a proportion of funding from the Department of Health's National Institute for Health Research Biomedical Research Centre's funding scheme. J.M. was supported by the National Brain Appeal, P.N. is supported by the Wellcome Trust and the University College London Hospitals National Institute for Health Research Biomedical Research Centre. T.X. is supported by the University College London Hospitals National Institute for Health Research Biomedical Research Centre.

Competing interests

The authors report no competing interests.

References

1. Shallice T, Cipolotti L. The prefrontal cortex and neurological impairments of active thought. Annu Rev Psychol. 2018;69:157–180.
2. Shallice T, Stuss DT, Picton TW, et al. Mapping task switching in frontal cortex through neuropsychological group studies. Front Neurosci. 2008;2(1):79–85.
3. Robinson G, Shallice T, Bozzali M, et al. The differing roles of the frontal cortex in fluency tests. Brain. 2012;135(Pt 7):2202–2214.
4. Coslett HB, Bowers D, Verfaellie M, et al. Frontal verbal amnesia: Phonological amnesia. Arch Neurol. 1991;48(9):949–955.
5. Baldo JV, Shimamura AP. Letter and category fluency in patients with frontal lobe lesions. Neuropsychology. 1998;12(2):259–267.
6. Butler RW, Rorsman I, Hill JM, et al. The effects of frontal brain impairment on fluency: Simple and complex paradigms. Neuropsychology. 1993;7(4):519–529.
7. Pendleton MG, Heaton RK, Lehman RA, et al. Diagnostic utility of the Thurstone Word Fluency Test in neuropsychological evaluations. J Clin Exp Neuropsychol. 1982;4(4):307–317.
8. Janowsky JS, Shimamura AP, Kritchevsky M, et al. Cognitive impairment following frontal lobe damage and its relevance to human amnesia. Behav Neurosci. 1989;103(3):548–560.
9. Szatkowska I, Grabowska A, Szymańska O. Phonological and semantic fluencies are mediated by different regions of the prefrontal cortex. Acta Neurobiol Exp. 2000;60(4):503–508.
10. Perret E. The left frontal lobe of man and the suppression of habitual responses in verbal categorical behaviour. Neuropsychologia. 1974;12(3):323–330.
11. Vilkki J, Holst P. Speed and flexibility on word fluency tasks after focal brain lesions. Neuropsychologia. 1994;32(10):1257–1262.
12. Baldo JV, Schwartz S, Wilkins DP, Dronkers NF. Double dissociation of letter and category fluency following left frontal and temporal lobe lesions. Aphasiology. 2010;24(12):1593–1604.
13. Miceli G, Caltagirone C, Gainotti G, Masullo C, Silveri MC. Neuropsychological correlates of localized cerebral lesions in non-aphasic brain-damaged patients. J Clin Exp Neuropsychol. 1981;3(1):53–63.
14. Stuss DT, Alexander MP, Hamer L, et al. The effects of focal anterior and posterior brain lesions on verbal fluency. J Int Neuropsychol Soc. 1998;4(3):265–278.
15. Troyer AK, Moscovitch M, Winocur G, et al. Clustering and switching on verbal fluency: The effects of focal frontal-and temporal-lobe lesions. Neuropsychologia. 1998;36(6):499–504.
16. Martin RC, Loring DW, Meador KJ, et al. The effects of lateralized temporal lobe dysfunction on normal and semantic word fluency. Neuropsychologia. 1990;28(8):823–829.
17. Loring DW, Meador KJ, Lee GP. Effects of temporal lobectomy on generative fluency and other language functions. Arch Clin Neuropsychol. 1994;9(3):229–238.
18. Joanette Y, Goulet P. Criterion-specific reduction of verbal fluency in right brain-damaged right-handers. Neuropsychologia. 1986;24(6):875–879.
19. Davidson PS, Gao FQ, Mason WP, et al. Verbal fluency, Trail Making, and Wisconsin Card Sorting Test performance following right frontal lobe tumor resection. J Clin Exp Neuropsychol. 2008;30(1):18–32.
20. Andrade SP, Brucki SM, Bueno OF, et al. Neuropsychological performance in patients with subcortical stroke. Arq Neuro-Psiquiatr. 2012;70(5):341–347.
21. Peterburs J, Bellebaum C, Koch B, Schwarz M, Daum I. Working memory and verbal fluency deficits following cerebellar lesions: Relation to interindividual differences in patient variables. Cerebellum. 2010;9(3):375–383.
22. Leggio MG, Silveri MC, Petrosini L, et al. Phonological grouping is specifically affected in cerebellar patients: A verbal fluency study. J Neurol Neurosurg Psychiatry. 2000;69(1):102–106.
23. Scherr M, Krenn Y, Sorg C, et al. Patterns of cognitive performance in subcortical ischemic vascular disease (SIVD). J Neuropsychiatr Clin Neurosci. 2014;26(2):150–154.
24. Turunen KE, Kauranen TV, Laari SP, et al. Cognitive deficits after subcortical infarction are comparable with deficits after cortical infarction. Eur J Neurol. 2013;20(2):286–292.
25. Schmidt CS, Nitschke K, Bormann T, et al. Dissociating frontal and temporal correlates of phonological and semantic fluency in a large sample of left hemisphere stroke patients. Neuroimage Clin. 2019;23:101840.
26. Jankowska AM, Klimkiewicz R, Kubsik A, et al. Location of the ischemic focus in rehabilitated stroke patients with impairment of executive functions. Adv Clin Exp Med. 2017;26(5):767–776.
27. Cipolotti L, Molenberghs P, Dominguez J, et al. Fluency and rule breaking behaviour in the frontal cortex. Neuropsychologia. 2020;137:107308.
28. Milner B. Some effects of frontal lobectomy in man. In: JM Warren, K Alkert, eds. The frontal granular cortex and behavior. New York: McGraw-Hill, 1964:313–334.
29. Baldo JV, Schwartz S, Wilkins D, et al. Role of frontal versus temporal cortex in verbal fluency as revealed by voxel-based lesion symptom mapping. J Int Neuropsychol Soc. 2006;12(6):896–900.
30. Biesbroek JM, van Zandvoort MJ, Kappelle LJ, et al. Shared and distinct anatomical correlates of semantic and phonemic fluency revealed by lesion-symptom mapping in patients with ischemic stroke. Brain Struct Funct. 2016;221(4):2123–2134.
31. Chouiter L, Holmberg J, Manuel AL, et al. Partly segregated cortico-subcortical pathways support phonologic and semantic verbal fluency: A lesion study. Neuroscience. 2016;329:275–283.
32. Gläscher J, Adolphs R, Damasio H, et al. Lesion mapping of cognitive control and value-based decision making in the prefrontal cortex. Proc Natl Acad Sci USA. 2012;109(36):14681–14686.
33. Mah YH, Husain M, Rees G, et al. Human brain lesion-deficit inference remapped. Brain. 2014;137(Pt 9):2522–2531.
34. Nachev P. The first step in modern lesion-deficit analysis. Brain. 2015;138(Pt 6):e354.
35. Xu T, Jha A, Nachev P. The dimensionalities of lesion-deficit mapping. Neuropsychologia. 2018;115:134–141.
36. Herbet G, Lafargue G, Duffau H. Rethinking voxel-wise lesion-deficit analysis: A new challenge for computational neuropsychology. Cortex. 2015;64:413–416.
37. Karnath HO, Smith DV. The next step in modern brain lesion analysis: Multivariate pattern analysis. Brain. 2014;137(Pt 9):2405–2407.
38. Smith DV, Clithero JA, Rorden C, et al. Decoding the anatomical network of spatial attention. Proc Natl Acad Sci USA. 2013;110(4):1518–1523.
39. Pustina D, Avants B, Faseyitan OK, et al. Improved accuracy of lesion to symptom mapping with multivariate sparse canonical correlations. Neuropsychologia. 2018;115:154–166.
40. Kristinsson S, Zhang W, Rorden C, et al. Machine learning-based multimodal prediction of language outcomes in chronic aphasia. Hum Brain Mapp. 2021;42(6):1682–1687.
41. Cipolotti L, Healy C, Chan E, et al. The impact of different aetiologies on the cognitive performance of frontal patients. Neuropsychologia. 2015;68:21–30.
42. Nasrabadi NM. Pattern recognition and machine learning. J Electron Imaging. 2007;16:049901.
43. Sinha A, Hripcsak G, Markatou M. Large datasets in biomedicine: A discussion of salient analytic issues. J Am Med Inform Assoc. 2009;16(6):759–767.
44. Lee CH, Yoon HJ. Medical big data: Promise and challenges. Kidney Res Clin Pract. 2017;36(1):3–11.

45. Verleysen M, François D. The curse of dimensionality in data mining and time series prediction. In: J Cabestany, A Prieto, F Sandoval, eds. Computational intelligence and bioinspired systems. Berlin, Heidelberg: Springer; 2005:758–770.
46. de Schotten MT, Foulon C, Nachev P. Brain disconnections link structural connectivity with function and behaviour. Nat Commun. 2020;11:1–8.
47. McKenna P, Warrington EK. Graded naming test. Windsor, Berks: NFER-Nelson Publishing Co. Ltd; 1983.
48. Warrington ETC. Memory tests manual. Hove, United Kingdom: Psychology Press; 1996.
49. Warrington EK. Recognition memory test: Manual. Windsor, United Kingdom: Nfer-Nelson; 1984.
50. Nelson HE. The national adult reading test. Windsor, Berks: NFER-Nelson Publishing Co. Ltd; 1982.
51. Raven JC. Advanced progressive matrices, Sets I and II. London: H.K. Lewis; 1965.
52. Wechsler DA. Wechsler adult intelligence scale. 3rd ed. London: The Psychological Corporation; 1997.
53. Cipolotti L, Warrington EK. Towards a unitary account of access dysphasia: A single case study. Memory. 1995;3(3-4):309–332.
54. Robinson G, Shallice T, Bozzali M, et al. Conceptual proposition selection and the LIFG: Neuropsychological evidence from a focal frontal group. Neuropsychologia. 2010;48(6):1652–1663.
55. Cipolotti L, MacPherson SE, Gharooni S, et al. Cognitive estimation: Performance of patients with focal frontal and posterior lesions. Neuropsychologia. 2018;115:70–77.
56. Faria AV, Joel SE, Zhang Y, et al. Atlas-based analysis of resting-state functional connectivity: Evaluation for reproducibility and multi-modal anatomy–function correlation studies. Neuroimage. 2012;61(3):613–621.
57. Winkler AM, Ridgway GR, Webster MA, Smith SM, Nichols TE. Permutation inference for the general linear model. Neuroimage. 2014;92:381–397.
58. Rojkova K, Volle E, Urbanski M, et al. Atlasing the frontal lobe connections and their variability due to age and education: A spherical deconvolution tractography study. Brain Struct Funct. 2016;221(3):1751–1766.
59. Dockès J, Poldrack RA, Primet R, et al. NeuroQuery, comprehensive meta-analysis of human brain mapping. Elife. 2020;9:e53385.
60. Inoue K, Madhyastha T, Rudrauf D, Mehta S, Grabowski T. What affects detectability of lesion–deficit relationships in lesion studies? Neuroimage Clin. 2014;6:388–397.
61. Makalic E, Schmidt DF. High-dimensional Bayesian regularised regression with the BayesReg package. 2016;arXiv. 1611.06649.
62. Aron AR, Monsell S, Sahakian BJ, et al. A componential analysis of task-switching deficits associated with lesions of left and right frontal cortex. Brain. 2004;127(Pt 7):1561–1573.
63. Stuss DT, Alexander MP, Shallice T, et al. Multiple frontal systems controlling response speed. Neuropsychologia. 2005;43(3):396–417.
64. Roca M, Parr A, Thompson R, et al. Executive function and fluid intelligence after frontal lobe lesions. Brain. 2010;133(Pt 1):234–247.
65. Urbanski M, Bréchemier ML, Garcin B, et al. Reasoning by analogy requires the left frontal pole: Lesion-deficit mapping and clinical implications. Brain. 2016;139(Pt 6):1783–1799.
66. Stamenova V, Gao F, Black SE, et al. The effect of focal cortical frontal and posterior lesions on recollection and familiarity in recognition memory. Cortex. 2017;91:316–326.
67. Aridan N, Pelletier G, Fellows LK, et al. Is ventromedial prefrontal cortex critical for behavior change without external reinforcement? Neuropsychologia. 2019;124:208–215.
68. MacPherson SE, Allerhand M, Gharooni S, et al. Cognitive reserve proxies do not differentially account for cognitive performance in patients with focal frontal and non-frontal lesions. J Int Neuropsychol Soc. 2020;26:739–748.
69. Cox DE, Heilman KM. Dynamic-intentional thalamic aphasia: A failure of lexical-semantic self-activation. Neurocase. 2011;17(4):313–317.
70. Silveri MC. Contribution of the cerebellum and the basal ganglia to language production: Speech, word fluency, and sentence construction—evidence from pathology. Cerebellum. 2021;20(2):282–283.
71. Alexander GE, DeLong MR, Strick PL. Parallel organization of functionally segregated circuits linking basal ganglia and cortex. Annu Rev Neurosci. 1986;9:357–381.
72. Cummings JL. Frontal-subcortical circuits and human behavior. Arch Neurol. 1993;50(8):873–880.
73. Middleton FA, Strick PL. Basal ganglia output and cognition: Evidence from anatomical, behavioral, and clinical studies. Brain Cogn. 2000;42(2):183–200.
74. Mega MS, Alexander MP. Subcortical aphasia: The core profile of capsulostriatal infarction. Neurology. 1994;44(10):1824–1829.
75. Crosson B. Subcortical functions in language: A working model. Brain Lang. 1985;25(2):257–292.
76. Chen CH, Fremont R, Arteaga-Bracho EE, et al. Short latency cerebellar modulation of the basal ganglia. Nat Neurosci. 2014;17(12):1767–1775.
77. Stuss DT, Shallice T, Alexander MP, et al. A multidisciplinary approach to anterior attentional functions. Ann N Y Acad Sci. 1995;769(1):191–212.
78. Jay T. Why we curse: A neuro-psycho-social theory of speech. Philadelphia: John Benjamins Publishing; 1999.
79. Ringman JM, Kwon E, Flores DL, et al. The use of profanity during letter fluency tasks in frontotemporal dementia and Alzheimer's disease. Cogn Behav Neurol. 2010;23(3):159–164.
80. Hansen SJ, McMahon KL, de Zubicaray GI. The neurobiology of taboo language processing: FMRI evidence during spoken word production. Soc Cogn Affect Neurosci. 2019;14(3):271–279.
81. Foley JA, Niven EH, Abrahams S, et al. Phonemic fluency quantity and quality: Comparing patients with PSP, Parkinson's disease and focal frontal and subcortical lesions. Neuropsychologia. 2021;153:107772.
82. Benton AL. Differential behavioral effects in frontal lobe disease. Neuropsychologia. 1968;6(1):53–60.
83. Shallice T. Cognitive neuropsychology and its vicissitudes: The fate of Caramazza's axioms. Cogn Neuropsychol. 2015;32(7-8):385–411.
84. Van der Linden PM. Supervisory attentional system in patients with focal frontal lesions. J Clin Exp Neuropsychol. 2001;23(2):225–239.
85. Bullmore E, Sporns O. Complex brain networks: Graph theoretical analysis of structural and functional systems. Nat Rev Neurosci. 2009;10(3):186–198.
86. Hornberger M, Bertoux M. Right lateral prefrontal cortex—specificity for inhibition or strategy use? Brain. 2015;138(Pt 4):833–835.