

Dissociating frontal and temporal correlates of phonological and semantic fluency in a large sample of left hemisphere stroke patients

Charlotte S.M. Schmidt^{a,b,c,d,e,*}, Kai Nitschke^{a,b,c,d,e}, Tobias Bormann^{a,e}, Pia Römer^{b,e}, Dorothee Kümmerer^{a,b,d,e}, Markus Martin^{a,b,d,e}, Roza M. Umarova^{a,b,d,e,i}, Rainer Leonhart^f, Karl Egger^{b,e,g}, Andrea Dressing^{a,b,d,e}, Mariachristina Musso^{a,b,d,e}, Klaus Willmes^h, Cornelius Weiller^{a,b,d,e}, Christoph P. Kaller^{a,b,d,e,g,*}

^a Dept. of Neurology, University Medical Center Freiburg, Germany

^b Freiburg Brain Imaging Center, University of Freiburg, Germany

^c Biological and Personality Psychology, Dept. of Psychology, University of Freiburg, Germany

^d BrainLinks-BrainTools Cluster of Excellence, University of Freiburg, Germany

^e Faculty of Medicine, University of Freiburg, Germany

^f Social Psychology and Methodology, Dept. of Psychology, University of Freiburg, Germany

^g Dept. of Neuroradiology, University Medical Center Freiburg, Germany

^h Dept. of Neurology, University Hospital, RWTH Aachen University, Germany

ⁱ Department of Neurology, Inselspital, Bern University Hospital, University of Bern, Switzerland

ARTICLE INFO

Keywords:

Verbal fluency
Brain lesion
Frontal lobe
Temporal lobe
Double dissociation

ABSTRACT

Previous lesion studies suggest that semantic and phonological fluency are differentially subserved by distinct brain regions in the left temporal and the left frontal cortex, respectively. However, as of yet, this often implied double dissociation has not been explicitly investigated due to mainly two reasons: (i) the lack of sufficiently large samples of brain-lesioned patients that underwent assessment of the two fluency variants and (ii) the lack of tools to assess interactions in factorial analyses of non-normally distributed behavioral data. In addition, previous studies did not control for task resource artifacts potentially introduced by the generally higher task difficulty of phonological compared to semantic fluency.

We addressed these issues by task-difficulty adjusted assessment of semantic and phonological fluency in 85 chronic patients with ischemic stroke of the left middle cerebral artery. For classical region-based lesion-behavior mapping patients were grouped with respect to their primary lesion location. Building on the extension of the non-parametric Brunner-Munzel rank-order test to multi-factorial designs, ANOVA-type analyses revealed a significant two-way interaction for *cue type* (semantic vs. phonological) by *lesion location* (left temporal vs. left frontal vs. other as stroke control group). Subsequent contrast analyses further confirmed the proposed double dissociation by demonstrating that (i) compared to stroke controls, left temporal lesions led to significant impairments in semantic but not in phonological fluency, whereas left frontal lesions led to significant impairments in phonological but not in semantic fluency, and that (ii) patients with frontal lesions showed significantly poorer performance in phonological than in semantic fluency, whereas patients with temporal lesions showed significantly poorer performance in semantic than in phonological fluency.

The anatomical specificity of these findings was further assessed in voxel-based lesion-behavior mapping analyses using the multi-factorial extension of the Brunner-Munzel test. Voxel-wise ANOVA-type analyses identified circumscribed parts of left inferior frontal gyrus and left superior and middle temporal gyrus that significantly double-dissociated with respect to their differential contribution to phonological and semantic fluency, respectively. Furthermore, a main effect of lesion with significant impairments in both fluency types was found in left inferior frontal regions adjacent to but not overlapping with those showing the differential effect for phonological fluency.

The present study hence not only provides first explicit evidence for the anatomical double dissociation in verbal fluency at the group level but also clearly underlines that its formulation constitutes an oversimplification as parts of left frontal cortex appear to contribute to both semantic and phonological fluency.

* Corresponding authors at: Dept. of Neurology, University Medical Center, University of Freiburg, Breisacher Str. 64, 79106 Freiburg, Germany.

E-mail addresses: charlotte.schmidt@uniklinik-freiburg.de (C.S.M. Schmidt), christoph.kaller@uniklinik-freiburg.de (C.P. Kaller).

<https://doi.org/10.1016/j.nicl.2019.101840>

Received 18 June 2018; Received in revised form 15 January 2019; Accepted 24 April 2019

Available online 25 April 2019

2213-1582/ © 2019 Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Verbal fluency is one of the most frequently used neuropsychological measures of language abilities and executive functioning (Moscovitch, 1994; Strauss et al., 2006; Lezak et al., 2012; Shao et al., 2014), requiring the examinee to generate as many words as possible to a given category cue (semantic fluency) or letter cue (phonological fluency) within a pre-set time interval (e.g. 60 s; Strauss et al., 2006; Lezak et al., 2012).

Previous region-based lesion studies revealed that patients with left frontal lesions produce fewer words in phonological fluency tasks as compared to healthy controls (e.g., Perret, 1974; Pendleton et al., 1982; Miller, 1984; Janowsky et al., 1989; Stuss et al., 1994; Gershberg and Shimamura, 1995; Tucha et al., 1999; Channon and Crawford, 2000; Jurado et al., 2000; Baldo et al., 2001) and to patients with non-frontal lesions (e.g., Milner, 1964; Perret, 1974; Helmstaedter et al., 1998), whereas patients with lesions in left temporal areas produce fewer words in semantic fluency tasks as compared to healthy controls (Martin et al., 1990; Troyer et al., 1998; Luckhurst and Lloyd-Jones, 2001). Based on this evidence, it hence appears that semantic and phonological fluency are differentially subserved by left temporal and left frontal lobe, respectively (Baldo et al., 2006, 2010). Yet, to the best of our knowledge, none of the previous studies has directly investigated this implied double dissociation by statistically assessing the respective interaction between the type of fluency task (semantic vs. phonological) and lesion location (left temporal vs. left frontal).

Furthermore, establishing this double dissociation at the group level is considerably hampered by a potential task resource artifact (Shallice, 1988; Davies, 2010) due to general differences in task difficulty between semantic and phonological fluency for which – to the best of our knowledge – none of the previous studies has explicitly controlled for. Semantic fluency is commonly reported to be easier than phonological fluency, as becomes evident from the descriptive overview of performance scores reported in previous lesion studies (see Table 1), although this was not tested statistically in most studies. As this pattern can be observed across various samples of brain lesioned patients as well as healthy controls (Table 1; see also Katzev et al., 2013; Schmidt et al., 2017), frontal lesion sites associated with poorer performance in phonological fluency may hence simply reflect higher task demands on general processes, which nonetheless subserve both types of verbal fluency rather than a functional specificity for phonological fluency (Shallice, 1988; Davies, 2010).

Finally, left frontal and temporal association cortices constitute large entities subsuming different, functionally heterogeneous subregions. Impairments in semantic and phonological fluency are therefore likely to anatomically dissociate with circumscribed lesions at the level of functionally specific subregions rather than at the coarse level of the overall left temporal and left frontal lobe. In turn, left frontal and/or temporal lobes may also entail subregions, which subserve common processes (Schmidt et al., 2017) and whose integrity is associated with successful performance in both types of verbal fluency (cf. Biesbroek et al., 2016; Chouiter et al., 2016). In this respect, evidence from neuroimaging studies suggests contributions of frontal areas in semantic fluency (e.g., Frith et al., 1991; Hirshorn and Thompson-Schill, 2006; Katzev et al., 2013; see Costafreda et al., 2006, and Wagner et al., 2014, for meta-analyses). Likewise, several region-based lesion studies reported semantic fluency to be affected not only in patients with left temporal but also with left frontal lesions (e.g., Stuss et al., 1996, 1998, 1999; Baldo and Shimamura, 1998; Baldo and Shimamura, 1998; Rogers et al., 1998; Troyer et al., 1998; Szatkowska et al., 2000; Sylvester and Shimamura, 2002; Reverberi et al., 2006; Robinson et al., 2012; see Henry and Crawford, 2004, for a meta-analytic review). Applying voxel-based lesion-behavior mapping (VLM; Bates et al., 2003; Rorden et al., 2007) as a complement to the common region-based approaches may hence reveal more specific insights into the differential functional anatomy underlying semantic and

phonological fluency, but it requires larger samples of patients and was therefore only rarely applied (see Baldo et al., 2006; Almairac et al., 2015; Biesbroek et al., 2016; Chouiter et al., 2016, for notable exceptions).

Taken together, as of now, empirical support for the proposed double dissociation of semantic vs. phonological fluency and left temporal vs. left frontal cortex is promising but limited, as explicit statistical confirmation is still lacking and would also have to account for the above raised issues of task difficulty and anatomical specificity. In addition, systematic assessment of the effects of left frontal and left temporal brain lesions on phonological and semantic fluency in a non-parametric factorial analysis does not only require a large sample of well-described patients but also calls for appropriate statistical tools to investigate interaction effects in non-normally distributed behavioral data of brain-lesioned patients (cf. Nitschke et al., forthcoming).

In the present study, we therefore addressed all these challenges by explicitly studying the proposed double dissociation between fluency type (semantic vs. phonological) and lesion location (temporal vs. frontal) at the group level in a large sample of 85 chronic left hemisphere stroke patients (see Section 2.1). To control for a potential task resource artifact, we administered a German version of the verbal fluency task that comprised a subset of items with comparable task difficulty for both the semantic and phonological fluency condition (cf. Katzev et al., 2013; see Section 2.2). In a first series of classical region-based analyses patients were grouped with respect to their primary lesion location (i.e., left frontal, left temporal, and neither of them as stroke control group). Non-parametric factorial analyses were conducted based on an extension of the Brunner-Munzel rank-order test for multi-factorial designs (Brunner and Munzel, 2000, 2002; see Section 2.5). To further understand the hypothesized interaction effects between lesion location and fluency type, subsequent contrast analyses followed an established predefined operationalization of a double dissociation (see Section 2.6). In a second series of voxel-based analyses we applied a newly developed toolbox (NIX-toolbox; Nitschke et al., forthcoming; cf. Dressing et al., 2018) which also builds on the multi-factorial extension of the Brunner-Munzel rank-order test. This allowed us to directly assess interaction effects in anatomically more specific VLM analyses. Besides establishing the implied double dissociation, the present study further aimed at revealing common brain regions, which are comparably crucial for both semantic and phonological fluency (cf. Biesbroek et al., 2016; Chouiter et al., 2016).

2. Materials and methods

2.1. Participants

Chronic stroke patients were recruited from the Department of Neurology at the University Medical Center Freiburg and tested at least 5 months post-stroke as part of a larger study on their recovery after ischemic stroke (e.g., Dressing et al., 2018; Beume et al., 2017; Martin et al., 2016). The patient-specific inclusion criterion was first presentation of an ischemic stroke of the middle cerebral artery of the left hemisphere without a hemorrhagic event. Exclusion criteria in the acute phase were age over 90 years, illiteracy, as well as previous infarcts, intracerebral hemorrhage, traumatic brain injury or contemporary re-infarct. Any major cognitive impairment (e.g. dementia), hearing and visual deficits, or alcohol abuse constituted further exclusion criteria. Patients with an inability to tolerate the MRI examination or neuropsychological testing were also excluded from the study. Every eligible patient was asked to participate and, once consented, tested at the Department of Neurology. The study was approved by the local ethics committee and conducted in compliance with the Helsinki Declaration of the World Medical Association (<http://www.wma.net>).

From the resulting sample of 101 attending chronic stroke patients, 4 patients were excluded because of too severe aphasia (i.e. patients were unable to speak). Another 5 patients were excluded because they

Table 1
Overview of performance scores for semantic and phonological fluency reported in previous lesion studies.

Study	Authors/Year	Sample description	Semantic fluency		Phonological fluency		Effect of task difficulty
			Items	Performance M \pm SD	Items	Performance M \pm SD	
1	Joanette et al. (1986)	RH vascular lesion (n = 35) HC (n = 20)	Animals, furniture Animals, furniture Animals	Σ 32.5 \pm 10.5 Σ 43.2 \pm 7.9	B, R B, R F, A, S	Σ 20.0 \pm 11.3 Σ 24.5 \pm 13.1 Σ 23.2 \pm 7.8	sem > phon sem > phon sem > phon
2	Loring et al. (1994)	LH temporal lobectomy (n = 12, pre-operative) RH temporal lobectomy (n = 11, pre-operative)	Animals	16.9 \pm 5.1	F, A, S	Σ 28.8 \pm 8.7	sem > phon
3	Vilkkii et al. (1994)	LH anterior lesion (n = 19) LH posterior lesion (n = 16) RH anterior lesion (n = 10) RH posterior lesion (n = 15) RH brain damage (n = 15)	Animals (20) ¹ Animals (20) ¹ Animals (20) ¹ Animals (20) ¹ Animals, clothes, sports, vegetables, tools, weapons	176 \pm 196 s 175 \pm 242 s 111 \pm 94 s 76 \pm 26 s \emptyset 12.71 \pm 3.52	S (10) ¹ S (10) ¹ S (10) ¹ S (10) ¹ P, M, T, V, L, N	122 \pm 194 s 68 \pm 76 s 57 \pm 58 s 30 \pm 11 s \emptyset 10.62 \pm 3.83	sem > phon ¹ phon > sem ¹ sem > phon ¹ phon > sem ¹ sem > phon
4	Goulet et al. (1997)	HC (n = 15)	Animals, clothes, sports, vegetables, tools, weapons	\emptyset 15.53 \pm 4.13	P, M, T, V, L, N	\emptyset 14.69 \pm 7.57	sem > phon
5	Baldo et al. (1998)	LH + RH frontal lesion (n = 12) HC (n = 12)	Animals, fruits, occupations Animals, fruits, occupations	\emptyset 11.7 \emptyset 18.5	F, A, S F, A, S	\emptyset 7.6 \emptyset 16.7	sem > phon sem > phon
6	Jurado et al. (2000)	TBI all (n = 13) RH TBI lesion (n = 3) LH TBI lesion (n = 3) Bilateral TBI lesion (n = 7) HC (n = 26)	Animals, supermarket items Animals, supermarket items Animals, supermarket items Animals, supermarket items Animals, supermarket items	Σ 44.0 \pm 13.8 Σ 31.7 \pm 1.5 Σ 57.0 \pm 13.2 Σ 43.7 \pm 12.6 Σ 51.4 \pm 10.6	F, A, S F, A, S F, A, S F, A, S F, A, S	Σ 37.2 \pm 12.6 Σ 36.3 \pm 4.2 Σ 40.0 \pm 25.2 Σ 36.3 \pm 9.7 Σ 47.7 \pm 12.2	sem > phon sem > phon sem > phon sem > phon sem > phon
7	Szatkowska et al. (2000)	LH DLPPC (n = 6) RH DLPPC (n = 6) LH ventromedial PFC (n = 6) RH ventromedial PFC (n = 6)	Animals Animals Animals Animals	N.A. N.A. N.A. N.A.	K K K K	N.A. N.A. N.A. N.A.	sem \neq phon *** sem > phon *** across all patients
8	Baldo et al. (2006)	LH temporal and frontal lesion (n = 48)	Fruits, animals, supermarket items (90 s)	\emptyset 30.5 \pm 22.1	F, A, S (90 s)	\emptyset 18.8 \pm 17.6	sem > phon
9	Baldo et al. (2010)	LH temporal lesion (n = 1) LH frontal lesion (n = 1) LH + RH posterior lesions (n = 20) LH frontal lesion (n = 20) RH frontal lesion (n = 27) HC (n = 35)	Fruits, animals, supermarket items Fruits, animals, supermarket items Fruits, vegetables Fruits, vegetables Fruits, vegetables Fruits, vegetables	Σ 6 Σ 35 \emptyset 14.2 \pm 5.2 \emptyset 12.4 \pm 6.1 \emptyset 14.3 \pm 6.0 \emptyset 20.5 \pm 5.5	F, A, S F, A, S S S S S	Σ 34.5 Σ 2 13.8 \pm 5.8 6.8 \pm 5.2 12.6 \pm 4.2 16.9 \pm 4.7	sem > phon sem > phon sem > phon sem > phon sem > phon sem > phon
10	Robinson et al. (2012)	LH + RH posterior lesions (n = 20) LH frontal lesion (n = 20) RH frontal lesion (n = 27) HC (n = 35)	Animals (2 min)	23.1 \pm 10.4	A (1 min) N (1 min)	8.0 \pm 4.3 7.7 \pm 4.4	sem > phon sem > phon
11	Biesbroek et al. (2016)	Ischemic stroke patients (n = 93)	Animals	N.A.	M or S	N.A.	sem > phon ***
12	Chouiter et al. (2016)	LH lesion (n = 108) RH lesion (n = 83)	Animals	N.A.	M or S	N.A.	sem > phon
13	Li et al. (2017)	Stroke patients (n = 51) HC (n = 39)	Animals, fruits and vegetables, tools Animals, fruits and vegetables, tools	\emptyset 8.39 ³ \emptyset 17.26 ³	/Da4/, /Bu4/ /Da4/, /Bu4/	5.67 ³ 11.88 ³	sem > phon sem > phon

Note. M \pm SD, mean \pm standard deviation; HC, healthy control; LH, left hemisphere; RH, right hemisphere; TBI, traumatic brain injury; N.A., not available; *** p < .001. Note that the number of applied items for phonological and semantic fluency differ in some of the studies. If reported, the sum (Σ) and the average (\emptyset) symbols indicate the type of aggregation applied to the performance scores to allow for relative comparisons.¹ In this study performance was assessed as the time subjects needed to generate 20 and 10 semantically (animals) and phonologically (S) cued words, respectively.² In this study Chinese syllables were used as phonological cues.³ Average scores calculated on the basis of the mean number of words produced for each cue.

either did not complete or were unable to perform the task (i.e. task abortion at the request of the patient). Data of one patient was excluded because of technical difficulties with segmentation and normalization of the lesion data and 4 other patients were excluded because of very small lesions (i.e. lesion volume of < 1 ml). Given the influence of education on performance in verbal fluency tasks (Tombaugh et al., 1999; Strauss et al., 2006) another 2 patients were excluded due to an extraordinary low educational attainment of < 8 years. Prior to the main analysis, individual data were inspected for outliers within the patient sample. In detail, the total number of words produced with the semantic and the phonological fluency task and their respective inter-quartile ranges were separately computed but revealed no patients with deviant performance.

The final sample comprised 85 chronic stroke patients (23 female) with mean age (\pm SD) of 63.97 ± 14.20 years (range 22.4–85.8 years), mean education (\pm SD) of 13.27 ± 3.46 years (range 8–23 years), and average post-stroke duration (\pm SD) of 16.98 ± 18.20 months (range

5.0–66.8 months). Lesion overlay of the final overall sample is displayed in Fig. 1A.

2.2. Verbal fluency task

Participants were administered a German version of the verbal fluency task recently established by Katzev et al. (2013). The task comprised 8 semantic cues (categories, e.g. vegetables) and 8 phonological cues (letters, e.g. V) that were further classified as being easy or hard (cf. Katzev et al., 2013). Four different items were presented for each combination of cue type (semantic vs. phonological) and level of difficulty (easy vs. hard), yielding a total of 16 tasks (i.e., 8 semantic categories and 8 letters). Given that semantic fluency is commonly easier than phonological fluency (Katzev et al., 2013), applying this manipulation of task difficulty was done to allow for matching the two different fluency types for difficulty, which constitutes a central prerequisite for avoiding task resource artifacts (cf. Shallice, 1988) and for

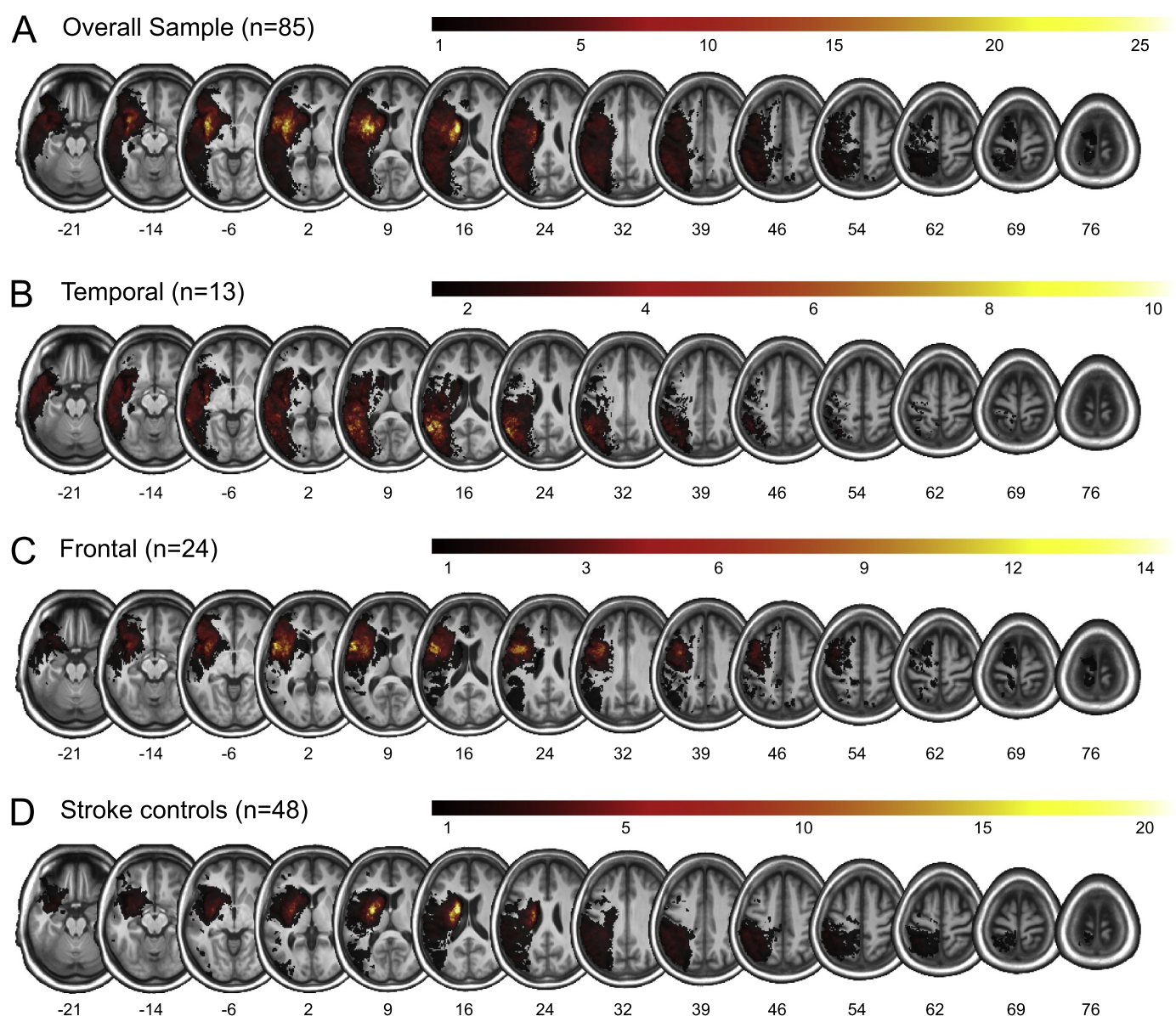


Fig. 1. Lesion overlays (A) for the overall sample of stroke patients ($n = 85$, maximum overlap = 25) as well as for the subsample of patients with lesions mainly (B) in left temporal cortex ($n = 13$, maximum overlap = 10) and (C) left frontal cortex ($n = 24$, maximum overlap = 14), and (D) the stroke control patients with their main lesion sites neither in temporal nor frontal cortex ($n = 48$, maximum overlap = 20).

identifying a functionally specific contribution of left temporal and left frontal brain regions independent of differences in task difficulty. In this respect, only the four semantic hard (thereafter semantic) items (i.e., Flüssigkeiten [English: liquids/fluids]; Spielzeuge [toys], Möbelstücke [furniture]; Gemüsearten [vegetables]) and the four phonological easy (thereafter phonological) items (i.e., T; B; S; K) as specified in the study of Katzev et al. (2013) were used for the present analyses as these were expected to be comparable in task difficulty (see also Schmidt et al., 2017).

Items and presentation order were identical for all patients (easy semantic, hard semantic, easy phonological, hard phonological items). Instructions for the verbal fluency task were given orally by the experimenter (CS or PR). Participants were told that the verbal fluency task would comprise two different parts (semantic and phonological) and that they were to generate as many nouns as possible within a time limit of 60 s following either a category or a letter. Task rules were explained with exemplar items (i.e., category: Lebensmittel [food or groceries]; letter: E). For both conditions only words common in German should be uttered. No words should be produced twice, no proper names, and no words beginning or ending with the same word stem were allowed. In the phonological condition, legitimate responses were limited to nouns to make the task more comparable to the semantic task. The total number of correct words for the semantic and for the phonological condition was recorded and served as outcome measures for data analyses.

2.3. Magnetic resonance imaging (MRI)

All participants were administered T1-weighted high-resolution anatomical brain imaging in the acute stroke phase usually 2–3 days after the stroke event on a 3-Tesla TIM TRIO whole-body MRI scanner (SIEMENS, Erlangen, Germany) applying T1-weighted magnetization-prepared rapid gradient echo (MPRAGE) imaging with the following scan acquisition parameters: repetition time (TR), 2200 ms; echo time (TE), 2.15 ms; inversion time (TI), 1100 ms; voxel size, $1 \times 1 \times 1 \text{ mm}^3$; 176 sagittal slices. Furthermore, fluid attenuated inversion recovery (FLAIR) images (TR, 9000 ms; TE, 93.0 ms, flip angle, 140° ; voxel size, $0.94 \times 0.94 \times 5.00 \text{ mm}^3$, 23 axial slices) were taken from all patients. In addition, diffusion-weighted images (DWI) were acquired in the acute phase for the delineation of lesion location and size with the following standard sequence: TR, 3700 ms; TE, 100 ms, flip angle, 90° ; voxel size, $1.2 \times 1.2 \times 5 \text{ mm}^3$; 23 axial slices; 3 diffusion-encoding gradient directions with a b-factor of 1000 s/mm^2 . The same MRI sequences were also acquired in the chronic phase to control for potential re-infarction.

2.4. Lesion demarcation and analysis

First, lesions were semi-automatically delineated in the diffusion-weighted images (DWI) acquired 2–3 days after the stroke event, which is considered to reflect the irreversible core of the ischemic infarct. Lesion delineation was based on a customized region-of-interest toolbox implemented in SPM8 (release r4667; <http://www.fil.ion.ucl.ac.uk/spm/software/spm8>) running on MATLAB (R2012a; The MathWorks, Inc., Natick, MA). Individual intensity thresholds were applied to find the best match between the binary lesion map and the diffusion-restricted area. The resulting lesion map was subsequently inspected with MRIcron (<http://people.cas.sc.edu/rorden/mricron/index.html>) and manually adjusted if necessary (cf. Hoeren et al., 2014; Martin et al., 2016). For spatial normalization, the lesion map and the underlying DWI scan were co-registered to the anatomical T1 scan in order to apply the normalization parameters estimated from the T1 image onto the lesion map derived from the DWI scan. For four patients no T1 scan with sufficient quality for normalization purposes was available. In these cases, co-registration and normalization was based on the acute FLAIR. High-resolution T1 (or FLAIR) images were segmented using the

VBM8 toolbox (release r435; <http://dbm.neuro.uni-jena.de/vbm/download/>) for SPM8. Deformation field parameters for nonlinear normalization into the stereotactic Montreal Neurological Institute (MNI) standard space were then computed using the DARTEL (Diffeomorphic Anatomical Registration Through Exponentiated Lie algebra; Ashburner, 2007) approach implemented in VBM8. Normalization quality of lesion maps was visually checked (cf. Hoeren et al., 2014; Martin et al., 2016).

2.5. Data analysis

In a first series of region-based lesion-behavior mapping analyses, the hypothesized double dissociation was investigated by classifying patients according to the primary location of their lesion. Classification was done using in-house custom software written in MATLAB (The MathWorks, Inc., Natick, MA) based on the predefined regions-of-interest (ROIs) for the frontal and temporal lobes (Lancaster et al., 2000) as implemented in the WFU PickAtlas (Maldjian et al., 2003) available for SPM8. Intersections between the atlas' ROIs and the individual lesions were calculated. The ROI containing the proportionately biggest lesion volume was regarded as primary lesion location (i.e., the ROI with the highest percentage share of the overall lesion). Patients with the primary lesion location neither in the frontal nor in the temporal lobe were regarded as stroke control group. Visual examination of the lesion overlays (see Fig. 1) ensured that the resulting groups distinctly reflected the respective ROIs and were not biased (e.g. due to patients with large fronto-temporal lesions). Analyses comprised an extension of the non-parametric Brunner-Munzel rank test (Brunner and Munzel, 2002) for multi-factorial designs with within- and between-subjects factors. The ANOVA-type test statistic was chosen, which is based on a $F(df, \infty)$ -distribution and robust even for small sample sizes (Brunner and Munzel, 2002, p. 72).

The region-based analysis comprised the between-subjects factor *lesion location* (i.e. frontal vs. temporal vs. stroke controls) and the within-subject factor *cue type* (i.e. semantic vs. phonological). The non-parametric Brunner-Munzel rank test was chosen for the analyses (i) as the distributions within the cells of the factorial design in some cases deviated from normality, and (ii) to keep the approach for the region-based analyses identical to the subsequent voxel-based analyses for which single voxels most often violate the assumptions of a normal distribution and hence for parametric testing (Rorden et al., 2007). Given that the factor *lesion location* comprised three levels, a significant main effect was planned to be followed up by two predefined contrasts that tested for worse performance in the frontal or temporal subgroup against the stroke control patients (one-tailed with an α -level of 0.05). Significant effects for the interaction were planned to be followed up by predefined contrast analyses as explicated in more detail in Section 2.6 (see below). Contrasts were computed using the non-parametric Brunner-Munzel rank test.

In a second series of voxel-based analyses, differential VLBM analyses for semantic and phonological fluency were performed using the recently developed NIX-toolbox (Nitschke et al., forthcoming; <https://github.com/kainitschke/NIX>). The NIX ('Non-parametric Interaction Effects')-toolbox is an open-source toolbox implemented in MATLAB which enables voxel-wise testing of interaction effects in non-normally distributed data. As for the region-based analysis (see above), also the VLBM analyses were based on the ANOVA-type extension of the non-parametric Brunner-Munzel rank test (Brunner and Munzel, 2002) for multi-factorial designs with within- and between-subjects factors. More specifically, the present VLBM analyses comprised the between-subjects factor *lesion* (i.e. lesion vs. no lesion) and the within-subjects factor *cue type* (i.e. semantic vs. phonological fluency). Only voxels with 5 or more patients per group (e.g. lesion vs. no lesion) with a cluster size of 10 voxels (33.75 mm^3) or larger and passing an uncorrected threshold of $p_{\text{uncorr}} < 0.001$ were considered significant (cf. Dressing et al., 2018; Martin et al., 2016). Significant interaction effects of *cue type* by *lesion*

were planned to be followed up using predefined contrasts for voxel-wise establishing single dissociations (see also Section 2.6). These contrast analyses were conducted using the non-parametric Brunner-Munzel rank test as implemented in the NIX toolbox.

Given that significant interaction effects of the factors *cue type* by *lesion* in these voxel-wise analyses only allow to reveal single dissociations, the full picture of a double dissociation was tested in a separate analysis based on the data from two clusters, namely the main left frontal and the main left temporal cluster showing a significant interaction. As explicated below, patients either had a lesion in the main frontal or the main temporal cluster or in neither of them, but none of the patients had a lesion overlapping these two main clusters. Patients were consequently assigned to the *lesion location* (i.e., main frontal cluster, main temporal cluster, none/stroke controls) for which their individual lesion showed the highest overlap. Similar to the region-based analyses, the resulting model hence comprised a two-factorial design with the between-subjects factor *lesion location* (i.e. frontal vs. temporal vs. stroke controls) and the within-subject factor *cue type* (i.e. semantic vs. phonological). As the factor *lesion location* comprised three levels, a significant main effect was planned to be followed up by two predefined contrasts that tested for worse performance in the frontal or temporal subgroup against the stroke control patients (one-tailed; see also above). The rationale behind predefined contrasts for following up a significant interaction effect of *lesion location* by *cue type* is specified below (see Section 2.6).

2.6. Operationalization for establishing single and double dissociations

Several concepts were previously suggested for establishing a double dissociation to infer separate localizations for different cognitive functions (e.g. Teuber, 1955; Shallice, 1988; Crawford et al., 2003; Davies, 2010). A double dissociation emerges if a brain-injured patient P_1 is impaired on task T_A but performs normal on task T_B whereas a patient P_2 with a different lesion site shows the opposite pattern with normal performance on task T_A and impairment in task T_B (see also Davies, 2010). However, as this definition of a *classical double dissociation*¹ (Teuber, 1955; Shallice, 1988) crucially hinges on the applied reference for the normal range, it was extended by the requirement for a significant *within-patient comparison* (Crawford et al., 2003; Davies, 2010) so that, in addition to the pattern above, patient P_1 performs task T_B significantly better than task T_A and that patient P_2 performs task T_A significantly better than task T_B .

In the present paper, we therefore adopted the combined approach of testing for a significant *classical double dissociation* (Shallice, 1988) that also requires significant *within-patient comparisons* (Crawford et al., 2003). More specifically, the implied double dissociation between left frontal and left temporal brain lesions and differential performance impairments in phonological and semantic fluency, respectively, was tested in both the region-based and the voxel-based lesion-behavior mapping analyses as follows: In a first step, we tested for a significant cross-over interaction effect between the two factors *cue type* (semantic vs. phonological) and *lesion location* (frontal vs. temporal vs. stroke controls). Provided a significant interaction was revealed, we then tested the requirements for a *classical double dissociation* in terms of a priori defined contrasts between the levels of the factor *lesion location* (stroke controls > frontal and stroke controls > temporal) separately for the two levels of the factor *cue type*. In addition, we also formulated and tested one-sided contrasts reflecting the required *within-patient comparison* separately for the two levels of interest of the factor *lesion location*, namely phonological < semantic fluency for frontal patients

and semantic < phonological fluency for temporal patients. That is, given that all these contrasts were based on a priori hypotheses and formulated in a directed fashion, the respective statistical tests were consequently one-tailed. *Within-patient comparisons* of semantic and phonological fluency in the stroke control patients were however formulated in a two-tailed fashion given that no a priori expectations existed. In this respect, the stroke control patients hence served as an unbiased assessment of potential differences in task difficulty between the two cue types.

For the VLBM analyses, the *single dissociations* emerging at the voxel-level were planned to be followed up in a similar fashion: Providing a significant interaction of the factors *cue type* by *lesion*, we then tested for single dissociations (see Section 2.5) using two a priori defined contrasts that tested one-sided for worse performance in the presence of a lesion separately for the two levels of the factor *cue type*. Within-patient comparisons for the two levels of cue type were only conducted in patients with a lesion in a given cluster with one-sided contrasts being complementary for temporal (semantic > phonological) and frontal clusters (phonological > semantic).

3. Results

3.1. Region-based lesion-behavior analysis

In order to investigate the implied double dissociation in classical region-based lesion-behavior analyses, patients were classified with respect to the primary location of their lesion. Of the 85 chronic stroke patients $n = 24$ and $n = 13$ had a lesion primarily in the left frontal and the left temporal lobe, respectively, whereas the remaining patients had primarily parietal ($n = 17$) and subcortical lesions ($n = 31$) and were considered as stroke control group ($n = 48$) in the following. The individual lesion overlays of the three *lesion location* groups (left temporal, left frontal, stroke controls) are depicted in Fig. 1B, C, and D, respectively. The factorial extension of the non-parametric Brunner-Munzel rank-order test with the between-subjects factor *lesion location*, the within-subjects factor *cue type* (semantic vs. phonological), and the total number of words produced as dependent variable revealed a significant two-way interaction of *cue type* \times *lesion location* ($F_{1.85, \infty} = 6.310$, $p = .003$) whereas the main effects of *lesion location* ($F_{1.66, \infty} = 1.875$, $p = .161$) and *cue type* ($F_{1.00, \infty} = 2.459$, $p = .117$) failed to reach significance.

For the significant interaction effect as illustrated in Fig. 2A, we first tested whether the above formulated contrasts fulfilled the requirements for a *classical double dissociation*: Patients with a left temporal lesion (red) as compared to patients in the stroke control group (gray) produced significantly fewer words in the semantic fluency condition (rank mean \pm standard error of the rank mean, $M_R \pm SEM_R$, 71.89 ± 12.70 vs. 97.91 ± 5.49 , $p = .049$) but showed no significant differences in the phonological fluency condition ($M_R \pm SEM_R$, 83.89 ± 15.44 vs. 93.15 ± 8.05 , $p = .257$). The opposite pattern emerged for the patients with a left frontal lesion (blue) who – when compared to patients in the stroke control group (gray) – produced significantly fewer words in the phonological fluency condition ($M_R \pm SEM_R$, 54.46 ± 10.47 vs. 93.15 ± 8.05 , $p = .004$), but showed no significant difference in the semantic fluency condition ($M_R \pm SEM_R$, 84.69 ± 8.09 vs. 97.91 ± 5.49 , $p = .108$). In addition, we further computed the above formulated contrasts for the *within-patient comparison* of the two fluency types separately for the left frontal and the left temporal patients: Patients with a left temporal lesion (red) showed a trend for producing fewer words in the semantic as compared to the phonological fluency condition ($M_R \pm SEM_R$, 71.89 ± 12.70 vs. 83.89 ± 15.44 , $p = .054$), whereas patients with a left frontal lesion (blue) produced significantly fewer words in the phonological as compared to the semantic fluency condition ($M_R \pm SEM_R$, 54.46 ± 10.47 vs. 84.69 ± 8.09 , $p < .001$). For patients in the stroke control group (gray), no significant difference was found between the semantic

¹ Please note that the definition of a classical double dissociation in the sense of Shallice (1988) is based on the comparison of patients with normal controls, whereas in the present study the patients of interest with primarily frontal and temporal lesions were contrasted with stroke control patients.

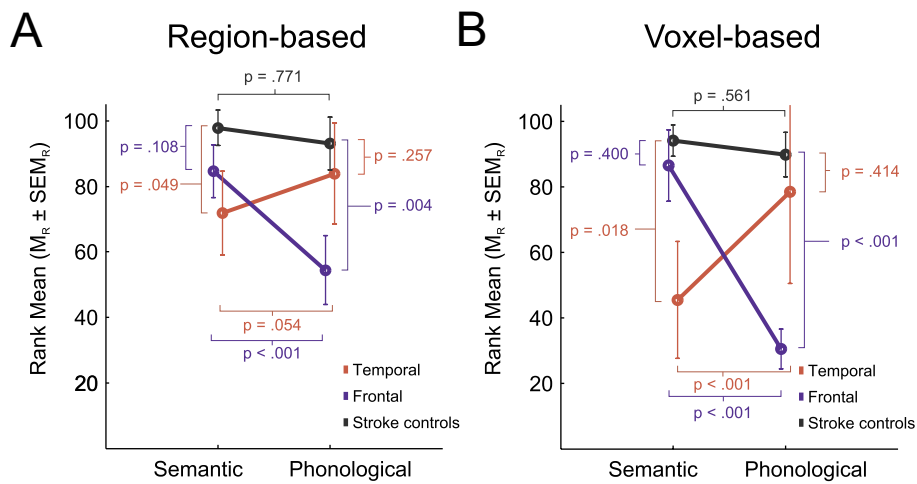


Fig. 2. Illustrations of the double dissociation based on the significant two-way interaction *cue type* \times *lesion location* as revealed in (A) the region-based lesion-behavior analysis and in (B) the voxel-based lesion-behavior mapping analysis.

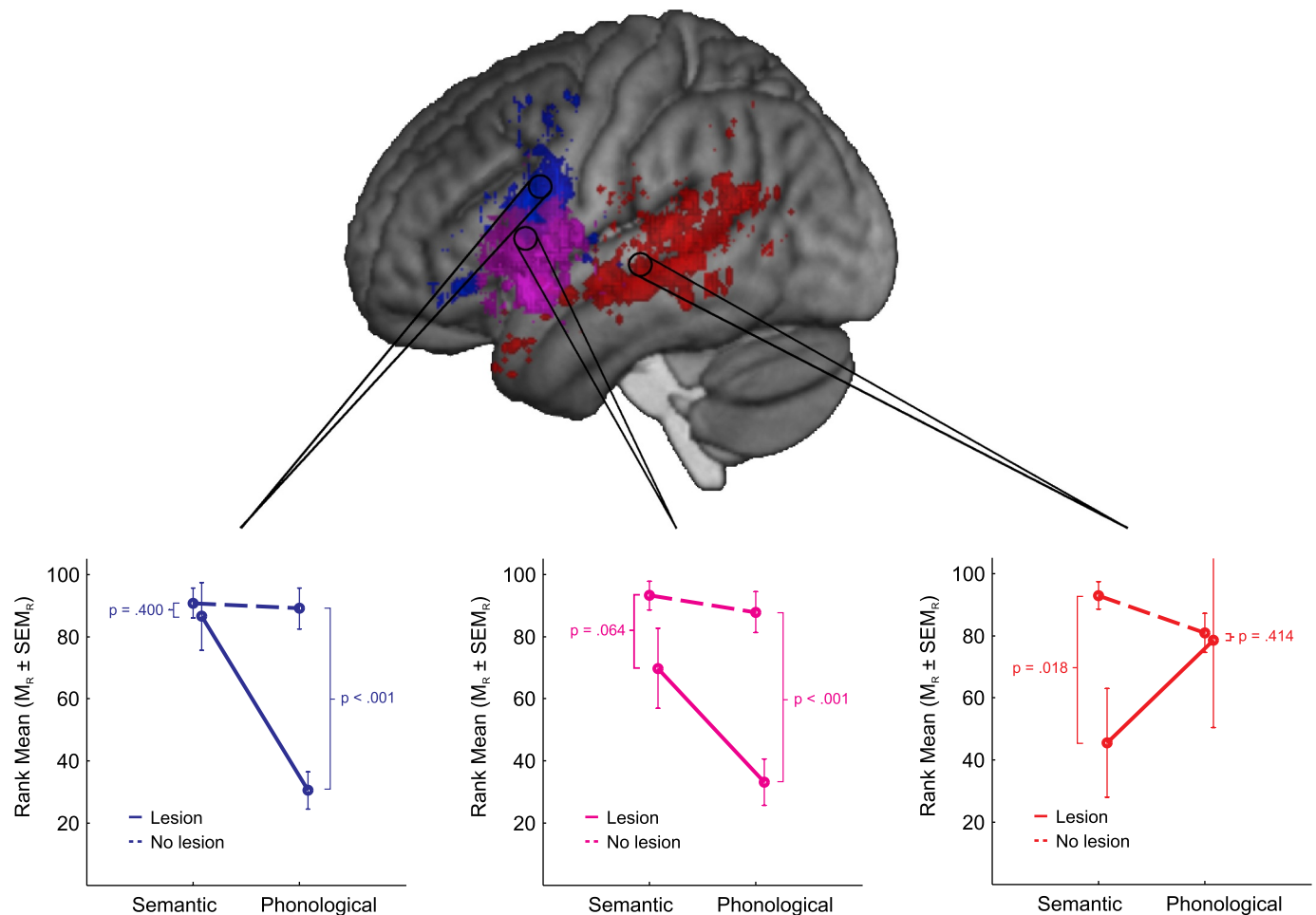


Fig. 3. Overview of the voxel-wise results in the voxel-based lesion-behavior mapping analysis. Voxels colored in magenta indicate the distribution of the main effect of *lesion* with significant differences between patients with and without a lesion irrespective of the type of verbal fluency. Voxels colored in red and blue indicate the distribution of the significant interaction effect of *lesion* by *cue type* and the two resulting single dissociations with their opposing impairments in semantic and phonological fluency, respectively. Note that only voxels passing a threshold of $p_{\text{uncorr}} < 0.001$ are displayed. The three panels below the brain rendering illustrate the observed patterns of lesion effects on performance in semantic and phonological fluency for each of the main clusters for the main effect of lesion (magenta) as well as for the directions of the interaction effect (red, blue). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 2
Overview of clusters with significant voxels for the main effect of lesion.

#	Cluster descriptives			Patient Groups				Behavior																																																																																																																																																																																																																																																																																					
	Peak	Anatomical Distribution	Size	w/o lesion		w/ lesion		Semantic Fluency		Phonological Fluency		Contrasts																																																																																																																																																																																																																																																																																	
				w/o lesion	n	w/ lesion	n	M _R ± SEM _R	M _R ± SEM _R	w/o lesion (1)	w/ lesion (2)	M _R ± SEM _R	M _R ± SEM _R	w/ lesion (3)	w/ lesion (4)	(2) < (1)	(4) < (3)	(2) < (4)	(4) < (2)																																																																																																																																																																																																																																																																										
x,y,z	WFU Pick Atlas labels	k	n	n	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	M _R

Note. w/o, without; w/, with; k, cluster size in voxels; $M_R \pm SEM_M$, rank mean \pm standard error of the rank mean; n.a., not assessed. Coordinates of peak voxels (x,y,z) are provided in MNI space. Anatomical labels were specified based on the predefined regions-of-interest (ROIs) for the frontal and temporal lobes (Lancaster et al., 2000) as implemented in the WFU PickAtlas (Maldjian et al., 2003) available for SPM8. The percentage of overlap for the cluster with the ROI. Note that clusters consisted of voxels passing a threshold of $p_{uncorr} < 0.001$ and had a minimum size of $k > 10$ voxels. Note that in clusters #3, #8, and #9 ($p < .05$) or at least marginally significant ($p < .10$) one-sided contrasts (w/ lesion < w/o lesion) for both cue types are highlighted with p-values in bold font. Clusters revealing significant differences in the opposite direction, thus resulting in p-values $> .999$ for the predefined contrasts testing for a difference in the opposite direction. Clusters revealing significant differences in the opposite direction, thus resulting in p-values $> .999$ for the predefined contrasts testing for a difference in the opposite direction.

Table 3
Overview of clusters with significant voxels for the interaction cue type × lesion.

#	Cluster Descriptives			Behavior									
				Patient Groups		Semantic Fluency		Phonological Fluency		Contrasts			
				w/o lesion	w/ lesion	w/o lesion (1)	w/ lesion (2)	w/o lesion (3)	w/ lesion (4)	(2) < (1)	(4) < (3)	(2) < (4)	(4) < (2)
Peak	Anatomical Distribution			Size									
x,y,z	WFU Pick Atlas labels			k	n	n	M _R ± SEM _R	M _R ± SEM _R	M _R ± SEM _R	p	p	p	p
1	-60.0, -55.5, 21.0	Superior temporal gyrus (44.2%)	middle temporal gyrus (34.8%)	2895	80	5	92.98 ± 4.36	45.50 ± 17.85	80.95 ± 6.32	78.50 ± 27.99	0.018	0.414	< 0.001 n.a.
2	-54.0, 1.5, 24.0	Precentral gyrus (40.0%)	inferior frontal gyrus (21.1%)	916	73	12	90.79 ± 4.81	86.54 ± 10.88	89.08 ± 6.58	30.50 ± 6.07	0.400	< 0.001 n.a.	< 0.001
3	-48.0, 31.5, -12.0	inferior frontal gyrus (68.9%)	middle frontal gyrus (6.2%)	180	80	5	88.99 ± 4.61	109.40 ± 6.79	84.11 ± 6.31	28.00 ± 8.64	> 0.999	0.008 n.a.	< 0.001
4	-57.0, -10.5, -30.0	Middle temporal gyrus (63.2%)	inferior temporal gyrus (35.5%)	76	80	5	92.98 ± 4.36	45.50 ± 17.85	80.95 ± 6.32	78.50 ± 27.99	0.018	0.414	< 0.001 n.a.
5	-40.5, 21.0, -31.5	Superior temporal gyrus(85.5%)		62	79	6	92.22 ± 4.51	63.50 ± 15.86	79.51 ± 6.42	97.92 ± 20.32	0.064	> 0.999	< 0.001 n.a.
6	-37.5, -3.0, 52.5	Middle frontal gyrus (73.2%)		56	78	7	90.83 ± 4.67	83.07 ± 11.65	86.05 ± 6.32	22.36 ± 6.42	0.241	< 0.001 n.a.	< 0.001
7	-52.5, 24.0, 7.5	Inferior frontal gyrus (64.1%)	middle frontal gyrus (35.9%)	39	80	5	89.94 ± 4.61	94.30 ± 11.73	83.55 ± 6.38	36.90 ± 6.18	> 0.999	0.010 n.a.	< 0.001
8	-57.0, -63.0, 16.5	Middle temporal gyrus (65.7%)	middle occipital gyrus (28.6%)	35	80	5	90.02 ± 4.39	93.00 ± 27.93	79.38 ± 6.29	103.70 ± 27.39	> 0.999	> 0.999	< 0.001 n.a.
9	-45.0, 18.0, -31.5	Superior temporal gyrus(94.1%)		34	80	5	92.98 ± 4.36	45.50 ± 17.85	80.95 ± 6.32	78.50 ± 27.99	0.018	0.414	< 0.001 n.a.
10	-31.5, 1.5, 31.5	Not assignable, frontal white matter		34	76	9	91.25 ± 10.47	81.28 ± 9.76	87.02 ± 9.98	28.33 ± 6.32	0.151	< 0.001 n.a.	< 0.001
11	-40.5, 13.5, 36.0	Middle frontal gyrus (83.9%)		31	80	5	90.06 ± 4.59	92.30 ± 14.50	84.77 ± 6.23	17.40 ± 7.96	> 0.999	0.002 n.a.	< 0.001
12	-36.0, 43.5, -1.5	Inferior frontal gyrus (33.3%)		24	80	5	88.99 ± 4.61	109.40 ± 6.79	84.11 ± 6.31	28.00 ± 8.64	> 0.999	0.008 n.a.	< 0.001
13	-52.5, -64.5, 1.5	Middle temporal gyrus (100.0%)		17	78	7	93.40 ± 4.43	54.43 ± 15.04	81.01 ± 6.50	78.50 ± 18.29	0.015	0.418	< 0.001 n.a.
14	-55.5, 6.0, -28.5	Superior temporal gyrus (50.0%)	middle temporal gyrus (50.0%)	12	80	5	92.98 ± 4.36	45.50 ± 17.85	80.95 ± 6.32	78.50 ± 27.99	0.018	0.414	< 0.001 n.a.
15	-34.5, 10.5, 49.5	Middle frontal gyrus (16.7%)		12	80	5	89.59 ± 4.60	99.90 ± 12.06	84.37 ± 6.28	23.80 ± 8.38	> 0.999	0.004 n.a.	< 0.001

Note. w/o, without; w/, with; k, cluster size in voxels; M_R ± SEM_R, rank mean ± standard error of the rank mean; n.a., not assessed. Coordinates of peak voxels (x,y,z) are provided in MNI space. Anatomical labels were specified based on the predefined regions-of-interest (ROIs) for the frontal and temporal lobes (Lancaster et al., 2000) as implemented in the WFU PickAtlas (Maldjian et al., 2003) available for SPM8. The percentage indicates the amount of overlap for the cluster with the ROI. Note that clusters consisted of voxels passing a threshold of P_{uncorr} < 0.001 and had a minimum size of k > 10 voxels. Clusters in frontal and temporal cortex significantly (p < .05) or at least marginally significantly (p < .10) conforming to the criteria for single dissociations in the expected directions are highlighted with p-values in bold font.

compared to the phonological fluency condition ($M_R \pm SEM_R$, 97.91 ± 5.49 vs. 93.15 ± 8.05 , $p = .771$) (see Fig. 2A).

Given that the stroke controls comprised a heterogeneous selection of patients with parietal and subcortical lesions (see also Supplementary Fig. S1–1), the present results might be at least partly driven by the larger variance in the lesion distribution of the reference group of stroke controls ($n = 48$) compared to the patients of interest with lesions in either left frontal ($n = 24$) or left temporal lobe ($n = 13$). In order to preclude this potential bias, we repeated the above analyses only for the patients with parietal ($n = 17$) or for the patients with subcortical lesions ($n = 31$) as reference group of stroke controls (see Supplementary Materials, Sections S1). These control analyses corroborated the double dissociation, which did also hold for separately testing in the two subgroups of patients with parietal and subcortical lesions (Supplementary Fig. S1–2).

Taken together, the results of the region-based lesion-behavior analysis conform to the requirements for establishing a double dissociation both in the classical sense as well as based on significant within-patient comparisons or a trend thereof. These data hence strongly support the notion that lesions to the left temporal and left frontal lobe differentially affect performance in semantic and phonological fluency, respectively (Fig. 2A). Notably, this double dissociation cannot be attributed to potential differences in task difficulty between the two types of verbal fluency as these were effectively controlled for (Fig. 2A, stroke control group). The anatomical specificity of these findings was elucidated in subsequent voxel-based lesion-behavior mapping analyses.

3.2. Voxel-based lesion-behavior mapping (VLBM) analyses

In a second series of voxel-based analyses a VLBM analysis with the factor *lesion* (yes vs. no) and *cue type* (semantic vs. phonological) revealed voxels with a significant main effect of *lesion* as well as voxels with a significant interaction of *cue type* \times *lesion*.

The spatial distribution of the significant main effect of *lesion* (Fig. 3; depicted in magenta) comprised voxels mainly in left frontal cortex that showed a significant difference between patients with and without a lesion irrespective of *cue type* (semantic or phonological). Note that in order to identify voxels showing a main effect that was not solely driven by a potentially underlying (subthreshold) interaction effect, the assessment of the significant main effect ($p < .001$) was restricted only to voxels concurrently showing a p -value for the test of an interaction effect above $p > .05$. For most clusters of voxels patients with a lesion performed worse as compared to patients without a lesion. Subsequent contrast analyses further confirmed that in a majority of clusters this pattern was evident irrespective of the type of fluency, although the relative effect sizes often appeared stronger for phonological fluency. Detailed information on the location, extent, and direction of the main effect of *lesion* is reported in Table 2.

The spatial distribution for the significant two-way interaction for *cue type* \times *lesion* revealed anatomically distinct patterns of two different *single dissociations*: Performance differences between patients with vs. without a lesion in semantic but not in phonological fluency (Fig. 3; depicted in red) were primarily located in left temporal brain areas, whereas differential performance differences between patients with vs. without a lesion in phonological but not in semantic fluency (Fig. 3; depicted in blue) were primarily found in left frontal brain areas. Detailed information on the location, extent, and direction of the two-way interaction for *cue type* \times *lesion* as well as on the contrast analyses relevant for establishing single dissociations is presented in Table 3.

In order to further assess whether the single dissociations in left frontal and left temporal areas indeed reflected a double dissociation, the data from the two largest clusters in left frontal and left temporal cortex from the significant *cue type* \times *lesion* interaction were used to test for the three-way interaction of *lesion* (yes vs. no) \times *location*

(temporal vs. frontal) \times *fluency type* (semantic vs. phonological) (see Table 3). However, as there was no patient with a lesion in both the frontal and temporal cluster, the $2 \times 2 \times 2$ design was incomplete and therefore rearranged into a 2×3 design with *cue type* (semantic vs. phonological) \times *lesion location* (patients with a lesion in main left temporal cluster, patients with a lesion in main left frontal cluster, remaining patients as stroke controls). The extended Brunner-Munzel rank-order test for this 2×3 design revealed non-significant main effects of *cue type* ($F_{1,00,\infty} = 2.498$, $p = .114$) and *lesion location* ($F_{1,21,\infty} = 1.726$, $p = .189$) but, most importantly, a significant two-way interaction for *cue type* \times *lesion location* ($F_{1,73,\infty} = 20.083$, $p < .001$) fulfilling the requirements of a *classical double dissociation*: As illustrated in Fig. 2B, patients with a lesion in the main left temporal cluster (red) as compared to the stroke control patients (gray) produced significantly fewer words in the semantic fluency condition ($M_R \pm SEM_R$, 45.5 ± 17.85 vs. 94.13 ± 4.78 , $p = .018$), whereas no significant difference was found for the phonological fluency condition ($M_R \pm SEM_R$, 78.5 ± 27.99 vs. 89.85 ± 6.18 , $p = .414$). In contrast, patients with a lesion in the main left frontal cluster (blue) as compared to the stroke control patients (gray) produced significantly fewer words in the phonological fluency condition ($M_R \pm SEM_R$, 30.5 ± 6.07 vs. 89.85 ± 6.81 , $p < .001$) but not in the semantic fluency condition ($M_R \pm SEM_R$, 86.5 ± 10.88 vs. 94.13 ± 4.78 , $p = .400$). The *within-patient comparisons* further revealed that patients with a lesion in the main left temporal cluster (red) produced significantly fewer words in the semantic as compared to the phonological fluency condition ($M_R \pm SEM_R$, 45.5 ± 17.85 vs. 78.5 ± 27.99 , $p < .001$), whereas the opposite pattern was found for the patients with a lesion in the main left frontal cluster (blue) who produced significantly fewer words in the phonological as compared to the semantic fluency condition ($M_R \pm SEM_R$, 30.5 ± 6.07 vs. 86.5 ± 10.88 , $p < .001$). For the stroke control patients (gray), no significant difference was found between the semantic and the phonological fluency condition ($M_R \pm SEM_R$, 94.13 ± 4.78 vs. 89.85 ± 6.81 , $p = .561$) (Fig. 2B).

Taken together, the results of the voxel-based lesion-behavior mapping analysis not only corroborate the pattern of a double dissociation (Fig. 2B) that was established in the region-based analysis (Fig. 2A) but also substantially extend these findings due to the higher anatomical specificity of the VLBM approach. That is, differentially impaired performance in semantic fluency compared to phonological fluency was particularly observed for lesions of the left middle and superior temporal gyri (Fig. 3, red), whereas the opposite pattern with differentially impaired performance in phonological fluency compared to semantic fluency was observed following lesions of the left inferior frontal gyrus (IFG) (Fig. 3, blue) with a particular focus on pars opercularis but also on pars orbitalis. Beyond this double dissociation, it also became evident that lesions in other parts of left IFG, namely pars triangularis (Fig. 3, magenta), led to impairments in both types of verbal fluency.

4. Discussion

The present study revealed strong evidence for the proposed double dissociation at the group level between differential impairments in semantic and phonological fluency and lesions in left temporal and left frontal lobe (Fig. 2). To our knowledge this is the first study that explicitly addressed and statistically demonstrated this double dissociation in a large sample of chronic stroke patients using both region-based as well as voxel-based lesion-behavior mapping analyses. Our findings hence corroborate previous assumptions from both the lesion and neuroimaging literature which suggested this double dissociation at the group level (e.g., Mummery et al., 1996; Baldo et al., 2006, 2010). However, the present study goes substantially beyond the experimental demonstration of a since long proposed double dissociation: The results from the VLBM analysis further revealed a significant main effect for *lesion* particularly in the pars triangularis of the left IFG, showing that

patients with a lesion in these voxels performed worse than patients without a lesion irrespective of the type of verbal fluency (semantic or phonological) (see Fig. 3). Taking the double dissociation as evidence for a coarse attribution of cognitive processing in semantic and phonological fluency to the functions of the temporal and frontal cortex, respectively, thus seems to be an oversimplification.

4.1. Double dissociation between left frontal and temporal lesions and impairments in phonological and semantic fluency

The present study's main objective concerned the statistical demonstration of doubly-dissociating differential contributions of left frontal and temporal cortex in phonological and semantic fluency that were previously implied but had not been shown explicitly. In this respect, one might of course argue that – in view of partial evidence for the underlying single dissociations (e.g., Milner, 1964; Perret, 1974; Henry and Crawford, 2004; Baldo et al., 2006, 2010) – incidental proof for the double dissociation at the group level is little surprising but coercive. However, although previous lesion studies have demonstrated that comparing healthy controls to patients with left frontal and temporal lesions resulted in impairments in phonological (Perret, 1974; Pendleton et al., 1982; Miller, 1984; Janowsky et al., 1989; Stuss et al., 1994; Gershberg and Shimamura, 1995; Tucha et al., 1999; Channon and Crawford, 2000; Jurado et al., 2000; Baldo et al., 2001; Sylvester and Shimamura, 2002) and semantic fluency (Martin et al., 1990; Troyer et al., 1998; Luckhurst and Lloyd-Jones, 2001), respectively, the meta-analysis of Henry and Crawford (2004) has pointed out that the sensitivity of semantic fluency for frontal lesions is comparable to that of phonological fluency (cf. Stuss et al., 1996, 1998, 1999; Baldo and Shimamura, 1998; Baldo and Shimamura, 1998; Rogers et al., 1998; Troyer et al., 1998; Szatkowska et al., 2000; Robinson et al., 2012; Biesbroek et al., 2016). In addition, previous research has provided only limited direct evidence for differential patterns in terms of significant *within- or between-patient comparisons* which were neither independently nor fully covering the overall pattern of the proposed double dissociation (but see Baldo et al., 2006, 2010, for notable exceptions). In this respect, patients with frontal lesions were reported to have worse performance in phonological fluency compared to semantic fluency (Rogers et al., 1998) and compared to patients with temporal lesions (Milner, 1964; Perret, 1974; Helmstaedter et al., 1998). Thus, the (implicit) evidence for the double dissociation from previous studies has been less clear than it may appear at first glance. Furthermore, while most previous lesion studies relied on region-based lesion-behavior mapping on the level of lobes, the findings on contributions of left frontal cortex in semantic fluency contrasting the proposed pattern of a double dissociation particularly indicated the need for analyses at a higher spatial resolution to resolve the equivocality of potentially distinct and shared neural correlates of semantic and phonological fluency in left frontal cortex.

As of yet, only four studies have applied voxel-based lesion-behavior mapping (VLBM) analyses either in chronic (Baldo et al., 2006) or (sub) acute stroke patients (Biesbroek et al., 2016; Chouiter et al., 2016) or in patients with low-grade glioma (Almairac et al., 2015) that were assessed with both semantic and phonological fluency. In line with the present results, Baldo et al. (2006) revealed differential associations of impaired performance in phonological and semantic fluency with lesions primarily in left frontal (mainly precentral gyrus and IFG pars opercularis) and left temporal cortex (mainly superior and middle temporal gyrus), respectively, but due to the unavailability of appropriate statistical tools they provided only qualitative subtraction maps for visualization of the differential lesion distributions and did not explicitly test for the significance of the apparent double dissociation. Additional analyses by Baldo et al. (2006) further implicated that performance in hard semantic items also relied on left frontal cortex. Almairac et al. (2015) reported associations of deficits in semantic fluency with lesions in the deep white matter underlying the left IFG

and left superior temporal gyrus, the deep sylvian fissure, the posterior orbito-frontal areas, the striatum and in the insula, whereas no significant associations were found for phonological fluency. Biesbroek et al. (2016) reported again overlapping anatomical correlates in the left insula and left frontal (inferior frontal, middle frontal, and precentral gyri, rolandic operculum) cortex and discordant correlates of semantic and phonological fluency in left temporal (inferior temporal, lingual, and fusiform gyri, medial temporal lobe) and left frontal cortex (middle frontal gyrus), respectively, but they did also not statistically compare these differential patterns. Chouiter et al. (2016) reported associations with lesions shared by both fluency types that however concerned left parietal (angular and parts of supramarginal gyrus) and left temporal cortex (superior and middle temporal gyri). In addition, semantic fluency was differentially associated with lesions in middle temporal gyrus and the pallidum, whereas phonological fluency was associated with lesions in anterior middle and superior temporal areas, the rolandic operculum and the supramarginal gyrus (Chouiter et al., 2016).

Although none of these studies directly assessed the double dissociation in statistical terms, the findings from these VLBM studies either are fully in line (Baldo et al., 2006) or at least partly concur (Almairac et al., 2015; Biesbroek et al., 2016; Chouiter et al., 2016) with the anatomy underlying the present differential patterns. Yet, as systematic differences of task difficulty between phonological and semantic fluency have not been explicitly controlled for in any of these studies (see also Table 1), it cannot be formally excluded that findings of differential contributions in these VLBM studies could be fully or partly driven by the consistently reported higher task difficulty of phonological fluency and might hence reflect an unspecific task-resource artifact (cf. Shallice, 1988; Davies, 2010). The same argument holds also true for all other fluency studies which applied region-based lesion-behavior mapping approaches without controlling for the immanent differences in task difficulty between both types of verbal fluency (Table 1).

Taken together and to the best of our knowledge, the present study hence comprises the first explicit evidence for the double dissociation at the group level between differential contributions of left frontal and left temporal cortex in phonological and semantic fluency, respectively. The present findings also highlight the value of applying non-parametric factorial analyses in the context of lesion studies (cf. Nitschke et al., forthcoming). In this respect, by exploiting the higher spatial resolution of voxel-based compared to region-based lesion-behavior mapping analyses, the present study further allowed to resolve apparent discrepancies on distinct versus shared contributions of left frontal areas in both fluency types which are discussed below.

4.2. Distinct and shared contributions of left IFG in phonological and semantic fluency

The present results indicate that specificity and/or sensitivity of frontal lesions for deficits in the two different types of verbal fluency appears to hinge on their precise location. That is, while lesions in the pars opercularis (and partly in pars orbitalis) of the IFG lead to an isolated impairment in phonological fluency, lesions in pars triangularis lead to general impairments in both semantic and phonological fluency (Fig. 3). In this respect, several previous lesion studies have suggested that the left frontal cortex and in particular the left IFG is not only involved in phonological but also in semantic fluency (e.g., Troyer et al., 1998; Reverberi et al., 2006; see Henry and Crawford, 2004, for a meta-analytic overview). However, most of these region-based lesion-behavior mappings were conducted at the level of lobes (but see Robinson et al., 2012; Biesbroek et al., 2016) and hence did not allow differentiating between shared or distinct neural correlates (or both) of phonological and semantic fluency in the left frontal lobe.

Anatomically more specific, Costafreda et al. (2006) proposed functional segregation of the left IFG in semantic and phonological

processing. Based on a systematic review of functional MRI studies of verbal fluency, they suggested that the posterior-dorsal part of the left IFG (Brodmann area [BA] 44; roughly corresponding to pars opercularis) is associated with phonological fluency, whereas the anterior-ventral part of the left IFG (BA 45; roughly corresponding to pars triangularis) is associated with semantic fluency (Costafreda et al., 2006). Testing the proposed dissociation directly in a within-subjects fMRI experiment, Heim et al. (2008) only found overlapping activations in BA 45 for phonological and semantic fluency, whereas the former led to specific activation in BA 44 (see also Wagner et al., 2014). Although this seems partly in contrast with the proposal of Costafreda et al. (2006), it nonetheless appears to mirror the present findings as well as those of Biesbroek et al. (2016), who also reported common impairments in semantic and phonological fluency following lesions in left anterior IFG (pars triangularis, BA 45; see also Robinson et al., 2012) and specific phonological impairments following lesions in left posterior IFG (pars opercularis, BA 44).

Baldo et al. (2006) and Reverberi et al. (2006) suggested that the degree of frontal involvement in semantic fluency may rely on the task demands exerted by different categories, as an efficient strategy may particularly affect fluent word retrieval from smaller categories (i.e. hard semantic items). Taking item difficulty and individual ability into account in a recent fMRI study in healthy participants, Katzev et al. (2013) were able to elaborate on these two determinants of the proposed differential involvement of BA 44 and 45 for phonological and semantic fluency, respectively (cf. Costafreda et al., 2006): Higher activation for semantic than phonological fluency was evident in posterior BA 45 particularly in low-performing respondents and in anterior BA 45 only for semantic hard items, whereas higher activation for phonological than semantic fluency was robustly found in BA 44, but to a larger extent in high-performing respondents (Katzev et al., 2013). Compared against implicit baseline, BA 44 and posterior BA 45 were rather selectively activated by phonological or semantic fluency, respectively. In contrast, anterior BA 45 was activated by both types of fluency with the strongest activations for semantic hard items, intermediate levels for phonological hard and easy items, but with no significant activation for semantic easy items.

The present finding of impairments in (comparably difficult) semantic hard items and phonological easy items following left IFG lesions in pars triangularis (BA 45) hence concur with the fMRI data of Katzev et al. (2013). Although neither the focus nor part of the present analyses, one would further expect differential impairment comparing semantic hard vs. easy items, since the latter did not yield any activation of BA 45 in the study of Katzev et al. (2013). Exploratory VLBM analyses in semantic hard and easy items (not shown) however failed to reveal such an interaction between lesion and item difficulty in addition to (or instead of) the here reported main effect of lesion in left IFG pars triangularis.

Besides a potential lack of statistical power, the substantial age differences between the samples of Katzev et al. (mean \pm SD, 26.1 \pm 6.6 years) and the present study (63.97 \pm 14.20 years) might indicate that semantic easy items possibly exert also higher task demands on strategically controlled retrieval in older adults (in addition to semantic hard items and at least partly reliant on anterior BA 45) as compared to the rather automatically triggered responses in semantic easy items in younger adults (without essential recruitment of anterior BA 45). In accord with this, older adults have been reported to generate smaller clusters and to switch less frequently in semantic fluency than younger adults (Troyer et al., 1997; Lanting et al., 2009; but see Mayr and Kliegl, 2000). Directly comparing fMRI activation of older and younger adults has further demonstrated higher activation for phonological compared to semantic fluency in BA 44 and anterior BA 45 in younger adults, whereas older adults failed to show this distinction but recruited these areas in both conditions to a similar extent (Meinzer et al., 2009). As task difficulty of phonological and semantic fluency in this fMRI study was not matched with the latter being significantly

easier than the former in the younger but not the older adults (Meinzer et al., 2009) this finally closes the circle towards the differential findings of Katzev et al. (2013) in younger adults by again implying that the performance of older healthy adults and most likely also older stroke patients (as in the present sample) relies on BA 45 not only for semantic hard but also for semantic easy items.

Taken together, the present voxel-based findings of shared contributions of left IFG pars triangularis in semantic and phonological fluency not only render previous region-based attributions more precisely but also concur with and extend previous insights from neuroimaging studies. The potentially underlying cognitive correlates of dissociable and overlapping lesion-behavior mappings will be discussed in the following.

4.3. Cognitive correlates underlying lesion-specific impairments in verbal fluency

Dissociable neural contributions in left frontal and left temporal cortex most likely reflect marked differences in the retrieval processes underlying phonological and semantic fluency (Katzev et al., 2013) – particularly as a potential task resource artifact can be excluded. The present differential effect for lesions in left IFG pars opercularis (or BA 44) leading to specific impairment in phonological fluency is hence likely to be related to the serial search based on systematic syllabification of initial letters (Mummery et al., 1996; Rende et al., 2002; Henry and Crawford, 2004). In this respect, it was previously suggested that the left IFG performs sensorimotor encoding of auditory phonetic input (Demonet et al., 1994; Bookheimer, 2002) which overlaps with processes of inner speech such as motor programming and articulation (Indefrey and Levelt, 2000) and may thus also support proper subvocal syllabification of initial letters in phonological fluency.

In contrast to the rather artificial kind of search in phonological fluency, retrieval of words in semantic fluency tasks relies on the natural organization of conceptual knowledge stored in the temporal lobe (Gruenewald and Lockhead, 1980; Katzev et al., 2013). Already Gruenewald and Lockhead (1980) proposed that retrieving words from a given semantic category is a two-stage process consisting of (i) the top-down identification of a task-relevant subcategory as the foundation for (ii) the subsequent bottom-up retrieval of appropriate category members triggered by automatic associations (or semantic proximity) within that subcategory. Expanding on this and in an attempt to disentangle executive from semantic processes, Troyer et al. (1997) suggested that clustering in terms of the (automatic) word retrieval within subcategories and (controlled) switching between these subcategories constitute two important components of semantic fluency (see also Abwender et al., 2001) with the latter and the former being differentially associated with the left temporal and the left frontal lobe, respectively (Troyer et al., 1998). More specifically, switching between semantic subcategories has been related to the integrity of left frontal lobe (Troyer et al., 1998) and the functional activation of pars triangularis in particular (Hirshorn and Thompson-Schill, 2006). It has been further argued that not switching per se is impaired in patients with frontal lesion, but rather the application of an efficient search strategy as patients were found to exhibit increased switching and reduced semantic proximity (Reverberi et al., 2006). The concept of switching and clustering is however not specific for semantic fluency but can also be applied to phonological fluency to disentangle executive from lexical processes (Troyer et al., 1998), which – in light of the present findings – are likely to be reflected by lesion-behavior mappings with left IFG pars triangularis and pars opercularis, respectively.

With regard to the shared contributions in left IFG pars triangularis, mainly two different types of executive control processes have been discussed to be associated with semantic fluency (Katzev et al., 2013): either (post-retrieval) selection from among competing co-activated representations or words (Hirshorn and Thompson-Schill, 2006; see also Thompson-Schill et al., 1997, 1998) or top-down controlled

semantic search when bottom-up (automatic) retrieval (based on association chains; Collins and Loftus, 1975) is insufficient (e.g., Badre et al., 2005; Wagner et al., 2001). While the present lesion data do not allow deciding between these two alternative explanations, the fMRI data of Katzev et al. (2013) were in favor for the latter interpretation of the functional role of anterior BA 45 (but did not exclude the former for posterior BA 45; see also Robinson et al., 1998). This would also concur with the proposed role of BA 45 in switching between (semantic or phonological) subcategories (but see Hirshorn and Thompson-Schill, 2006, and Reverberi et al., 2006, for alternative accounts).

Furthermore, given that anterior BA 45 was found to exert intermediate activation in phonological hard and easy items (Katzev et al., 2013) and that also continuous word generation in phonological fluency is at least to some extent semantically structured (Abwender et al., 2001; Schwartz et al., 2003; Azuma, 2004), the present finding of a main effect of lesion in pars triangularis leading to general impairments in both semantic and phonological fluency may hence reflect impairments in top-down controlled semantic search and/or retrieval. However, it may alternatively signify overlapping deficits in domain-independent control processes as switching subcategories was found to be affected in phonological fluency following left lateral frontal lesions (Troyer et al., 1998). For instance, the post-retrieval selection hypothesis (Thompson-Schill et al., 1997, 1998) would predict deficits that are at least partly due to an impaired suppression of inappropriate responses automatically activated by semantic association chains which may similarly affect performance in phonological and semantic fluency (Katzev et al., 2013). In addition, overlapping impairments for both fluency types may also be directly caused or indirectly mediated by deficits in less specific domain-general control functions such as energization, self-monitoring, attention and processing speed (cf. Biesbroek et al., 2016).

Finally, it should be noted that evidence from the clustering and switching approach as put forward by Troyer et al. (1997, 1998) may not allow to unequivocally disentangle executive control processes from genuine semantic processes (Mayr and Kliegl, 2000; Mayr, 2002; see also Reverberi et al., 2006). In particular, the two resulting components are not independent (particularly if fluency is assessed in a time-restricted manner as usual; cf. Troyer et al., 1997) so that a reduced number of switches per se may not only reflect difficulties in (top-down controlled) accessing a new subcategory but may also depend on difficulties in (bottom-up driven automatic) semantic retrieval within subcategories (Mayr, 2002). In other words, the more time is spent on one category, the less time remains for switching to new categories. Following Mayr (2002), this interpretation uncertainty can only be resolved by the analysis of detailed timing protocols (i.e., within- and between-cluster retrieval latencies) – thus indicating not only the need but also the methodology for future research on the cognitive correlates underlying lesion-specific impairments in phonological and semantic fluency.

4.4. Limitations

There are several limitations to the interpretation of the present results. With respect to the formal criteria for claiming a valid double dissociation, Shallice (1988) argued that Teuber's (1955) formulation of a classical double dissociation (i.e. differential normal/impaired performance of two patients P_1 and P_2 in tasks T_A and T_B) is “not sufficient grounds for inferring separate subsystems” (p. 234) if the two tasks have different difficulties and may thus differentially call on given resources. He therefore proposed to consider the *between-patient comparison* with patient P_1 performing significantly worse than patient P_2 on task T_A and with patient P_2 performing significantly worse than patient P_1 on task T_B . In the present paper, we adopted an alternative but also established approach based on the extension of the tests for a *classical double dissociation* (on patient- and task-specific differential performance below normal) by significant *within-patient comparisons*

(Crawford et al., 2003; Davies, 2010) so that patient P_1 performs task T_B significantly better than task T_A and that patient P_2 performs task T_A significantly better than task T_B . This was done mainly for two reasons: First, by explicitly controlling for comparable levels of task difficulty between semantic and phonological fluency, we have directly addressed and resolved the problem of potential task resource artifacts put forward by Shallice (1988). Second, and more pragmatically, by accounting for systematic effects in inter-individual variance the *within-patient approach* is superior to the *between-patient approach* in terms of statistical test power, thus reducing the Type II error of false negative findings. Based on Monte-Carlo simulations McIntosh (2017) even argued that solely the significant *within-patient comparison* would be logically necessary to establish a (double) dissociation whilst sufficiently controlling for Type I error of false positive findings, whereas the additional tests for a *classical (double) dissociation* (i.e., impaired performance significant from normal on task T_A but not on task T_B in patient P_1 and vice versa in patient P_2 ; cf. Teuber, 1955; Crawford et al., 2003; Davies, 2010) applied here would only constitute further qualifying information.

Another limitation of the present study concerns the obvious ambiguity in attributing impairments in verbal fluency following circumscribed lesions (for instance, in anterior or posterior pars triangularis) to specific cognitive processes or functions (e.g., switching semantic/phonological subcategories vs. post-retrieval selection), which results at least partly from the fact that the majority of previous studies used rather coarse classifications on the level of lobes (e.g., frontal vs. non-frontal). This lacking spatial acuity hence hampered any direct reference between the present and previous results particularly as none of the few extant studies addressing impairments in verbal fluency on the level of processes had applied spatially resolved VLBMs analyses. In addition, all considerations with reference to microstructural cytoarchitecture (i.e., in terms of Brodmann areas) are purely speculative as the present data and methodology permit only macroanatomical mappings. In this respect it has to be further noted that the large inter-individual anatomical variability of BA 44 and 45 (Amunts et al., 1999) may have also constrained the present analyses (see also Amunts and Willmes, 2006). In addition, the present focus on dissociable and overlapping contributions of individual brain regions completely neglects the important role of the interconnecting fiber tracts (e.g., Saur et al., 2008; Klein et al., 2016) and the effects of lesions on the network level, which should be more systematically addressed in future studies (see Almairac et al., 2015; Chouiter et al., 2016). Likewise, there is also evidence for a functional role of the right hemisphere in semantic and phonological fluency that was neglected in the present study by its focus on left-hemispheric stroke patients and that warrants future research (cf. Biesbroek et al., 2016).

Finally, the task procedure of the verbal fluency task was partly different from other studies, because in the present study only nouns were allowed for responding to the phonological fluency condition (see also Schmidt et al., 2017, for a related discussion). This procedure was adopted from Katzev et al. (2013), who ran pilot experiments for final item selection based on frequency estimations for German nouns (excluding compound nouns) with given initial letters derived from the Mannheim Corpus of the German language as an indication of the expected difficulty level of individual letter cues (see also Heim et al., 2008, 2009). This constraint may have resulted in increased item difficulty and increased task demands for phonological cues compared to the common assessment of phonological fluency that was, however, compensated for difficulty-adjusted item selection following Katzev et al. (2013). However, the difference in the task procedures has not artificially introduced the difference in task difficulty between both fluency types in the first place, as this constitutes a general phenomenon that can be also observed without restricting responses in the phonological fluency condition to only nouns (see Table 1). Furthermore, given that previous evidence from the lesion, neuroimaging, and cognitive literature already suggested a potential dissociation between

semantic and phonological fluency in the underlying neural and cognitive processes (Baldo et al., 2006, 2010; Benton, 1968; Costafreda et al., 2006; Gourovitch et al., 2000; Henry and Crawford, 2004; Schmidt et al., 2017; Thompson-Schill et al., 1998), it is unlikely that the present findings of distinct and shared contributions of left frontal and left temporal cortex are artificially introduced or biased by constraining the phonological fluency task to nouns only.

5. Conclusion

The present region-based and voxel-based lesion-behavior mapping analyses in a large sample of chronic stroke patients corroborate and extend previous findings and also resolve discrepancies in the extant literature on the neural correlates of verbal fluency by demonstrating distinct as well as shared contributions. More precisely, the demonstration of a significant double dissociation at the group level showed that the integrity of left superior and middle temporal gyri is specifically crucial for semantic fluency, whereas the integrity of pars opercularis of left IFG specifically subserves phonological fluency. Furthermore, a main effect of lesion in pars triangularis of left IFG demonstrated that its integrity is generally critical for word retrieval in both semantic and phonological fluency.

Acknowledgements

The authors report no conflicts of interest. The present research was partly supported by a grant of the BrainLinks-BrainTools Cluster of Excellence (C.W., C.P.K., project #36) funded by the German Research Foundation (DFG; grant # EXC 1086) and by a grant of the Müller-Fahnenberg Foundation (C.P.K., C.W.). C.S.M.S. received scholarship funds from the State Law on Graduate Funding of the University of Freiburg, Germany.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.nicl.2019.101840>.

References

- Abwender, D.A., Swan, J.G., Bowerman, J.T., Connolly, S.W., 2001. Qualitative analysis of verbal fluency output: review and comparison of several scoring methods. *Assessment* 8 (3), 323–338.
- Almairac, F., Herbet, G., Moritz-Gasser, S., de Champfleury, N.M., Duffau, H., 2015. The left inferior fronto-occipital fasciculus subserves language semantics: a multilevel lesion study. *Brain Struct. Funct.* 220 (4), 1983–1995.
- Amunts, K., Willmes, K., 2006. From intersubject variability in clinical syndromes to anatomical variability. *Brain Lang.* 96 (2), 147–150.
- Amunts, K., Schleicher, A., Bürgel, U., Mohlberg, H., Uylings, H., Zilles, K., 1999. Broca's region revisited: cytoarchitecture and intersubject variability. *J. Comp. Neurol.* 412 (2), 319–341.
- Ashburner, J., 2007. A fast diffeomorphic image registration algorithm. *Neuroimage* 38 (1), 95–113.
- Azuma, T., 2004. Working memory and perseveration in verbal fluency. *Neuropsychology* 18 (1), 69.
- Badre, D., Poldrack, R.A., Paré-Blagoev, E.J., Insler, R.Z., Wagner, A.D., 2005. Dissociable controlled retrieval and generalized selection mechanisms in ventrolateral prefrontal cortex. *Neuron* 47 (6), 907–918.
- Baldo, J.V., Shimamura, A.P., 1998. Letter and category fluency in patients with frontal lobe lesions. *Neuropsychology* 12 (2), 259.
- Baldo, J.V., Shimamura, A.P., Delis, D.C., Kramer, J., Kaplan, E., 2001. Verbal and design fluency in patients with frontal lobe lesions. *J. Int. Neuropsychol. Soc.* 7 (5), 586–596.
- Baldo, J.V., Schwartz, S., Wilkins, D., Dronkers, N.F., 2006. Role of frontal versus temporal cortex in verbal fluency as revealed by voxel-based lesion symptom mapping. *J. Int. Neuropsychol. Soc.* 12 (06), 896–900.
- Baldo, J.V., Schwartz, S., Wilkins, D.P., Dronkers, N.F., 2010. Double dissociation of letter and category fluency following left frontal and temporal lobe lesions. *Aphasiology* 24 (12), 1593–1604.
- Bates, E., Wilson, S.M., Saygin, A.P., Dick, F., Sereno, M.I., Knight, R.T., Dronkers, N.F., 2003. Voxel-based lesion-symptom mapping. *Nat. Neurosci.* 6 (5), 448.
- Benton, A.L., 1968. Differential behavioral effects in frontal lobe disease. *Neuropsychologia* 6 (1), 53–60.
- Beume, L.A., Martin, M., Kaller, C.P., Klöppel, S., Schmidt, C.S., Urbach, H., Umarova, R.M., 2017. Visual neglect after left-hemispheric lesions: a voxel-based lesion-symptom mapping study in 121 acute stroke patients. *Exp. Brain Res.* 235 (1), 83–95.
- Biesbroek, J.M., van Zandvoort, M.J., Kappelle, L.J., Velthuis, B.K., Biessels, G.J., Postma, A., 2016. Shared and distinct anatomical correlates of semantic and phonemic fluency revealed by lesion-symptom mapping in patients with ischemic stroke. *Brain Struct. Funct.* 221 (4), 2123–2134.
- Bookheimer, S., 2002. Functional MRI of language: new approaches to understanding the cortical organization of semantic processing. *Annu. Rev. Neurosci.* 25 (1), 151–188.
- Brunner, E., Munzel, U., 2000. The nonparametric behrens-fisher problem: asymptotic theory and a small-sample approximation. *Biom. J.* 42 (1), 17–25.
- Brunner, E., Munzel, U., 2002. *Nichtparametrische Datenanalyse*. vol. 34. Springer, Berlin, Heidelberg, pp. 59.
- Channon, S., Crawford, S., 2000. The effects of anterior lesions on performance on a story comprehension test: left anterior impairment on a theory of mind-type task. *Neuropsychologia* 38 (7), 1006–1017.
- Chouiter, L., Holmberg, J., Manuel, A.L., Colombo, F., Clarke, S., Annoni, J.M., Spierer, L., 2016. Partly segregated cortico-subcortical pathways support phonologic and semantic verbal fluency: a lesion study. *Neuroscience* 329, 275–283.
- Collins, A.M., Loftus, E.F., 1975. A spreading-activation theory of semantic processing. *Psychol. Rev.* 82 (6), 407.
- Costafreda, S.G., Fu, C.H., Lee, L., Everitt, B., Brammer, M.J., David, A.S., 2006. A systematic review and quantitative appraisal of fMRI studies of verbal fluency: role of the left inferior frontal gyrus. *Hum. Brain Mapp.* 27 (10), 799–810.
- Crawford, J.R., Garthwaite, P.H., Gray, C.D., 2003. Wanted: fully operational definitions of dissociations in single-case studies. *Cortex* 39 (2), 357–370.
- Davies, M., 2010. Double dissociation: understanding its role in cognitive neuropsychology. *Mind Lang.* 25 (5), 500–540.
- Demonet, J.F., Price, C., Wise, R., Frackowiak, R.S.J., 1994. A PET study of cognitive strategies in normal subjects during language tasks: influence of phonetic ambiguity and sequence processing on phoneme monitoring. *Brain* 117 (4), 671–682.
- Dressing, A., Nitschke, K., Kümmerer, D., Bormann, T., Beume, L., Schmidt, C.S., Kaller, C.P., 2018. Distinct contributions of dorsal and ventral streams to imitation of tool-use and communicative gestures. *Cereb. Cortex* 28 (2), 474–492.
- Frith, C.D., Friston, K.J., Liddle, P.F., Frackowiak, R.S.J., 1991. A PET study of word finding. *Neuropsychologia* 29 (12), 1137–1148.
- Gershberg, F.B., Shimamura, A.P., 1995. Impaired use of organizational strategies in free recall following frontal lobe damage. *Neuropsychologia* 33 (10), 1305–1333.
- Gourovitch, M.L., Kirkby, B.S., Goldberg, T.E., Weinberger, D.R., Gold, J.M., Esposito, G., Berman, K.F., 2000. A comparison of rCBF patterns during letter and semantic fluency. *Neuropsychology* 14 (3), 353–360.
- Gruenewald, P.J., Lockhead, G.R., 1980. The free recall of category examples. *J. Exp. Psychol. Hum. Learn. Mem.* 6 (3), 225.
- Heim, S., Eickhoff, S.B., Amunts, K., 2008. Specialisation in Broca's region for semantic, phonological, and syntactic fluency? *Neuroimage* 40 (3), 1362–1368.
- Heim, S., Eickhoff, S.B., Amunts, K., 2009. Different roles of cytoarchitectonic BA 44 and BA 45 in phonological and semantic verbal fluency as revealed by dynamic causal modelling. *Neuroimage* 48 (3), 616–624.
- Helmstaedter, C., Gleissner, U., Zentner, J., Elger, C.E., 1998. Neuropsychological consequences of epilepsy surgery in frontal lobe epilepsy. *Neuropsychologia* 36 (7), 681–689.
- Henry, J.D., Crawford, J.R., 2004. A meta-analytic review of verbal fluency performance following focal cortical lesions. *Neuropsychology* 18 (2), 284–295.
- Hirshorn, E.A., Thompson-Schill, S.L., 2006. Role of the left inferior frontal gyrus in covert word retrieval: neural correlates of switching during verbal fluency. *Neuropsychologia* 44 (12), 2547–2557.
- Hoeren, M., Kümmerer, D., Bormann, T., Beume, L., Ludwig, V.M., Vry, M.S., Weiller, C., 2014. Neural bases of imitation and pantomime in acute stroke patients: distinct streams for praxis. *Brain* 137 (10), 2796–2810.
- Indefrey, P., Levelt, W.J., 2000. The neural correlates of language production. In: *The New Cognitive Neurosciences*, 2nd ed. MIT press, pp. 845–865.
- Janowsky, J.S., Shimamura, A.P., Kritchewsky, M., Squire, L.R., 1989. Cognitive impairment following frontal lobe damage and its relevance to human amnesia. *Behav. Neurosci.* 103 (3), 548.
- Jurado, M.A., Mataro, M., Verger, K., Bartumeus, F., Junque, C., 2000. Phonemic and semantic fluencies in traumatic brain injury patients with focal frontal lesions. *Brain Inj.* 14 (9), 789–795.
- Katzev, M., Tuscher, O., Hennig, J., Weiller, C., Kaller, C.P., 2013. Revisiting the functional specialization of left inferior frontal gyrus in phonological and semantic fluency: the crucial role of task demands and individual ability. *J. Neurosci.* 33 (18), 7837–7845.
- Klein, E., Suchan, J., Moeller, K., Karnath, H.-O., Knops, A., Wood, G., Willmes, K., 2016. Considering structural connectivity in the triple code model of numerical cognition: differential connectivity for magnitude processing and arithmetic facts. *Brain Struct. Funct.* 221, 979–995.
- Lancaster, J.L., Woldorff, M.G., Parsons, L.M., Liotti, M., Freitas, C.S., Rainey, L., Fox, P.T., 2000. Automated Talairach atlas labels for functional brain mapping. *Hum. Brain Mapp.* 10 (3), 120–131.
- Lanting, S., Haugrud, N., Crossley, M., 2009. The effect of age and sex on clustering and switching during speeded verbal fluency tasks. *J. Int. Neuropsychol. Soc.* 15 (2), 196–204.
- Lezak, M.D., Howieson, D.B., Bigler, E.D., Tranel, D., 2012. *Neuropsychological Assessment*, 5th ed. Oxford University Press, New York, NY, US.
- Li, M., Zhang, Y., Song, L., Huang, R., Ding, J., Fang, Y., ... Han, Z., 2017. Structural connectivity subserving verbal fluency revealed by lesion-behavior mapping in stroke

- patients. *Neuropsychologia* 101, 85–96.
- Luckhurst, L., Lloyd-Jones, T.J., 2001. A selective deficit for living things after temporal lobectomy for relief of epileptic seizures. *Brain Lang.* 79 (2), 266–296.
- Maldjian, J.A., Laurienti, P.J., Kraft, R.A., Burdette, J.H., 2003. An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *Neuroimage* 19 (3), 1233–1239.
- Martin, R.C., Loring, D.W., Meador, K.J., Lee, G.P., 1990. The effects of lateralized temporal lobe dysfunction on normal and semantic word fluency. *Neuropsychologia* 28 (8), 823–829.
- Martin, M., Beume, L., Kümmerer, D., Schmidt, C.S., Bormann, T., Dressing, A., Kaller, C.P., 2016. Differential roles of ventral and dorsal streams for conceptual and production-related components of tool use in acute stroke patients. *Cereb. Cortex* 26 (9), 3754–3771.
- Mayr, U., 2002. On the dissociation between clustering and switching in verbal fluency: comment on Troyer, Moscovitch, Winocur, Alexander and Stuss. *Neuropsychologia* 40 (5), 562–566.
- Mayr, U., Kliegl, R., 2000. Complex semantic processing in old age: does it stay or does it go? *Psychol. Aging* 15 (1), 29.
- McIntosh, R.D., 2017. Simple Dissociations for a Higher-Powered Neuropsychology. PrePrint from. <https://psyarxiv.com/mnhct/>.
- Meinzer, M., Flaisch, T., Wilser, L., Eulitz, C., Rockstroh, B., Conway, T., ... Crosson, B., 2009. Neural signatures of semantic and phonemic fluency in young and old adults. *J. Cogn. Neurosci.* 21 (10), 2007–2018.
- Miller, E., 1984. Verbal fluency as a function of a measure of verbal intelligence and in relation to different types of cerebral pathology. *Br. J. Clin. Psychol.* 23 (1), 53–57.
- Milner, B., 1964. Some effects of frontal lobectomy in man. In: *The Frontal Granular Cortex and Behavior*, pp. 313–334.
- Moscovitch, M., 1994. Cognitive resources and dual-task interference effects at retrieval in normal people: The role of the frontal lobes and medial temporal cortex. *Neuropsychology* 8 (4), 524.
- Mummery, C.J., Patterson, K., Hodges, J.R., Wise, R.J., 1996. Generating 'tiger' as an animal name or a word beginning with T: differences in brain activation. *Proc. R. Soc. B* 263 (1373), 989–995.
- Pendleton, M.G., Heaton, R.K., Lehman, R.A., Hulihan, D., 1982. Diagnostic utility of the Thurstone word fluency test in neuropsychological evaluations. *J. Clin. Exp. Neuropsychol.* 4 (4), 307–317.
- Perret, E., 1974. The left frontal lobe of man and the suppression of habitual responses in verbal categorical behaviour. *Neuropsychologia* 12 (3), 323–330.
- Rende, B., Ramsberger, G., Miyake, A., 2002. Commonalities and differences in the working memory components underlying letter and category fluency tasks: a dual-task investigation. *Neuropsychology* 16 (3), 309–321.
- Reverberi, C., Laiacona, M., Capitani, E., 2006. Qualitative features of semantic fluency performance in mesial and lateral frontal patients. *Neuropsychologia* 44 (3), 469–478.
- Robinson, G., Blair, J., Cipolotti, L., 1998. Dynamic aphasia: an inability to select between competing verbal responses. *Brain J. Neurol.* 121 (1), 77–89.
- Robinson, G., Shallice, T., Bozzali, M., Cipolotti, L., 2012. The differing roles of the frontal cortex in fluency tests. *Brain* 135 (7), 2202–2214.
- Rogers, R.D., Sahakian, B.J., Hodges, J.R., Polkey, C.E., Kennard, C., Robbins, T.W., 1998. Dissociating executive mechanisms of task control following frontal lobe damage and Parkinson's disease. *Brain J. Neurol.* 121 (5), 815–842.
- Rorden, C., Karnath, H.O., Bonilha, L., 2007. Improving lesion-symptom mapping. *J. Cogn. Neurosci.* 19 (7), 1081–1088.
- Saur, D., Kreher, B.W., Schnell, S., Kümmerer, D., Kellmeyer, P., Vry, M.S., Weiller, C., 2008. Ventral and dorsal pathways for language. *Proc. Natl. Acad. Sci. U. S. A.* 105 (46), 18035–18040.
- Schmidt, C.S., Schumacher, L.V., Römer, P., Leonhart, R., Beume, L., Martin, M., Kaller, C.P., 2017. Are semantic and phonological fluency based on the same or distinct sets of cognitive processes? Insights from factor analyses in healthy adults and stroke patients. *Neuropsychologia* 99, 148–155.
- Schwartz, S., Baldo, J., Graves, R.E., Brugger, P., 2003. Pervasive influence of semantics in letter and category fluency: a multidimensional approach. *Brain Lang.* 87 (3), 400–411.
- Shallice, T., 1988. *From Neuropsychology to Mental Structure*. Cambridge University Press.
- Shao, Z., Janse, E., Visser, K., Meyer, A.S., 2014. What do verbal fluency tasks measure? Predictors of verbal fluency performance in older adults. *Front. Psychol.* 5, 772.
- Strauss, E., Sherman, E.M.S., Spreen, O., 2006. *A Compendium of Neuropsychological Tests*. Oxford University Press, New York.
- Stuss, D.T., Alexander, M.P., Palumbo, C.L., Buckle, L., Sayer, L., Pogue, J., 1994. Organizational strategies with unilateral or bilateral frontal lobe injury in word learning tasks. *Neuropsychology* 8 (3), 355.
- Stuss, D.T., Craik, F.I., Sayer, L., Franchi, D., Alexander, M.P., 1996. Comparison of older people and patients with frontal lesions: evidence from word list learning. *Psychol. Aging* 11 (3), 387.
- Stuss, D.T., Alexander, M.P., Hamer, L., Palumbo, C., Dempster, R., Binns, M., ... Izukawa, D., 1998. The effects of focal anterior and posterior brain lesions on verbal fluency. *J. Int. Neuropsychol. Soc.* 4 (3), 265–278.
- Stuss, D.T., Toth, J.P., Franchi, D., Alexander, M.P., Tipper, S., Craik, F.I., 1999. Dissociation of attentional processes in patients with focal frontal and posterior lesions. *Neuropsychologia* 37 (9), 1005–1027.
- Sylvester, C.Y.C., Shimamura, A.P., 2002. Evidence for intact semantic representations in patients with frontal lobe lesions. *Neuropsychology* 16 (2), 197.
- Szatkowska, I., Grabowska, A., Szymanska, O., 2000. Phonological and semantic fluencies are mediated by different regions of the prefrontal cortex. *Acta Neurobiol. Exp. (Wars)* 60 (4), 503–508.
- Teuber, H.L., 1955. Physiological psychology. *Annu. Rev. Psychol.* 6 (1), 267–296.
- Thompson-Schill, S.L., D'Esposito, M., Aguirre, G.K., Farah, M.J., 1997. Role of left inferior prefrontal cortex in retrieval of semantic knowledge: a reevaluation. *Proc. Natl. Acad. Sci.* 94 (26), 14792–14797.
- Thompson-Schill, S.L., Swick, D., Farah, M.J., D'Esposito, M., Kan, I.P., Knight, R.T., 1998. Verb generation in patients with focal frontal lesions: a neuropsychological test of neuroimaging findings. *Proc. Natl. Acad. Sci.* 95 (26), 15855–15860.
- Tombaugh, T.N., Kozak, J., Rees, L., 1999. Normative data stratified by age and education for two measures of verbal fluency: FAS and animal naming. *Arch. Clin. Neuropsychol.* 14 (2), 167–177.
- Troyer, A.K., Moscovitch, M., Winocur, G., 1997. Clustering and switching as two components of verbal fluency: evidence from younger and older healthy adults. *Neuropsychology* 11 (1), 138.
- Troyer, A.K., Moscovitch, M., Winocur, G., Alexander, M.P., Stuss, D., 1998. Clustering and switching on verbal fluency: the effects of focal frontal- and temporal-lobe lesions. *Neuropsychologia* 36 (6), 499–504.
- Tucha, O., Smely, C., Lange, K.W., 1999. Verbal and figural fluency in patients with mass lesions of the left or right frontal lobes. *J. Clin. Exp. Neuropsychol.* 21 (2), 229–236.
- Wagner, A.D., Paré-Blagoev, E.J., Clark, J., Poldrack, R.A., 2001. Recovering meaning: left prefrontal cortex guides controlled semantic retrieval. *Neuron* 31 (2), 329–338.
- Wagner, S., Sebastian, A., Lieb, K., Tüscher, O., Tadić, A., 2014. A coordinate-based ALE functional MRI meta-analysis of brain activation during verbal fluency tasks in healthy control subjects. *BMC Neurosci.* 15 (1), 19.