# The functional relevance of dorsal motor systems for processing tool nouns–evidence from patients with focal lesions

Felix R. Dreyer [a,b,*], Thomas Picht [b,c,d], Dietmar Frey [c], Peter Vajkoczy [c], Friedemann Pulvermüller [a,b,d,e]

[a] Freie Universität Berlin, Brain Language Laboratory, Department of Philosophy and Humanities, Berlin, Germany
[b] Humboldt Universität zu Berlin, Cluster of Excellence Matters of Activity, Image Space Material, Berlin, Germany
[c] Charite University Medicine Berlin, Department of Neurosurgery, Berlin, Germany
[d] Einstein Center for Neurosciences Berlin, Berlin, Germany
[e] Humboldt Universität zu Berlin, Berlin School of Mind and Brain, Berlin, Germany

## ARTICLE INFO

## ABSTRACT

This study asks whether lesions in different parts of the brain have different effects on the processing of words typically used to refer to objects with and without action affordances, for example tools and animal-related nouns. A cohort of neurological patients with focal lesions participated in a lexical decision paradigm where nouns semantically related to tools, foods and animals were presented along with matched pseudo-words. Differences in semantic features between the categories were confirmed using extensive semantic ratings whereas all semantic word categories were matched for relevant psycholinguistic variables. In a data-driven region of interest analysis, lesions in dorsal pre- and postcentral grey and white matter areas were associated to specific performance deficits for tool nouns when compared to animal nouns. In contrast, patients with lesions primarily affecting perisylvian inferior-frontal and/or temporal regions presented similar deficits across all semantic word categories tested and likewise a group of age and education matched healthy control participants did not show any category specific differences. These findings falsify brain language models denying the fronto-parietal cortex' role in word recognition and semantic understanding. They are best accounted for by frameworks that acknowledge a role of sensorimotor cortex in the semantic processing of action-related words.

## 1. Introduction

Evidence from neuroimaging, gathered in the course of the past few decades, points to an involvement of modality specific sensory and motor areas in the processing of concepts and word meaning. The 'lighting up' of sensory and motor cortices in the processing of meaningful linguistic symbols was not only observed arbitrarily, but in a systematic fashion, following the predictions of neurocognitive semantic theories postulating that a word's meaning is grounded in perceptual (or 'sensory') and action-related (or 'motor') experience, reflecting its meaning and use (Barsalou, 1999, 2008; Pulvermüller, 1999, 2005; Pulvermüller and Fadiga, 2010; Glenberg and Gallese, 2012; Kiefer and Pulvermüller, 2012; Strijkers and Costa, 2016). For example, words related to face-, hand- and leg-related actions were shown to recruit motor regions normally involved in actual effector-specific action

execution (Hauk et al., 2004; Hauk and Pulvermüller, 2004; Martin et al., 1996; Pulvermüller et al., 2005a, 2005b; Carota et al., 2012; Shtyrov et al., 2014) and words with pronounced semantic relationships to odours (González et al., 2006), tastes (Barrós-Loscertales et al., 2012) or sounds (Kiefer et al., 2008) were related to activation patterns specifically reaching into brain areas involved in corresponding perception processes. Furthermore, an involvement of motor systems in semantic processing could even be demonstrated for the domain of abstract words, as passive reading of abstract emotion words like "fear" activated not only limbic areas, involved in processing affective information, but also hand- and face-related motor areas (Moseley et al., 2012) and abstract words referring to mental states and processes like "thought" were shown to elicit specific activation in face motor areas (Dreyer and Pulvermüller, 2018; for a different position see Ghio et al., 2016). These results suggest that sensory and motor systems of the human brain take

---

* Corresponding author. Felix Dreyer, Ph.D Brain and Language Laboratory Department of Philosophy and Humanities, WE4 Freie Universität Berlin, 19145, Berlin, Germany.
*E-mail address:* felix.dreyer@fu-berlin.de (F.R. Dreyer).

part in, and contribute to semantic processing. They do not encourage a "strong embodiment" of semantics exclusively in sensory and motor modules or cortices, but they suggest that the distributed circuits carrying meaningful symbols also draw upon, and 'reach into', those areas of cortex that are otherwise engaged in perception and action per se (Pulvermüller, 1999, 2005; 2018). The semantic representation of an object word would therefore include neurons encoding the visual shape and color of the object the word is typically used to speak about and, if this object happens to be a tool, also neurons storing the action-related knowledge that the reference object is normally used in such and such a way.

The evidence about action and perception system activations reflecting aspects of the meaning of symbols remains unexplained by classic cognitive approaches viewing the seat of concepts and semantics in central modules divorced from encapsulated input modules for perception and output modules for motor function (see, for example, Fodor, 1983; Anderson, 1983; Ellis and Young, 1988). According to this strong 'amodal' semantics view, the semantic representation of an object word and the perceptual 3D representation of that object in the mind or brain are distinct from each other; the latter would not belong to the former (cf., for example, Caramazza et al., 1990). Meanwhile, this position is difficult to maintain, given the common neuroimaging results that, for example, color or form-related words 'light up' different parts of the visual perceptual brain system (Pulvermüller and Hauk, 2005). Therefore, a weak form of an amodal semantic model was proposed according to which amodal semantic and perceptual or action-related representations are still functionally independent in principle, but can be linked occasionally and flexibly, for example if there is need for relating symbols to the world (see, for example, Mahon and Caramazza, 2008; Mahon, 2015). In this perspective, any activation of sensory or motor areas in processing meaningful symbols should only emerge when required by the task. The optional and flexible additional contribution may 'enrich' or 'color' the concept, but never, according to this position, would a lesion in sensory or motor systems be predicted to impair the processing of symbols related to objects or action-affordances; never, unless one postulates that the semantic system itself is seated in motor or visual cortex.

The discussion between grounded models with interwoven action, perception and symbol systems and the weak amodal semantic accounts has led to critical and important debates. For example, it has been pointed out that the observed activations of sensory and motor cortices during symbol processing could be a result of optional second order conscious or unconscious imagery processes, which may happen spontaneously or as a result of the task, but in the aftermath of the actual semantic or conceptual comprehension process to which perceived symbols give rise. Although this interpretation may seem unlikely, given that some of the aforementioned approaches were able to show somatotopic dissociations in the motor system already very early in processing, within 200 ms after critical phonological information was perceived (Pulvermüller et al., 2005b; Shtyrov et al., 2014; Grisoni et al., 2016, 2019), it has been argued that evidence on the time domain is not useful with regards to the question of causality (Pulvermüller, 2005; Mahon and Caramazza, 2008). Hence, to address the causality question, one has to resort to neuropsychological research and study functional changes in the neural substrate (as the independent variable) and their resultant behavioural change (as the dependent one). Both neurostimulation and investigations of clinical populations allow one to draw direct conclusions on the causality and functional relevance of stimulated or lesioned areas for cognitive function.

TMS work by Pulvermüller and colleagues (2005a) and Willems and coworkers (2011) revealed effects of motor area stimulation on response times in a lexical decision task (LDT) for specific types of action-related verbs, whereas matched control verbs remained unaffected (see also Vukovic et al., 2017 for similar results in a semantic judgment task). These observations are indeed in line with a causal role of these motor areas in the processing of action-related semantics. However, this kind

of evidence is questioned by some authors (Mahon and Caramazza, 2008) and interpreted in terms of possible spreading of neuro-stimulation induced activity from modality specific motor systems to the neural substrate of a potentially underlying amodal concept. Still, it has been argued that this spreading activation account does not fully account for why focal stimulation of specific motor areas specifically influences the processing of particular semantic kinds of words (Pulvermüller, 2013).

The gold standard for demonstrating a causal role of cortical areas for cognitive processing are neuropsychological patient studies. If patients with lesions in an area of interest X show a significant deficit in processing a given semantic word type, whereas control words with different semantics are affected less, and if such a dissociation is absent in control subjects, and even in patients with lesions elsewhere, then there would be support for the causal nature of area X in semantic processing. In case that this could be shown for motor areas and for the processing of words typically used to refer to action-affording objects, it would be difficult to argue that the result is partly explained by activation spreading from motor areas to an alleged amodal meaning representation, unless this activation spreading were of functional relevance and constitutive for said amodal meaning representation. As this scenario would resemble a single dissociation, a standard inference scheme in neuropsychology (Crawford et al., 2003) would be that, if the lesion in this brain area leads to stronger impairments of one semantic category as compared with a second one, then this region is crucial, causal, or even exerts a necessary role in processing the more impaired items.

Previous research has shown that lesions involving motor areas led to specific deficits in processing words with action-related semantics. This could be demonstrated in stroke patients with inferior-frontal or fronto-parietal lesions (Damasio and Tranel, 1993; Neininger and Pulvermüller, 2001, 2003), and in patients suffering from motor neuron disease (Bak et al., 2001; Bak and Hodges, 2004) and Parkinson's disease (Cotelli et al., 2007; Boulenger et al., 2008, Fernandino et al., 2013). A similar finding came from research on individuals with Autism Spectrum Disorder, which is characterized by an anomaly of the long-distance cortical connections, in particular the arcuate fascicle connecting motor systems with posterior cortices (Moseley et al., 2013). These reports however, suffer from a methodological noun-verb confound, as the performance of action-related verbs was compared to that of object-related nouns. Comparisons between object nouns and action verbs render it difficult to disentangle effects of grammatical and semantic word type, i.e. in theory, there is still a possibility that motor areas are functionally involved in the processing of all kinds of verbs, irrespective of semantics, and their associated grammatical information is crucial. Kemmerer et al. (2012) used a more systematic approach, applying voxel-based lesion symptom mapping (VLSM) to investigate neural processing in large cohorts of neurological patients with different etiologies. Here, a functional role of precentral motor areas in the processing of actions and action verbs was demonstrated on a wide range of tasks (including word-picture matching, as well as word comparison and word attribute tasks) was shown, along with additional evidence that inferior-frontal and temporal regions are relevant, too. This approach however, was lacking a proper (matched) control condition, again rendering it difficult to ascribe a functional role of the observed areas specifically to the processing of action-related semantics. To avoid this issue, Arevalo and coworkers (2012) analyzed word-picture matching performance of nouns and verbs of both action-related and action unrelated semantics using VLSM in a cohort of stroke patients. Significant voxels for processing action-related semantics were found not only in posterior and superior temporal and inferior-frontal cortex, but also in premotor areas. An effector-specific semantic somatotopy in motor areas contribution could, however, not be observed. One possible reason for this absence of fine-grained topographical mappings might be seen in the typically large lesions of stroke patients (though see Neininger and Pulvermüller, 2001, 2003 for notable exceptions) which may have led to

contamination across lesion profiles across adjacent effector-specific motor areas.

Considering the aforementioned issues, Dreyer et al. (2015) compared lexical decision performance of action and non-action-related categories within a grammatical word class in a pair of two patients showing lesions of high focality in their motor systems. To this end, hand-action-related tool nouns, face-action-related food nouns, abstract emotion nouns and animal nouns, the latter to provide an action-unrelated baseline, were applied in a LDT. Notably, results from one patient with a focal lesion in white matter directly adjacent to hand motor areas showed a specific processing deficit for tool words relative to action-related and –unrelated control words. This result directly demonstrated the functional necessity of motor areas for processing of hand-action-related semantics, independent of confounds of grammatical class and with high spatial specificity. In the context of the debate on the role of sensorimotor systems in semantic processing, one can use these results to argue that in at least one case, a focal lesion impacting on the connections of dorsolateral motor areas led to a processing deficit for words with hand-action-affording object-related meaning. One may of course question these results because a single case could always be exceptional and not representative of language processing in the brain as it is present in the majority of the population. Therefore, we set out to validate observations of Dreyer et al. (2015) on the functional relevance of motor areas for processing words of concrete semantics by investigating a larger group of patients. We asked whether the selected populations showed an impairment in processing nouns from specific semantic categories, relative to matched nouns from other semantic types. We focused on tool names, which relate to objects affording hand-actions, food words, which relate to objects affording mouth-actions and animal names, which lack strong action-relatedness.

Based on the neurobiological model postulating words and their meaning to be represented by distributed cell assemblies spread across perisylvian inferior-frontal and temporal areas (Pulvermüller, 1999, 2005; Pulvermüller and Fadiga, 2010), extrasylvian multimodal sites and sensorimotor cortices, where their grounding referential semantic information is stored, we expect specific between-category differences in processing impairments and increased between-category performance heterogeneity to be present in patients with extrasylvian sensorimotor lesion but not in patients suffering from perisylvian lesions. Furthermore, we predict category specific impairments to follow a pattern of semantic somatotopy. Specifically, tool nouns referring to objects affording hand-actions are expected to be specifically impaired in comparison with the non-action animal baseline following focal lesions in the dorsal central cortices, especially in those necessary for moving the hands, i.e. dorsal hand motor areas in Brodmann Areas (BAs) 4 and 6, as well as anterior parietal areas (BA 40 and 7), previously reported to be relevant for the manipulation of tools (Moll et al., 2000; Choi et al., 2001; Johnson-Frey, 2004; Ohgami et al., 2004). At the same time, food nouns with a strong relation to face-actions are hypothesized to be selectively impaired following lesions in ventral face-related motor areas in BAs 4 and 6. In contrast, patients with lesions affecting predominantly the perisylvian inferior-frontal and temporal, but not (or much less so) the extrasylvian sensorimotor systems should not present any processing advantage of animal nouns over nouns with action-related semantics, but general category-unspecific processing deficits across categories instead. Likewise, healthy controls should also not exhibit any similar patterns of category specific performance impairments.

These results provide support for the position that sensorimotor areas and more generally modality-preferential ones are essential for the processing of specific semantic categories. They would allow us to rule out the possibility that these areas only serve an optional role in semantic processing. As we use a lexical decision task, we do not apply a task that could be said to make subjects process particular aspects of meaning, so as to emphasize a latent but unessential and entirely flexible link between word form and referential knowledge. So any category

differences after extrasylvian lesions would also argue against weak amodal semantic models. Furthermore, any such results would be impossible to explain in terms of second order processes, such as imagery consecutive to symbol understanding.

## 2. Methods

### 2.1. Participants

#### 2.1.1. Patient profiles and selection criteria

In total, forty-one patients participated in the LDT paradigm and, if time allowed, also in a subsequent shortened version of the AAT. Tumor patients were selected for this analysis due to their lesions being of smaller extent (56 cm$^3$ on average) and not restricted by vascular properties of the brain, as is the case for stroke patients investigated in earlier approaches, thus allowing for more fine grained examinations of the neural substrates of semantic processing. Four of those patients had to abort the LDT, before it was completed, due to medical examinations scheduled on short notice and were hence excluded from the current analysis and another patient was excluded due to pronounced left handedness (Oldfield laterality quotient of −69). Furthermore, 10 patients were excluded due to diagnoses of high grade tumors (Glioma of World Health Organization Grade IV), as these tumors tend to infiltrate surrounding tissue in a diffuse fashion, rendering it difficult to effectively map and interpret lesions in terms of lesion site specific functional involvement in cognitive processes. Another two patients were excluded from analysis due to tumors growing on the dorsal meninges, rather than infiltrating brain tissue itself. Lesion sites of the remaining 24 patients were mapped and were subjected to further analysis described below (see Table 1 for clinical and social-demographic characteristics). Lesion profiles were most focused in both, perisylvian (inferior-frontal, superior and middle temporal) and extrasylvian (inferior parietal, pre- and postcentral, inferior and middle frontal) grey and white matter with a maximum lesion overlay of 9 patients in voxels in left dorsal premotor cortex (BA 6) and left insula (see Fig. 2), all in the left hemisphere.

#### 2.1.2. Healthy controls

A group of 9 healthy participants (3 males) without neurological records served as control sample for the LDT paradigm. On average, controls were 47.1 years (*S.E.* = 4.7 years) old at the time of testing and received 16.8 years (*S.E.* = 1.1 years) on average of formal education. Age and years of formal education were matched between the healthy controls and the patient sample considered for voxel wise analysis (all *p* > .83).

### 2.2. Paradigm

The experimental paradigm was identical to that applied in a previous dual-case study (Dreyer et al., 2015). In essence, all participants were asked to first participate in a speeded lexical decision task, as described below. If time allowed, which was unfortunately not the case for all patients, subtests from the Aachen Aphasia Test (AAT), a standardized German aphasia test battery (Huber, 1983), were conducted and further socio-demographic data was collected using questionnaires. All procedures were approved by the ethics committee of the Charité University Hospital, Berlin, Germany.

### 2.3. LDT stimuli

One hundred twenty concrete target nouns were presented, along with 40 abstract nouns and 160 action and abstract verbs, which were of not of interest for the current analyses, in addition 160 matched pseudo-nouns, as well as 160 matched pseudo-verbs. To avoid an interference of the concreteness effect (James, 1975) on the comparison of performance heterogeneity between patients and controls (see Analysis), abstract nouns were not considered in the current analysis. Likewise verbs were

**Table 1**

Clinical and social-demographic characteristics of patients included in analysis.

| Patient | Patient Sample | Age | Years of Education | Lesion Type | Lesion Size in cm³ |
|---|---|---|---|---|---|
| 9 | Dorsal Central ROI | 40 | 16 | Astrocytoma. WHO Grace II | 119 |
| 13 | Dorsal Central ROI | 41 | 13 | Olig odendroglioma WHO Grade III | 49 |
| 15 | Dorsal Central ROI | 40 | n.a. | Astrocytoma. WHO Grade III | 19 |
| 17 | Dorsal Central ROI | 52 | 10 | Lung Cancer Metastasis | 66 |
| 20 | Dorsal Central ROI | 44 | n.a. | Astrocytoma. WHO Grace II | 31 |
| 28 | Dorsal Central ROI | 41 | 18 | ADEM | 2 |
| 29 | Dorsal Central ROI | 59 | 14 | Non-Hodgin Lymphoma | 38 |
| 32 | Dorsal Central ROI | 27 | 12 | Astrocytoma | 32 |
| 34 | Dorsal Central ROI | 28 | 23 | Astrocytoma | 138 |
| 2 | Perisylvian/ Temporal | 62 | 24 | Olig odendroglioma WHO Grade III | 187 |
| 4 | Perisylvian/ Temporal | 64 | 12 | Astrocytoma. WHO Grade III | 74 |
| 7 | Perisylvian/ Temporal | 35 | 16 | Astrocytoma. WHO Grace II | 8 |
| 22 | Perisylvian/ Temporal | 48 | 18 | Cavemoma | 4 |
| 24 | Perisylvian/ Temporal | 66 | 13 | Meningioma. WHO Grade II | 181 |
| 31 | Perisylvian/ Temporal | 69 | 18 | Glioma. WHO Grade I-II | 11 |
| 35 | Perisylvian/ Temporal | 58 | 13 | Astrocytoma | 77 |
| 8 | Control Patients | 46 | 15 | Cavemoma | 58 |
| 10 | Control Patients | 33 | 18 | Astrocytoma. WHO Grade III | 65 |
| 25 | Control Patients | 41 | n.a. | Astrocytoma. WHO Grade III | 23 |
| 27 | Control Patients | 52 | 22 | Astrocytoma. WHO Grade III | 38 |
| 36 | Control Patients | 56 | n.a. | Cavemoma | 5 |
| 37 | Control Patients | 31 | 22 | Glioma | 72 |
| 39 | Control Patients | 55 | 17 | Glioma | 34 |
| 40 | Control Patients | 35 | 15 | Glioma | 6 |

not subject to analysis, as they lacked a non-action reference category. Target nouns included 40 stimuli each from semantic categories of animals, foods, tools, which matched for a range of lexical and sublexical psycholinguistic variables, as determined by the dlex corpus (Heister et al., 2011). Matching was achieved for word length, number of syllables, phonological stress, normalized lemma frequency, character bigram frequency, character trigram frequency, initial character-, initial character bigram- and initial character trigram frequency as well as for number of orthographic neighbours in terms of Coltheart's and Levenshtein's N. F/t tests did not reveal differences between semantic category groups for any of these psycholinguistic variables (all p > .05, see Table 2 for details). Four pseudo-randomized stimulus lists were created for the LDT and counterbalanced between patients in order to avoid a systematic influence of stimulus context effects on task performance.

Pseudo-words were generated based on proper word stimuli using the "Wuggy" software (Keuleers and Brysbaert, 2010). To best mimic appearance of proper words, pseudo-nouns all started with a capital letter and were pronounceable, though not homophonous to proper words. In addition, pseudo-nouns were matched to proper word stimuli in their sublexical psycholinguistic properties of average word length, number of syllables, character bigram frequency, character trigram frequency, initial character frequency and initial bigram frequency, as determined by the dlex corpus (all p > .05, see Table 3 for details).

Semantic properties of word stimuli were empirically controlled by semantic ratings, collected from 20 healthy participants (monolingual native speakers of German aged 18–28) before the main experiment. Similar to previous studies (Pulvermüller et al., 2001; Hauk and Pulvermüller, 2004; Dreyer et al., 2015), semantic ratings were expressed on a Likert scales ranging from 1 (no relation) to 7 (strong relation). Each word was rated for its semantic relatedness to hand/arm-, face/-mouth-, leg/foot actions, to visual, olfactory, gustatory, and haptic/-tactile perceptions, as well as to emotions and mental processes. Ratings

**Table 3**

**Matching of semantic noun classes and pseudo-words.** P-values denote results of independent sample t-tests between both stimulus types.

| | Proper Nouns | | Pseudo Nouns | | P |
|---|---|---|---|---|---|
| | M | SD | M | SD | |
| Character-Big ram frequency p. Mio. | 230.46 | 123.52 | 227.02 | 118.52 | .79 |
| Character-Trig ram frequency p. Mio. | 129.67 | 76.34 | 129.94 | 77.89 | .98 |
| Initial Character frequency p. Mio | 13.89 | 6301 | 14.6 | 6193 | .31 |
| Initial Character-Bigram frequency p. Mio | 2312 | 2038 | 2296 | 2280 | .94 |
| Length | 5.73 | 1.47 | 5.63 | 1.27 | .49 |

**Table 2**

**Matching of word stimulus groups (animal, food, tool nouns) on psycholinguistic properties.** P-values denote results from one-way ANOVAs on the effect of semantic category.

| Variables | Semantics | | | | | | P |
|---|---|---|---|---|---|---|---|
| | Animals | | Foods | | Tools | | |
| | M | SD | M | SD | M | SD | |
| Lemma Frequency p. Mio. | 7.26 | 5.47 | 5.95 | 7.74 | 6.86 | 5.97 | .65 |
| Length | 5.5 | 1.66 | 5.78 | 1.37 | 5.93 | 1.47 | .44 |
| Number of Syllables | 1.7 | 0.46 | 1.78 | 0.42 | 1.88 | 0.33 | .17 |
| Character Bigram Frequency p. Mio. | 243.3 | 123.50 | 210.83 | 120.28 | 250.94 | 145.23 | .35 |
| Character Trigram Frequency p. Mio. | 148.48 | 68.33 | 124.3 | 78.41 | 125.52 | 88.83 | .31 |
| Initial Character Frequency p. Mio. | 13.974 | 5816 | 14.427 | 6163 | 14.992 | 7248 | .77 |
| Initial Bigram Frequency p. Mio. | 2349 | 1901 | 1956 | 2000 | 2599 | 2321 | .38 |
| Initial Trigram Frequency p. Mio. | 748 | 1703 | 473 | 1262 | 913 | 1882 | .48 |
| Coltheart Neighbours Frequency p. Mio. | 82 | 269 | 28 | 74 | 56 | 117 | .39 |
| Coltheart's N | 7 | 6.65 | 6.01 | 5.77 | 7.16 | 5.7 | .66 |
| Levenshtein Neighbours Frequency p. Mio. | 165.29 | 547.99 | 147.26 | 594.08 | 61.24 | 118.91 | .58 |
| Levenshtein N | 9.90 | 8.26 | 8.79 | 7.32 | 10.36 | 6.72 | .63 |

of concreteness and word familiarity were also obtained. The concreteness scale was designed with the extremes of high abstractness (1) to high concreteness (7). For inclusion into an effector-specific action word category (tool/food nouns), words had to achieve an average rating above the neutral mid-point of 4 for the related question while being rated lower on all other action semantic scales. Average ratings for all 10 semantic dimensions are shown in Fig. 1.

### 2.4. LDT procedures

Patients and healthy controls were positioned 70 cm in front of a computer screen and were instructed to decide whether or not a word flashing on screen is a meaningful German word, or a meaningless pseudo-word instead. Responses were given via left hand mouse clicks, to assure that responses were not affected by possible motor impairments caused by the left hemispheric lesions in the patient sample. Each trial started with a presentation of a central fixation cross. Its presentation time was pseudo-randomly varied between 2250 and 2750 ms (2500 ms on average) and it was followed by an acoustic 'beep' signal of 200 ms length. 800 ms after the offset of this acoustic signal, the fixation cross disappeared and a letter string, word or pseudoword, was presented tachistoscopically in the center of the screen for 130 ms. After word offset, the screen remained blank until a response was given, or for a maximum of 3000 ms after which the central fixation cross reappeared. All letter strings were printed in black letters on a light grey background, using monospaced Courier New font with a font size of 13.5 and were spanning a maximum of 2° horizontal and 0.6° vertical visual degree. Each test session began with a practice session, consisting of a series of 10 practice trials for the LDT, which applied stimuli that were not used in the actual experiment. Those trials were repeated until a task accuracy of 80% was achieved and until participants felt comfortable with the task, to assure that participants were sufficiently familiarized with task procedures.

The LDT was split up into 8 blocks, each consisting of 80 trials presenting 40 proper words (among those 5 tools, 5 foods and 5 animals) and 40 pseudowords. Furthermore, 2 words were presented as additional filler items at the beginning of each block, which were excluded from analysis. Each block lasted between 6 and 8 min, depending on participants' response speed. Between experimental blocks, participants were offered breaks for lengths of their own choosing.

### 2.5. Aphasia testing and further questionnaires

Following the LDT testing, patients conducted AAT subtests in the following order: Token Test, Verbal Repetition, Naming and Comprehension. For the sake of shorter testing, subjects who performed <7 age-corrected error points on the Token Test (no aphasia diagnosis) were only given the most difficult part of the other subtests and if their performance was flawless, the rest of the subtest was omitted. On average, this shortened version of the aphasia test battery could be conducted within 20 min. Each test session was concluded by the Edinburgh Handedness Inventory (Oldfield, 1971) and a basic demographics questionnaire.

## 3. Analysis

### 3.1. Lesion mapping procedures

Lesion templates were created manually for each individual patient on either T1 or T2 weighted structural MRI images in MRIcron (www.mricro.com/mricron), depending on availability and quality of images. Lesion mapping was conducted before the evaluation of LDT results by FRD. TP and DF, who were also blind to the LTD results, reviewed and corrected resulting lesion masks. Lesion maps included both, actual tumor tissue, if visible, and surrounding hemorrhages or edemas. Resulting lesion maps were used for lesion cost function masking (Brett et al., 2001), to normalize patients' structural images and corresponding lesion templates to a Montreal Neurological Institute (MNI) standard space using the Clinical toolbox (Rorden et al., 2012) of the Statistical Parametric Mapping software (SPM8, Wellcome Department of Cognitive Neurology).

### 3.2. LDT analysis

To allow response bias corrected comparisons between patients, task accuracies for individual semantic noun categories were converted into d-prime scores. D-prime values for each category were derived by considering the category specific hit rate and the overall false positive rate of the pseudo-nouns category (see also Pulvermüller et al., 2010). Furthermore, the root mean square (RMSQ) of all d-prime score differences between concrete noun categories (i.e. animals vs tool, animals vs. food, tool vs. food) was calculated as an index of performance heterogeneity across semantic categories and of category specificity of performance deficits. In order to check an influence speed accuracy
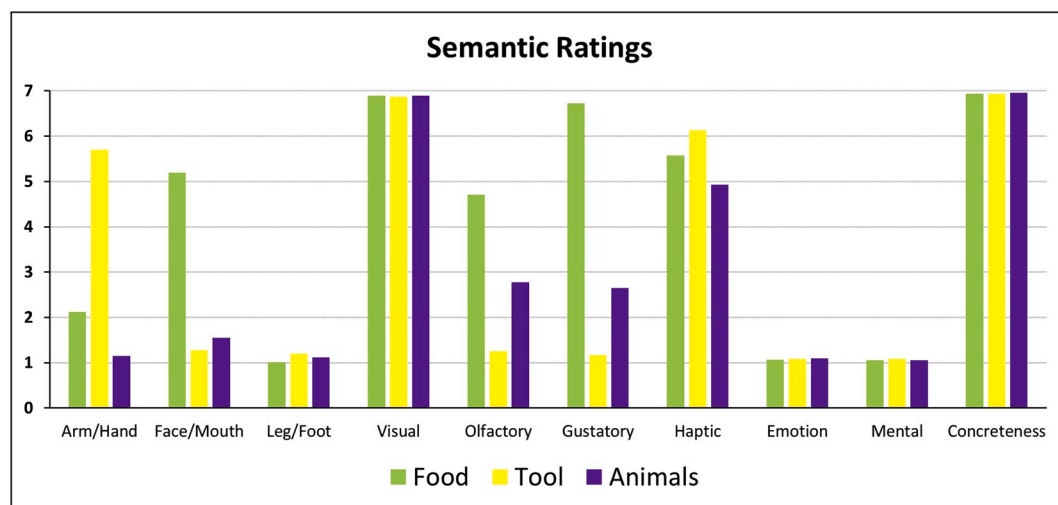


**Fig. 1.** **Average semantic ratings for noun categories, given on a scale from 1 (no semantic relation) to 7 (very strong semantic relation).** Green, yellow and purple bars respectively represent, food, tool and animal nouns. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

tradeoffs on response accuracies, average response times and accuracy were correlated across patients.

### 3.3. AAT analysis

Raw AAT performance scores for each subtest were determined, converted into age normalized standard T-scores and compared to control samples, according to the tests' instructions. In case a patient could not be tested on all items in a subtest, performance was interpolated based on the percentage of scores already achieved in this subtest.

### 3.4. Voxel-wise analysis and ROI definition

In a first step of analysis RMSQ scores between all three semantic noun categories were compared in a voxel-wise fashion to results from a group of healthy controls of all patients to determine data driven regions of interest (ROIs) for further analysis. For this analysis all voxels were considered that were lesioned in at least 5 patients. A lesion overlay of voxels included in analysis is shown in Fig. 2. As the inclusion of healthy controls in such an analysis is prone to provide results of only limited specificity in case too many control participants are included (Rorden et al., 2009), the size of the current control group was chosen to match the maximum lesion overlap in the current sample (N = 9) while at the same time matching the patients in terms of age and years of education. As the current patient sample size and lesion distribution does not allow one to assume data prerequisites for parametric testing for each individual voxel (Rorden et al., 2007), non-parametric two-sample permutation tests (Edgington, 1964; Nichols and Holmes, 2002) were chosen for voxel-wise comparisons to healthy controls. Rather than assuming specific properties of the analyzed data, the null-distribution of in this test is derived directly from the data investigated. To do so, data were randomly reallocated between groups (patients with a lesion in a given voxel and healthy controls) and the average difference of this random permutation was determined in each iteration, which was repeated 1000 times. Resulting p-scores thereby represent the proportion of randomly permutated samples that show a difference of averages between patients and controls that is at least as extreme as the one observed in the actual data. In light of previous results on category specific performance

deficits in clinical populations with extrasylvian lesions (e.g. Bak et al., 2001; Neininger and Pulvermüller, 2001, 2003, Mårtensson et al., 2014; Dreyer et al., 2015), we assumed brain lesions to be associated with more heterogeneous category performance (i.e. increased RMSQs), rather than with more homogeneous performance (i.e. reduced RMSQs), these analyses were conducted in a one-tailed fashion. Resulting images were thresholded at p < .05 on a voxel level and clusters were extracted using respective scripts from the voxel-based lesion symptom mapping toolbox (Bates et al., 2003).

### 3.5. ROI analysis

Clusters with an extent of at least 50 voxels in the above analysis were considered as ROIs and a subsample of patients was created showing lesions in either of those ROIs. Contributions of individual white matter fiber tracts by overlaying significant clusters with the Atlas of Human Brain Connections (Rojkova et al., 2016), using a probability of .7 as cutoff to binarize fiber tract definition in analysis. A further patient subsample was created by selecting patients with perisylvian and/or temporal lesions (n = 7).

LDT d-primes between action (tool and food) and animal nouns within both patient subsamples and within the healthy control group were compared using permutation tests for repeated measures. Similar to the procedure described above, the null distribution of this test is derived from randomly reallocating values between pairs and the average difference between pairs of this random permutation was determined for 1000 000 iterations. Given the results of Dreyer et al. (2015), we were interested whether action-related concrete categories (i.e. foods and tools) show stronger impairments than animal nouns and hence opted for one-tailed testing and Bonferroni correction was applied to adjust p-scores for multiple comparisons. In addition, LDT and AAT performance, as well as socio-demographic characteristics between patient subsamples, as well as between patients and controls, were compared using a two-sample permutation test as described above.

Please note that ROI definition and ROI analysis are performed in an orthogonal fashion, with ROI definition being based on comparisons to healthy controls and ROI analysis being based comparisons within samples. Nevertheless, this procedure might appear to be susceptible to
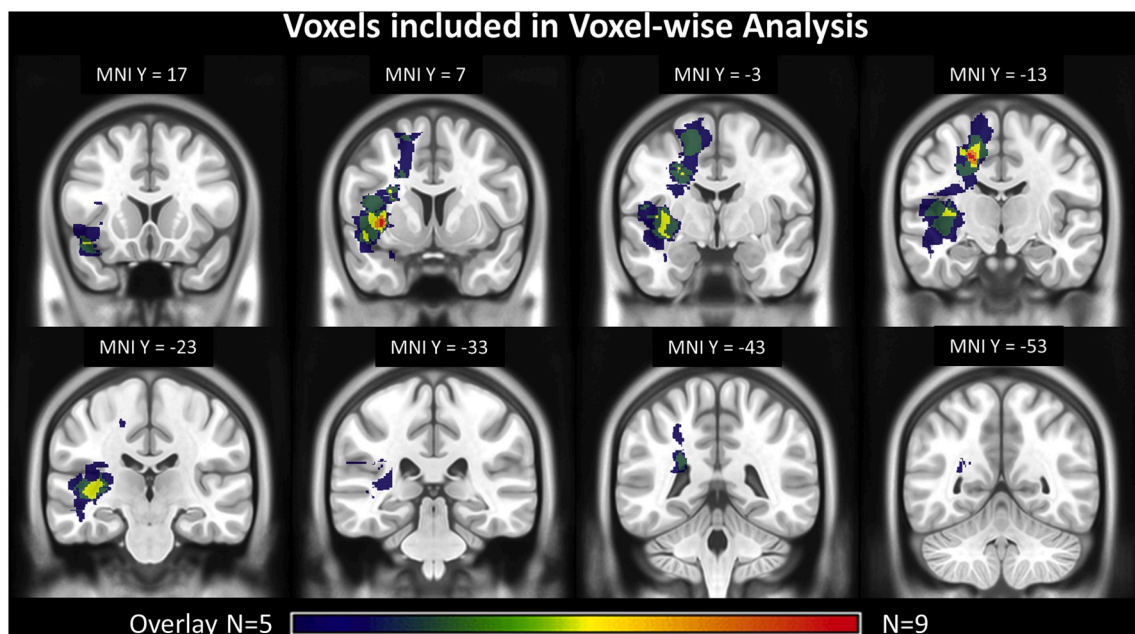


**Fig. 2. Lesion overlay of patients included in analysis.** Color indicates number of overlapping lesions per voxel, ranging from dark blue (N = 5) to red for maximum overlap (N = 9). Each coronal slide is presented with the respective Y coordinate in MNI space. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

potential processing differences between semantic categories, independent of lesion site, which may be masked by ceiling effects in the control sample. To control for any such influence, an additional control patient sample, consisting of patients not allocated to any of the aforementioned samples, was analyzed for between-category differences as described above.

As a further control for an influence of potential performance biases in the sample of healthy controls (like ceiling effects) on results of the initial voxel-wise and hence also the ROI analyses, a separate set of supplementary voxel-wise analyses on the animal-tool and animal-food d-prime differences was conducted. This analysis was based on point-biserial correlations between behavior and lesions in a given voxel, inspired by methods presented in Garcea et al. (2018). A permutation testing strategy was applied in order to determine p-values, similar to the initial voxel-wise analysis, but based on point-biserial correlation coefficients rather than difference of means between patients with and without lesions in a given voxel. As for the initial analysis described above, images were thresholded at a voxel-wise p < .05 and a cluster extent of at least 50 voxels.

## 4. Results

### 4.1. Voxel-wise comparison to healthy controls

Voxel wise comparison to healthy controls revealed three larger clusters (>50 voxels) of significant voxels in dorsal pre- and postcentral grey and white matter (see Table 4 and Fig. 3) to be associated with a RMSQ of between-category differences that was more extreme in patients than in healthy controls. Patient with lesions in these clusters showed an average RMSQ of 0.4 (SE = 0.06), compared to .08 (SE = 0.02) in the healthy controls. Grey matter lesions were situated in Brodmann Area (BA) 6 and BA 4 and significant clusters involved white matter of the Corpus callosum, Frontal Commissural fibers, Superior Longitutinal Fasciculus I and II, Cortical Spinal tract, Pons, Anterior Thalamic Projections, Fronto Striatal fibers, the Frontal Superior Longitudinal tract, as well as the Frontal Aslant tract on the left hemisphere.

Sociodemographic measures of the patient sample considered in VLSM analysis (n = 24) and healthy controls were revealed to be matched in terms of age (p = .96) and years of education (p = .82), thus rendering it unlikely that the voxel-wise comparison was influenced by these factors. Likewise, a strong influence of speed-accuracy tradeoffs on these results appears to be unlikely, as the correlation between average response speed and accuracy was not significant (Pearson $r = -0.24$, $p = 0.27$).

### 4.2. Patient subsample selection and matching

Nine patients were shown to have lesions in at least one of the aforementioned three clusters where evidence for significant heterogeneity across semantic categories was found. This patient sample was considered for ROI analysis of specific between-category differences. As all of those nine patients were shown to have lesions in ROI 1 (dorsal central white matter) and with only three exceptions also in ROI 2 or 3, no further tests for contributions of individual ROIs could be conducted.

**Table 4**
Cluster extent in voxels (k), center of mass (in MNI coordinates) and Brodmann Areas of significant voxels in the voxel-wise comparison between-category d-prime differences root mean squares of patients to healthy controls with k > 50 at p < .05.

|  | k | Center of Mass (MNI) | | | Brodmann Areas |
|---|---|---|---|---|---|
|  |  | x | y | z |  |
| Cluster 1 | 1859 | −13 | −13 | 52 | 6;4 |
| Cluster 2 | 252 | −24 | 0 | 60 | 6 |
| Cluster 3 | 94 | −21 | −14 | 65 | 6 |

Individual lesions of these patients predominantly covered dorsal frontal, central and parietal areas, most notably in BAs 4, 6, 3, 8 and 40. A lesion overlay of this patient selection is given in Fig. 4. The resulting dorsal central ROIs subsample was shown to match healthy controls in terms of age and years of education (all p > .29).

Likewise, the sample of patients with temporal/perisylvian lesions, involving BAs 20, 21, 37, 38, 44 and/or 45 (see Fig. 5 for a lesion overlay) did not differ significantly from healthy controls in neither age (p = .1) nor years of education (p = .13).

Analysis of socio-demographic features between two patient-subsamples showed no significant differences for years of education (p = .26) and lesion size (p = .44), but patients from the perisylvian/temporal lesion subsample were shown to be significantly older (mean age: 57.4 y, S.E. = 4.5 y) than their dorsal lesion counterparts (mean age: 41.33 y, S.E. = 3.4 y; p = . 006). An overview of patient characteristics for both subsamples is given in Table 5.

### 4.3. ROI and subsample analyses

#### 4.3.1. AAT results

Aphasia testing revealed mild to moderate aphasic symptoms in 5 of the 16 patients investigated in the different patient groups. Four of those were found in the sample of primarily perisylvian/temporal lesions and one in the patient sample of dorsal central ROI lesions, whereas the other patients were revealed to be a-symptomatic as far as neurolinguistics deficits are concerned. Furthermore, direct comparisons of T-Scores on the AAT subtest between the two samples showed patients with peri-sylvian lesions to be significantly more impaired than patients with dorsal central lesions in the Token Test (p < .001) and the AAT subscale for Verbal Repetition (p = .028), but both groups were equal in terms of object naming (p = .16), spoken (p = .26) and written word comprehension (p = .18). See Table 4 for AAT scores for both subsamples.

#### 4.3.2. LDT performance

##### 4.3.2.1. Within sample comparisons.
Analysis of between-category differences in the patient sample with lesions in dorsal central ROIs revealed performance for tool nouns (mean d-prime = 2.74, S.E. = 0.27) to be significantly worse than that for animal nouns (mean d-prime = 3.14, S.E. = 0.24; p = .04). At the same time, performance for food nouns (mean d-prime = 3, S.E. = 0.25) was not significantly different from animal nouns (p = .24).

In the subsample of 7 patients selected for their lesions in perisylvian and/or temporal regions, performance for animals (mean d-prime = 2.83, S.E. = 0.47) was not significantly different to that for food (mean d-prime = 2.82, S.E. = 0.45, p = .98) or tool nouns (mean d-prime = 2.65, S.E. = 0.43, p = .3).

In the group of healthy control participants, no significant difference between the performance for animal nouns to neither foods, nor tools was shown in analysis (all p > .62). Mean d-primes in each category and sample are depicted in Fig. 6. Also the additional control sample of patients not allocated to any of the other samples did not exhibit any category specific differences between animals (mean d-prime = 3.48, S. E. = 0.23) and tools (mean d-prime = 3.4, S.E. = 0.2; p = .44) or foods (mean d-prime = 3.57, S.E. = 0.23; p = .7).

##### 4.3.2.2. Between sample comparisons.
Additional post-hoc comparisons between patient subsamples and controls revealed the d-prime difference scores for animals - tools to be higher for patients with dorsal central lesions (mean difference = 0.41, S.E. = 0.17) than for healthy controls (mean difference = −0.05, S.E. = 0.08; p = .014), whereas d-prime differences for animal-foods were not significantly different (Dorsal central patients: mean difference = 0.14, S.E. = 0.11; Controls: mean difference = −0.03, S.E. = 0.11, p = .22) in this comparison. Note that the former finding is equivalent to an interaction effect of the group
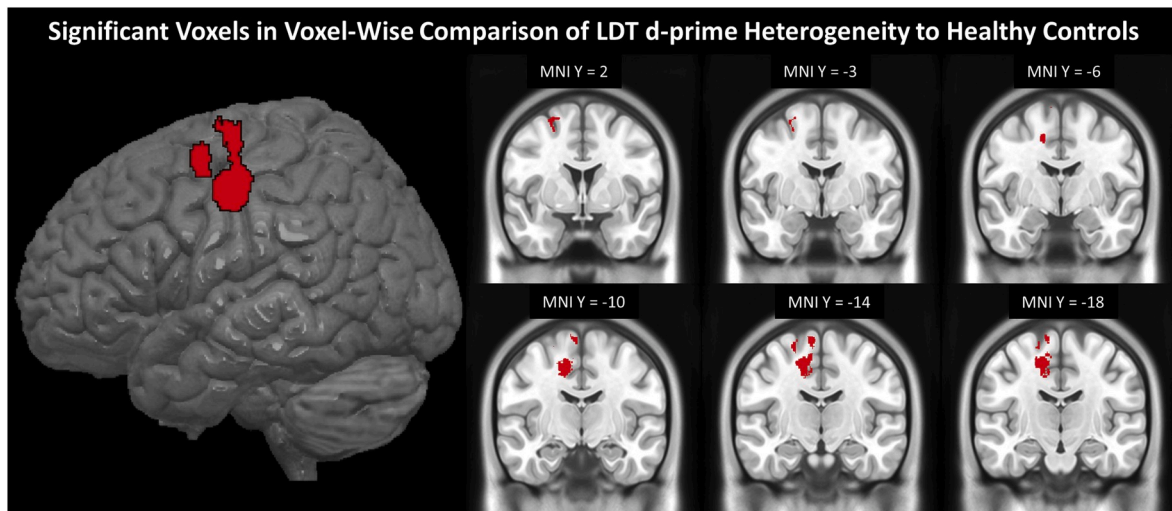
**Fig. 3.** Clusters of significant voxels in the voxel-wise comparison of between-category RMSQs of patients to healthy controls with k > 50 at p < .05.
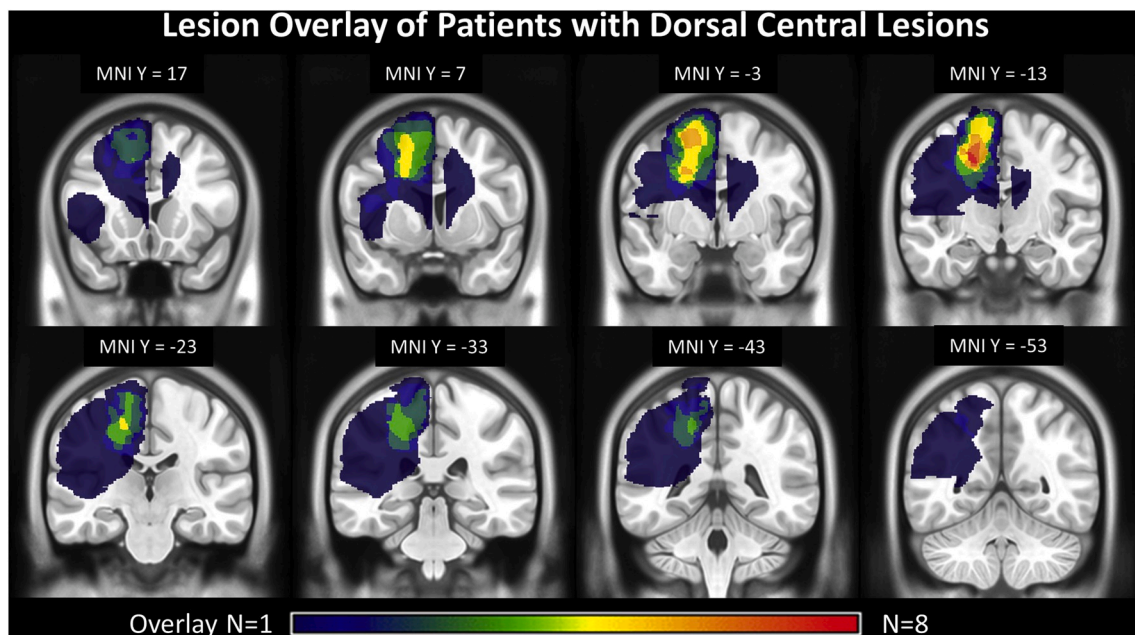


**Fig. 4.** Lesion overlay map of patients with lesions in dorsal central ROIs. Color indicates number of overlapping lesions per voxel, ranging from dark blue (N = 1) to red for maximum overlap (N = 8) in the sample. Each coronal slide is presented with the respective Y coordinate in MNI space. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

(i.e. healthy controls vs dorsal central lesions) and word type (animal vs tool nouns) factors emerging in an analysis of variance, although our results were obtained with non-parametric distribution-free statistics. These same difference scores were revealed to be not significantly different between patients with predominantly perisylvian lesions and healthy controls (animals-tool: mean difference = .18, S.E. = 0.17, p = .6; animals – food: mean difference = 0.01, S.E. = 0.23, p = 1). Direct comparisons between both patient subsamples on the animal-tool d-prime difference scores showed no significant result (p = .31), as it was the case also for animal-food d-prime differences (p = .48). See Fig. 7 for details of these between-sample comparisons.

*4.4. Supplementary voxel-wise point-biserial correlation analysis*

Results of the voxel-wise point-biserial correlations between lesions and animal-tool differences resulted in a single significant cluster in

dorsal central white matter, situated in ROI 1 of the initial analysis. Eight patients were observed to show lesions in this cluster, matching all patients included in the dorsal central lesion sub-sample, with the exception of Patient 15. These 8 patients however showed a stronger animal-tool d-prime difference than the original dorsal central lesion sub-sample (M = 0.46, S.E. = 0.18). The analysis for animal-food d-prime differences did not yield significant clusters. Results for these analyses are presented in Fig. 1. and Table 1 of the Supplementary Materials.

**5. Discussion**

A lexical decision task, probing a set of different semantic noun categories tightly matched on non-semantic psycholinguistic measures, was conducted in a cohort of brain tumor patients and matched healthy control participants. Patients with focal lesions affecting dorsal pre- and
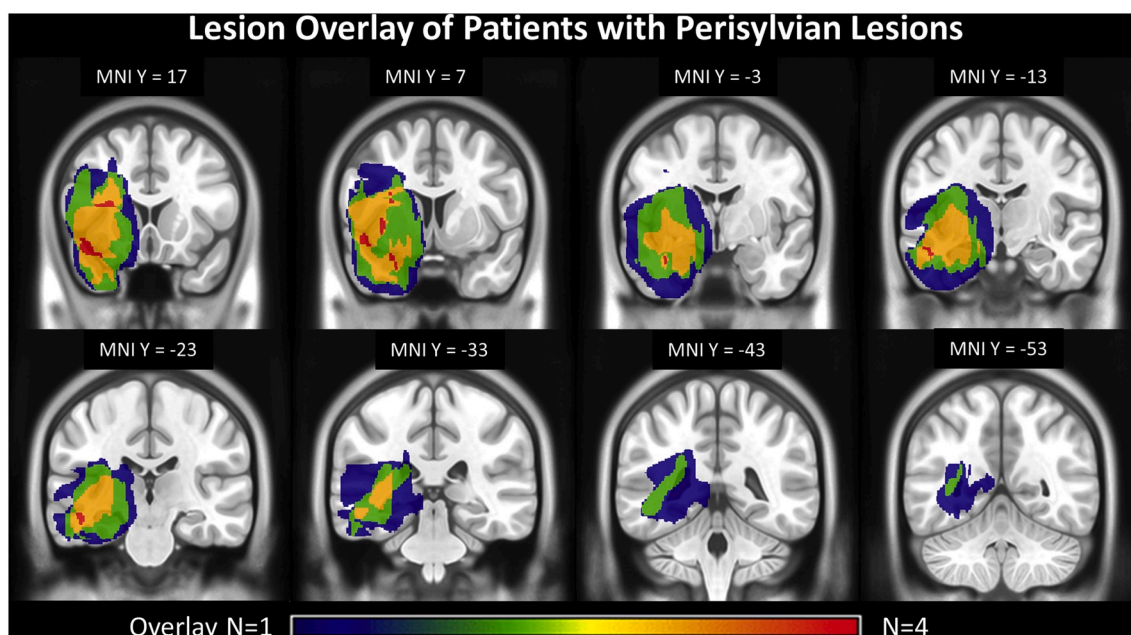
**Fig. 5. Lesion overlay map of patients with lesions in dorsal central ROIs.** Color indicates number of overlapping lesions per voxel, ranging from dark blue ($N = 1$) to red for maximum overlap ($N = 4$) in the sample. Each coronal slide is presented with the respective Y coordinate in MNI space. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

**Table 5**
**Patient characteristics and aphasia testing results patient samples.** AAT performances are given in age-corrected T-scores.

| | Perisylvian Lesion Patients | | Dorsal Central Lesion Patients | |
|---|---|---|---|---|
| | M | S.E. | M | S.E |
| Lesion Size (cm³) | 77.3 | 29.8 | 54.8 | 15.2 |
| Age | 57.4 | 4.5 | 41.3 | 3.4 |
| Handedness (Oldfield LQ) | 90 | 10 | 91.1 | 3.1 |
| Education *(years)* | 14.2 | 1.4 | 16.6 | 1.5 |
| AAT Errors Token Test T Score | 69.1 | 1.5 | 71.4 | .4 |
| AAT Repetition T Score | 69 | 2 | 73.3 | .8 |
| AAT Object Naming T Score | 63.9 | 5.1 | 72.8 | 4.4 |
| AAT Auditory Comprehension T Score | 63.3 | 5.1 | 68.9 | 1.8 |
| AAT Reading Comprehension T Score | 64.7 | 3.9 | 70 | 1.9 |
| AAT Language Comprehension T Score | 64.9 | 4.2 | 73.5 | 1.5 |

postcentral sensorimotor areas showed deficits in the recognition of tool nouns, which were significantly more pronounced than animal nouns used as action-unrelated semantic control category. In contrast, patients with tumors situated predominantly in temporal and inferior-frontal perisylvian regions showed similar performance on different semantic categories, without any significant between-category differences. This latter finding was also seen in a group of healthy control participants. When comparing the difference between the performance on animal nouns and that on tool nouns across groups, contrasting patients with lesions in dorsal pre- and postcentral areas with healthy controls, we found a significantly larger difference in the patients, whereas a similar interaction effect did not reach significance for animal-food noun performance differences. Likewise, the patient sample with temporal/perisylvian lesions failed to show any comparable significant between-category differences relative to healthy controls.

These results support the notion of an indeed functional and necessary - rather than just flexible and optional – role of fronto-parietal networks, including modality specific sensorimotor areas, in the processing of action-related words. We note that language performance in the population of patients with dorsal frontocentral lesions was only mildly to moderately affected in the lexical decision task applied and

still yielded d-prime scores of above 2, thus documenting retained ability to recognize most words correctly. However, and crucially, significantly more errors were made by the members of this dorsal sensorimotor group when the task was to recognize nouns related to objects with hand-action affordances as compared with matched but action-neutral animal nouns. The significant interaction effect showed that, at least in comparison with healthy control subjects, the difference in processing semantic word categories was specific to the dorsal patient group. These results motivate the conclusion that selective focal lesions overlapping in dorsal sensorimotor areas can impair the processing of specific action-related semantic categories. Still, no significant interaction effect emerged in the comparison of our two patient subgroups with dorsal sensorimotor and perisylvian lesions, an issue to which we will return below.

The results on tool nouns are in line with earlier observations of a functional activation of motor areas in processing words with action-related semantics (Hauk et al., 2004; Hauk and Pulvermüller, 2004; Martin et al., 1996; Pulvermüller et al., 2005b; Kemmerer et al., 2008; Carota et al., 2012; Shtyrov et al., 2014; Grisoni et al., 2016). These earlier studies just showed activations of sensorimotor areas, thus just suggesting a functional role of these areas in lexicosemantic processing. Over and above activation studies, lesion studies have indicated that sensorimotor systems also play a functional role in the processing of words and concepts related to action. Since they confirm a relevant role of motor systems for processing specific action-related linguistic types, the current results match those of previous neuropsychological studies of action verb deficits in patients with lesions involving the motor system and adjacent sites (Damasio and Tranel, 1993; Daniele et al., 1994; Bak et al., 2001; Neininger and Pulvermüller, 2001, 2003; Tranel et al., 2003; Bak and Hodges, 2004; Cotelli et al., 2006, 2007; Boulenger et al., 2008; Kemmerer et al., 2012). In contrast to these earlier patient reports, the current study used a strictly controlled paradigm with critical and control word types from the same lexical category (i.e., nouns) matched for a range of relevant psycholinguistic variables and differing only with regard to their action-related semantics. In this experimental context, it is obvious that the observed word type differences in the processing success (measured by d-prime values) of action- and non-action-related words is due to a semantic difference. Furthermore, the fact that the
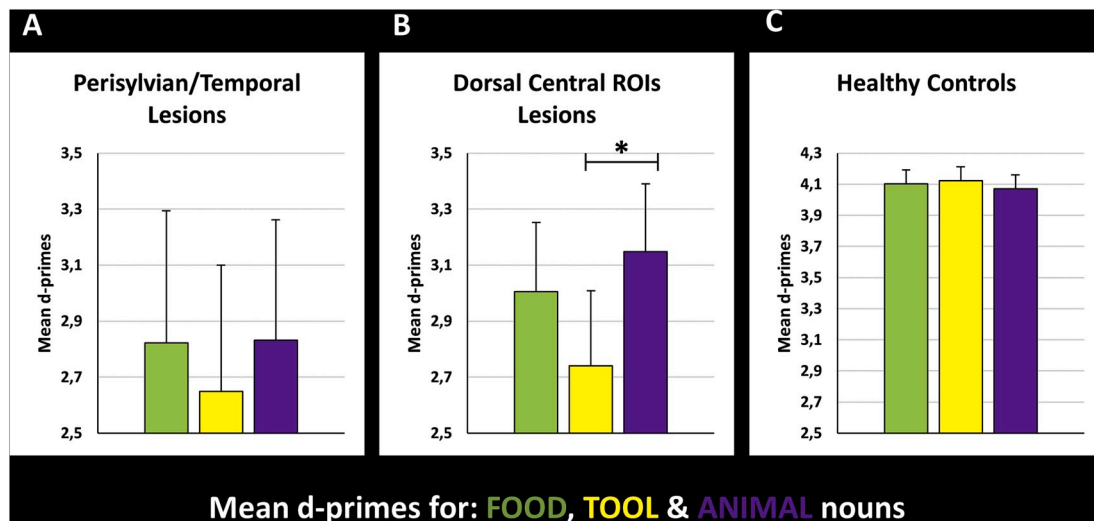
**Fig. 6. Lexical decision results per semantic type and sample of tested individuals.** Bar graphs depict mean d-primes for food (green), tool (yellow) and animal nouns (purple) in the patient sample of patients with predominantly perisylvian lesions (A), patients with dorsal central lesions (B) and healthy controls (C). Error bars depict the standard error of the mean and an asterisk indicates significant between-category differences. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)
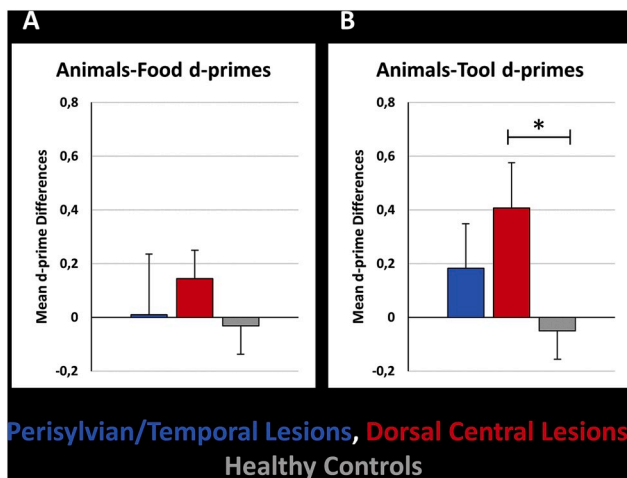


**Fig. 7. Performance differences between animal nouns and action-related semantic word types.** Bar graphs depict mean d-prime differences for animals vs food (A) and animals vs tool (B) nouns in all three populations tested. Results of patients with predominantly perisylvian lesions are depicted in blue, those of patients with dorsal central patients in red and healthy control data in grey. Error bars show the standard error of the mean and an asterisk indicates significant between sample differences. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

observed processing deficit was only seen for words used to speak about hand-action-affording tools, but not for mouth-action-affording food words, suggests at least some degree of semantic specificity. Only those words seem to be significantly affected that relate to objects affording hand actions; words for foods that primarily afford mouth activity do not seem to be affected to the same degree. These same observations had emerged from an earlier study reporting single cases (Dreyer et al., 2015), although any conclusions on somatotopic deficits in processing hand-action-related semantics after lesions in dorsal motor areas could not be generalized to the general population from the limited data presented there. Together with the earlier observation, the significantly reduced d-prime scores of the dorsal group relative to the control words show a necessary role of the motor system in the processing of words

from a specific semantic group. The results were predicted by the semantic topography model (see, e.g., Pulvermüller, 2005, 2013; 2018), according to which the distributed neuronal circuits processing meaningful symbols reach into those sensory and motor areas where information about the semantic grounding of these symbols in action and perception are processed. Therefore, the hand-action-affording symbols would reach into hand-motor areas and lesions of the latter or their connections impair the processing of these symbols specifically.

Importantly, the current observations do not motivate conclusions that extrasylvian dorsal central areas exclusively carry a functional role in processing specific semantic content. Action semantics is not "sitting in", or limited to, the motor system as a caricature of an embodiment approach to concepts might suggest. Such an interpretation would sharply contrast with most grounded theories of language processing (Glenberg and Gallese, 2012; Kiefer and Pulvermüller, 2012; Barsalou, 2016) and in particular with neurocognitive models proposing semantic representations to be realized in distributed cell assemblies spread across both modality specific and domain general areas. We note that mathematically exact neurocomputational modelling studies taking into account cortical structure and connectivity provide independent motivation for the claim that semantic learning leads to the formation of distributed circuits comprising neural elements across these different area types, sensorimotor sites included (Garagnani et al., 2017; Tomasello et al., 2017, 2018, 2019). Our present findings indicate that lesions in fronto-parietal cortex involving the sensorimotor system can cause selective processing deficits largely independent of the presence of lesions in perisylvian temporal, parietal and inferior-frontal regions.

### 5.1. Semantic category specificity of result profiles

In order to validly interpret any of the observed accuracy effects, one must first verify their specificity. Here, the critical issue is to assure that the observed semantic category specific results cannot be attributed to deficits in more basic, non-semantic functions, such as orthographic processing, visual perception or decision processes. This has been achieved by tight matching on psycholinguistic variables on the lexical and sublexical level between semantic categories. This allows to attribute any difference in LDT results to between-category differences in the semantic domain, despite the task design, which would allow proper task performance purely based on lexical knowledge (i.e., information about whether a given item is stored in the 'mental lexicon'), in absence

of any semantic processing. Please note that it is well known that semantic factors can influence performance in the LDT (see e.g. Chumbley and Balota, 1984; Shebani et al., 2017). We chose the LDT for this experiment because it may allow conclusions on implicit semantic processing not necessary for achieving the task, thus motivating conclusions on symbol recognition mechanisms.

As dependent variable, we preferred the d-prime measure to standard accuracy values. This is well motivated, because the use of accuracies comes with the risk of contaminating results by possible response biases and conversion to d-primes is considered to remove this bias. We believe that d-prime values are advantageous in the present case, although one may argue that any unspecific response bias in favor of either 'word' or 'pseudo-word decisions' could not explain category differences. The calculation of d-primes relative to comparable pseudo-words closely matched to the lexical stimuli of interest (in our case 'pseudo-nouns'; see also Shebani et al., 2017) furthermore allows to minimize the likelihood that visual/physical or graphematic differences between word and pseudo-word stimuli influence the results. Neither analysis of healthy controls, nor of patients in the additional control sample, revealed category specific deficits. On this background, it is remarkable that only patients with dorsal frontocentral lesions showed semantic category differences, with tool nouns being more affected than animal nouns. Furthermore, the supplementary voxel-wise point-biserial correlation analysis on the between category differences (animal-tool d-primes) showed overlap of results with those of the initial analysis, by identifying dorsal central lesions to be related to selectively pronounced deficits for tools compared to animal nouns. These observations indicate that the results of the patient sample with dorsal central ROI lesions are not merely driven by a general processing deficit for tool nouns.

Potentially, other non-linguistic lesion-induced cognitive deficits, as, for example, impaired vision, attention, or general executive function required to follow task instructions, may impair LDT performance and could be reflected in reduced general task accuracy. However, one would expect any such general and basal cognitive impairments to affect all semantic categories alike, especially in light of the aforementioned matching of the four different semantic categories. Hence, any of the reported differences between semantic categories, i.e. the comparison of action-related categories of food and tool items to the non-action category of animal nouns, are still interpretable even in the potential presence of some general (non-linguistic) cognitive impairments. The same holds for any preconditions of the patients that are independent of specific lesions, as, for example, the general level of psychological stress, given that the majority of patients had brain surgery scheduled for the days following experimental testing. The influence of these lesion site independent deteriorating effects was further controlled for by comparisons of results within patients groups, so that those non-lesion site dependent effects can be assumed to cancel each other out in analysis. The insufficient matching in terms of age between the patient groups with dorsal motor and perisylvian lesions (with the former being younger in average than the latter) could potentially have biased direct comparisons between these two patient samples. However, one would assume such a bias to result in improved performance for patients with dorsal central lesion when contrasted with the sample of predominantly perisylvian/temporal lesions, where the most severe linguistic deficits are normally observed. Hence, any potential selective deficit observed only in dorsal motor/parietal patients can still be soundly interpreted. When compared to the group of healthy controls, all patients as a whole, as well as individual patient samples were matched for age and education levels (all $p > .05$). Therefore, neither age nor education could explain any differences in LDT performance profiles between patients and healthy controls. Furthermore, the insufficient matching of general language impairments (as revealed by the AAT) between patient groups, with the perisylvian group showing more severe deficits, could have led to performance biases in the LDT for between-sample comparisons. In particular, one may argue that the more severe language deficits seen

after persylvian lesions lead to a floor effect reducing the degree of category-specificity of semantic deficits. Although we cannot fully rule out this possibility, we should still point to the relatively high d-prime values achieved by our perisylvian population (most values > 2.5), which makes a floor effect unlikely. In addition, it is important to note that comparisons of category specific performance (i.e. tool, food vs. animal nouns) within one sample can still be soundly derived, as the presence of mild to moderate general aphasic symptoms would not predict differences in performance between categories.

### 5.2. Local specificity of results

The study of neuropsychological deficits was, for a long time, dominated by work with stroke patients. This was so for good reason: the cortical tissue functionally affected by tumors is less easily determined as tissue affected by an ischemic or hemorrhagic stroke. In the latter cases, the lesion can be easily delineated and any edema or penumbra areas where function is also impaired are typically in close vicinity of the lesion. In the case of tumors however, the situation is more difficult, since intrinsic brain tumors have a tendency to diffusely infiltrate surrounding brain tissue. For this and other similar reasons, many neuropsychologists remained hesitant about strongly interpreting lesion studies with tumor patients. As brought forward by Shallice and Skrap (2011) however, the aforementioned disadvantages of tumor patient investigations also apply to patient populations of other etiologies (including strokes). Taking into account the potential drawbacks of tumor patient studies we spent much care to minimize the risk of unjustified inferences from our present patient cohort. To this end patients with glioblastomas (Gliomas of WHO Grade IV) were excluded from analysis, as these tumors are characterized by pronounced diffuse infiltration of surrounding tissue, and two experts in tumor surgery (TP and DF) were consulted when delineating the exact extent of the lesions.

As most of the present tumors started growing in subcortical tissue and extend into cortical areas, rather than affecting either white or grey matter exclusively (with the exception of one patient that presented an acute disseminated encephalomyelitis, a focal lesion of white matter tracts), white and grey matter lesions do not occur in isolation in individual patients of the current patient cohorts investigated. Hence, the potentially different effects of white and grey matter lesions on processing action-related words can, at least in the current patient samples, not be separated from another.

An additional issue concerning local lesion specificity lies in the fact that any edemas or hemorrhages surrounding the tumor were marked as lesioned tissue included in lesion maps, although it can generally not be inferred a priori whether, or which parts of the nervous tissue within those edemas or hemorrhages are fully dysfunctional, partly functional or still asymptomatic (Karnath and Steinbach, 2011). This very liberal method of lesion definition was chosen as a precaution to prevent under-estimation of lesion tissue. This also works against the spatially specific hypotheses tested by this study, addressing specific contributions of sensorimotor regions to the processing of concrete action-related words, though, in doing so, the spatial lesion profile specificity was decreased across the whole patient sample. The alternative would have been to include only those patients without surrounding edema or hemorrhages. However, this procedure would have resulted only in very small sample sizes. At the same time, the other alternative would have been to disregard non-tumor tissue entirely in lesion definition, but this procedure would have come with the risk that lesions were ignored. We opted against these latter options, as we aimed to avoid any possible overestimation of the spatial focality when interpreting our results. Following a further consideration brought forward by Karnath and Steinbach (2011), functional reorganization of cortical and subcortical regions during (in some cases very slow) tumor growth are likely to have occurred in the patient samples. As one cannot tell whether or not specific lesioned voxels have their original function restored in remaining healthy brain tissue, a one-to-one correspondence of lesion

maps and functional impairment faces serious difficulty. However, following Duffau (2011), those reorganization processes are likely to restore functions to levels below normal performance, so that deficit mapping would still be possible. In addition, benefits of reorganization processes should normally benefit all impaired word categories equally and do not predict the creation of category specific differences of results, but rather their removal. Hence, the category specific impairment for tool nouns present in the current sample of dorsal central lesions, can indeed be soundly interpreted in terms of a functional role of modality preferential systems for semantic processing, given that specificity of lesion profiles is warranted.

Contrary to the initial predictions, a similar category specific deficit was not observed for food nouns related to face-actions. The reason for this observation might be seen in the lesion profiles in the patient sample analyzed in this study. Reviewing the individual lesion profiles, only one patient showed a lesion in ventral face motor areas exclusively, whereas in the other cases the lesions affected perisylvian or dorsal motor areas, too. This low lesion focality in the face motor system might have occluded a category specific deficit for food-nouns, as predicted initially.

We note again that the strongest evidence provided by this study came from a comparison between groups. Whereas age-matched healthy control subjects processed all word categories alike in the LDT, the group of tumor patients with dorsal sensorimotor lesions showed the predicted deficit for tool nouns. This result supports the specificity of the action-affording word deficit in this patient population. In addition, the absence of similar interaction in the comparison between patients with perisylvian lesions and healthy controls further shows that the category-specificity in the dorsal group is not a general consequence of a tumor or focal lesion. However, the expected interaction effect resulting from comparing the animal-tool d-prime difference scores between patient samples with dorsal central and perisylvian lesions did not reach significance. This is an important issue, as the semantic circuit model would have predicted such an interaction, a prediction which our data did not confirm. Because of the relevance of this potential effect and its absence, we discuss below several possibilities why the predicted effect was not obtained.

According to the distributed semantic circuit account, nouns typically used to refer to action-affording tools would be cortically represented by distributed mini-networks comprising neurons in dorsal frontocentral sensorimotor areas, so that lesions there would functionally impair these circuits gradually relative to control words lacking semantic action knowledge – in our case the animal words. Patients with lesions in perisylvian cortex but without sensorimotor cortex involvement should not show such a dissociation, so that the comparison of the perisylvian and dorsal frontocentral groups should result in a significant interaction effect, with relatively strong animal vs. tool word performance differences in the dorsal group and no such difference in the perisylvian population. The present results did not provide support for such an interaction effect, although it is notable that the additional point-biserial correlation VLSM found significant clusters for animal-tool d-primes in dorsal central, but not perisylvian or temporal areas. Likewise, a further post-hoc comparison of animal-tool d-primes of the dorsal central subsample against all other patients (see Fig. 2 in Supplementary Materials) showed at least a near significant difference (p = .053). Results of both additional analyses thus suggest a degree of spatial specificity in the observed relatively pronounced tool deficit of the dorsal central sample.

Another explanation for the absent interaction in the animal-tool d-prime differences between the perisylvian and the dorsal sample may be seen in the individual lesion topographies. Ideally and according to the (sub-) sampling criteria, lesions in the perisylvian sample covered temporal and/or inferior frontal grey and white matter, while sparing dorsal central areas. However, the lesions in that group were in fact not confined to perisylvian regions. Instead, as depicted in Fig. 4, individual lesions spread into posterior temporal, occipital and extensive inferior to middle frontal areas as well. Some of those areas are according to

previous studies, involved in tool naming (left insula and posterior middle temporal gyrus); see for example a recent activation likelihood estimation meta-analysis (Ishibashi et al., 2016) and a review of the neural substrates of tool processing (Lewis, 2006). Therefore, any lesion here, as presented in some of the patients from the perisylvian sample, might have resulted in some impairment of tool-compared to animal nouns which in turn may have occluded a potential interaction with the results from the dorsal central sample. This interpretation sees further support in very recent findings by Riccardi and coworkers (2019) in a cohort of stroke patients. In this study voxel-based lesion–symptom mapping, resting-state functional connectivity and grey matter fractional anisotropy analyses on the performance in a semantic similarity judgement task provided converging evidence for an involvement of anterior parietal and posterior temporal regions (in addition to premotor areas) in tool noun processing. In this context, we have to ask, whether the lesion profile of our present patient population is open to the claim that parietal lesions co-occurring with the impairments in sensorimotor per- and post-central areas could account for the pattern of deficits observed. Looking more closely at the lesion profiles of our 9 patients in the dorsal central lesion group, we found that 3 of these patients had lesions extending into posterior inferior parietal cortex. However, the remaining 6 patients did not show substantial involvement of inferior parietal cortex. Therefore, our present results are not explainable as a consequence of inferior parietal lesion. Taken together, the study by Riccardi and coworkers (2019) and our present contribution show that category specific semantic deficits can arise from lesions in the dorsal stream of action processing, which includes classic key areas of the mirror neuron system.

### 5.3. Integrating category specific impairments into theory

The observed category specific deficit directly supports the notion of a causal role of sensorimotor areas to processing concrete hand-action-related tool nouns. Previous neurostimulation studies (e.g. Pulvermüller et al., 2005a, 2005b; Willems et al., 2011) already provided support for this claim, but the current study in patients with lesions in the critical sensorimotor areas go further than this earlier work. As these patients showed larger numbers of errors in the LDT when processing tool nouns as compared with animal nouns, a processing deficit is obvious. Therefore, we can claim that, at least for optimal errorless processing of this semantic type comparable to the processing of other semantic categories, intact sensorimotor cortices are necessary. Therefore, in contrast to earlier TMS studies, the current investigation points to a necessary, rather than merely fascilitatory, role of sensorimotor areas to semantic category processing, as demonstrated by the reported dissociations in error rates (i.e. d-primes) between tool and animal nouns.

The documented effect falsifies strong symbolic approaches claiming that sensorimotor systems and semantic systems are separated and situated in different modules. Strictly speaking, this requires the additional assumption that the cortical loci of these different systems are different, and, particularly, that the sensorimotor precentral and postcentral cortices are not the location where also the 'semantic module' or semantic hub is localized. However, the assumption that the sensorimotor system houses semantics (which was once discussed by Mahon and Hickok, 2016) appears not very plausible, given that proposals for 'semantic hub' localization include inferior frontal, anterior temporal and inferior parietal cortices, but not the pre- and post-central gyrus. Therefore, the strong symbolism approach lacks viability in light of the present data.

Still, one may argue that a weaker model of symbolic concepts, where these are allowed to occasionally and flexibly interact with sensorimotor information (see, for example, Mahon and Caramazza, 2008; Caramazza et al., 2014), for the purpose of 'enrichment' or 'coloring', could accommodate them. However, we are not convinced that such a construct provides a feasible explanation of the data reported.

Given that a processing deficit, even though only manifest in a slight increase in error rates and drop in d-primes, occurred, how could the removal of optional and therefore unessential processes explain it? The significantly increased number of errors in recognizing action-affording tool words in our patients with sensorimotor lesions shows that they lack something essential in the word recognition process, not a process that can be flexibly added or removed.

As proposed by Mahon and Hickok (2016) it could still be possible, theoretically and very speculatively, that a functional role of the motor systems is the result of entirely amodal representations which (despite their amodality) happen to be localized in the motor systems, potentially even following a pattern of effector-specific semantic somatotopy. We agree that this putative position exists: the motor system could indeed house abstract "amodal" (in our view multimodal) semantic processes. Only, just stipulating such a connection is a trivial consequence of taking into account the data, those about brain activation or the present neuropsychological ones. However, we believe that a key question in cognitive neuroscience is why such function-structure relationships develop. The topographically specific semantic circuits developing consequent to learning contingencies between body actions and word usage enforce the prediction that a lesion in sensorimotor cortices can have a category-specific effect on semantic processes. In contrast, Mahon & Hickok's proposal repeats established facts without attempting at an explanation. Purely cognitive theories without a neurobiological foundation are indeed open to any brain locus of their processing components, semantics included. However, in order to make progress in the brain language sciences, it is advantageous to prefer models about the nature of the neural bases of semantic representation and processing, which can be used to explain and even predict experimental data such as those observed in the current investigation.

One set of theories that provides exactly these features are those of grounded (sometimes also called 'embodied') semantic representations, according to which knowledge about referential links of signs is stored in (but not exclusively in) sensorimotor systems (Pulvermüller, 1999, 2005; Barsalou et al., 2003; Pulvermüller and Fadiga, 2010; Glenberg and Gallese, 2012; Kiefer and Pulvermüller, 2012; Strijkers and Costa, 2016). In neural terms this can be achieved by cell assemblies spreading over multi-modal perisylvian areas and extending into modality specific systems, incuding the motor system. The exact structure of these cell assemblies is believed to be shaped by basic principles of correlational neural learning, Hebbian and Anti-Hebbian learning (Hebb, 1949). For concrete tool words like 'pliers' or 'screwdriver', their usage would sometimes co-occur with actual motor performance and perception of the respective objects. Hence the linguistic sign, the word form, stored in inferior-frontal and temporal perisylvian regions, would gradually be connected to aspects of its meaning in terms of respective referenced sensorimotor information. Accordingly, the underlying cell assemblies would extend from perisylvian areas into the respective motor and sensory areas, thus directly reflecting the relation of sign and reference. In contrast, symbolic models view the activation of modality-preferential areas to be a consequence of genuine semantic processing but not partly constitutive for it (Mahon and Caramazza, 2008; Caramazza et al., 2014; Mahon, 2015). Once a base vocabulary of symbols has directly been grounded in action and perception knowledge, novel symbols can be acquired based on combinatorial learning, where semantic features may be transferred between symbol representations (Harnad, 1990; Cangelosi et al., 2002; Pulvermüller, 2002). However, the modality specific components grounding meaning in action and perception would still remain a necessary and functional part of the underlying semantic circuits – along with the cognitive processes they make possible, including the capacity to relate multimodal linguistic signs to their meaning via experience.

It has been argued previously that lesion of sensorimotor systems normally only result in "subtle, rather than catastrophic" general cognitive or language deficits (Binder and Desai, 2011) and also the current effect sizes seem to support this notion, as most patients still performed well above chance in the affected categories, despite the reported category specific impairments. However, this must not be seen as evidence against the aforementioned cell assemblies and the role of sensorimotor areas, as the widespread nature of these assemblies predict a large degree of redundancy, allowing the assembly to potentially still sufficiently ignite in case some of its extrasylvian motor nodes are missing (Neininger and Pulvermüller, 2003). As the patients investigated in the current analysis were either entirely a-symptomatic regarding their motor function or presented only motor impairments of mild severity, it is likely that effector-specific motor components were not lesioned in their entirety. Regarding the present study, this point applies to the motor circuits within the left hemisphere and of course particularly for the intact motor areas of the right hemisphere and subcortical structures in the basal ganglia and even the cerebellum, some of which are regularly active in action symbol processing (Hauk and Pulvermüller, 2011; Carota et al., 2012). It is therefore possible that representations of some tool concepts were left largely intact in the individual case, though to a degree that still allowed for category specific performance differences in a demanding and well-controlled lexical decision paradigm. Unfortunately, standard clinical neurological and neurophysiological investigations of the patients presented here did not contain any tests that would diagnose semantic category specific deficits, however from the LDT results alone, statements about the nature of underlying neural processing of single word recognition and related semantic representations can still be soundly derived.

## 5.4. Relation to earlier findings of neural correlates of tool usage and conceptual processing

Cortical areas affected most consistently in the patient sample with dorsal central lesions, i.e. dorsal motor and premotor cortex, as well as the supplementary motor cortex in the left hemisphere, have previously been identified to be part of a wider fronto-parietal network involved in actual tool usage in recent meta-analyses (Lewis, 2006; Ishibashi et al., 2016; Reynaud et al., 2016). Furthermore, white matter lesions in this sample included the superior longitudinal fasciculus and the anterior thalamic radiation, two fiber tracts which were indicated to be relevant for tool usage and conceptual processing in a recent diffusion tensor imaging (DTI) analyses (Bi et al., 2015). This overlap of neural substrates for tool usage and semantics of tool concepts is in line with predictions of the cell assembly account of lexical semantics, as outlined in the previous paragraphs. At the same time, previous literature on phenotypical dissociations between tool usage and their conceptual processing appear to be conflicting with the current results. Especially in the field of apraxia research, a range of studies report pronounced impairments of tool usage while the conceptual and semantic processing, as shown by tool picture naming was unimpaired, or impaired to a lesser degree than tool usage per se (Rumiati et al., 2001; Rosci et al., 2003; Negri et al., 2007; Garcea et al., 2013; Vannuscorps et al., 2016). In addition, those findings appear to be not limited to apraxia patients, as Vannuscorps and coworkes (2014) present a case report of an upper limp aplasia patient who showed no significant difference in picture naming of objects with and without motor knowledge. Such dissociations have previously been interpreted to be problematic for embodied or grounded accounts of semantic representations (e.g. Negri et al., 2007; Mahon and Caramazza, 2008; Vannuscorps et al., 2014, 2016; Mahon, 2015; Mahon and Hickok, 2016).

However, we do not believe that these results are incompatible but may be explainable independent of any theoretical framing. The shared substrates are consistent with models postulating that semantic and conceptual knowledge can draw upon knowledge about the shape of objects and about how to use these objects in bodily actions. The explanation of the dissociation seem straightforward. When using a tool, it is of course relevant to see the object, estimate its spatial position and its relationship to the own body. Such visual and spatial information is less relevant when naming the same object. On the other hand,

articulatory motor planning and fine grained motor movements are required in the naming exercise, abilities not so relevant when handling a tool. Furthermore, following the argument outlined in the final paragraphs of 5.3, the reported sparing of conceptual processing for tools may be explained by a level of redundancy of the underlying cell assemblies in the motor domain. In case of aplasia patients (as reported by Vannuscorps et al., 2014), motor knowledge may be absent, impaired or altered, and it is difficult to predict how the cortical reorganization processes in the deprived brain shape precisely (but see Tomasello et al., 2019 for a model). Even in case when the motor components of such a cell assembly may be lost completely following brain lesions (e.g. in apraxia patients), other unimpaired modality specific semantic components, like visual, haptic or auditory information (see also semantic ratings in Fig. 2), may be sufficient to ignite the (now lesioned) cell assembly representing a tool concept in order to name it in a picture naming paradigm. Thus, we see the previous results pointing to a dissociation of tool action and concept processing to be, if at all, problematic for unimodal accounts of semantic representation, rather than for the multimodal representations presented in the cell assembly model of semantic representations (Pulvermüller, 1999, 2005; 2018). To build on an analogy introduced by Yee and Thompson-Schill (2016), the loss of one finger does not entail losing use of the entire hand - but it may very well impair its function, in particular under conditions of high task demand or artificial constrains - as present for semantic processing of nouns in the speeded LDT applied in the current study.

*5.5. Limitations and need for future research*

With a sample size of 24, the patient sample under investigation was substantially smaller than previous VLSM approaches on the neural substrates of category specific semantics (e.g. Arevalo et al., 2012; Kemmerer et al., 2012). This small sample size, and especially the limited overlap of individual lesion profiles (maximum overlap N = 9), resulted in the need for non-parametric testing, which reduced the statistical power, so that no direct inferences could be drawn on the VLSM findings alone. In addition, the inferential power of the current analysis is restricted by the present lesion profiles, which allowed for careful conclusions on dorsal sensorimotor and perisylvian lesions, but did not allow for statements about the relevance of other cortical areas for semantics. Especially the absence of focal lesions in ventral motor areas related to face movements rendered it impossible to appropriately test predictions of grounded approaches for face-action related food items.

## 6. Conclusion

Category specific semantic deficits in a LDT were observed in a patient sample with focal fronto-parietal lesions, including sensorimotor areas. Processing of concrete tool nouns was selectively impaired when compared to a non-action baseline category of animal nouns after lesions of dorsal prefrontal, sensorimotor central and parietal areas. In contrast, patients with predominantly perisylvian and temporal lesions and also healthy age and education matched controls did not show such a category specificity of results. This confirms the functional necessity of extrasylvian dorsal sensorimotor areas for the processing of concrete hand-action-affording nouns, in line with grounded approaches towards semantic processing and representation.

## CRediT authorship contribution statement

**Felix R. Dreyer:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Software, Validation, Visualization, Writing - original draft, Writing - review & editing. **Thomas Picht:** Data curation, Funding acquisition, Investigation, Resources, Supervision, Validation, Writing - review & editing. **Dietmar Frey:** Data curation, Investigation, Resources, Supervision, Validation, Writing - review & editing. **Peter Vajkoczy:** Resources. **Friedemann Pulvermüller:** Conceptualization, Data curation, Methodology, Supervision, Validation, Writing - original draft, Writing - review & editing.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.neuropsychologia.2020.107384.

## References

Anderson, J.R., 1983. The Architecture of Cognition. Erlbaum, Hillsdale, NJ.

Arevalo, A.L., Baldo, J.V., Dronkers, N.F., 2012. What do brain lesions tell us about theories of embodied semantics and the human mirror neuron system? Cortex 48 (2), 242–254.

Bak, T.H., Hodges, J.R., 2004. The effects of motor neurone disease on language: further evidence. Brain Lang. 89 (2), 354–361.

Bak, T.H., O'Donovan, D.G., Xuereb, J.H., Boniface, S., Hodges, J.R., 2001. Selective impairment of verb processing associated with pathological changes in Brodmann areas 44 and 45 in the motor neurone disease–dementia–aphasia syndrome. Brain 124 (1), 103–120.

Barrós-Loscertales, A., González, J., Pulvermüller, F., Ventura-Campos, N., Bustamante, J.C., Costumero, V., Ávila, C., 2012. Reading salt activates gustatory brain regions: fMRI evidence for semantic grounding in a novel sensory modality. Cerebr. Cortex 22 (11), 2554–2563.

Barsalou, L.W., 1999. Perceptions of perceptual symbols. Behav. Brain Sci. 22 (4), 637–660.

Barsalou, L.W., 2008. Grounded cognition. Annu. Rev. Psychol. 59, 617–645.

Barsalou, L.W., 2016. On staying grounded and avoiding quixotic dead ends. Psychon. Bull. Rev. 23 (4), 1122–1142.

Barsalou, L.W., Simmons, W.K., Barbey, A.K., Wilson, C.D., 2003. Grounding conceptual knowledge in modality-specific systems. Trends Cognit. Sci. 7 (2), 84–91.

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.

Bi, Y., Han, Z., Zhong, S., Ma, Y., Gong, G., Huang, R., Caramazza, A., 2015. The white matter structural network underlying human tool use and tool understanding. J. Neurosci. 35 (17), 6822–6835.

Binder, J.R., Desai, R.H., 2011. The neurobiology of semantic memory. Trends Cognit. Sci. 15 (11), 527–536.

Boulenger, V., Mechtouff, L., Thobois, S., Broussolle, E., Jeannerod, M., Nazir, T.A., 2008. Word processing in Parkinson's disease is impaired for action verbs but not for concrete nouns. Neuropsychologia 46 (2), 743–756.

Brett, M., Leff, A.P., Rorden, C., Ashburner, J., 2001. Spatial normalization of brain images with focal lesions using cost function masking. Neuroimage 14 (2), 486–500.

Cangelosi, A., Greco, A., Harnad, S., 2002. Symbol grounding and the symbolic theft hypothesis. In: Simulating the Evolution of Language. Springer London, pp. 191–210.

Caramazza, A., Hillis, A.E., Rapp, B.C., Romani, C., 1990. The multiple semantics hypothesis: multiple confusions? Cogn. Neuropsychol. 7 (3), 161–189.

Caramazza, A., Anzellotti, S., Strnad, L., Lingnau, A., 2014. Embodied cognition and mirror neurons: a critical assessment. Annu. Rev. Neurosci. 37, 1–15.

Carota, F., Moseley, R., Pulvermüller, F., 2012. Body-part-specific representations of semantic noun categories. J. Cognit. Neurosci. 24 (6), 1492–1509.

Choi, S., Na, D.L., Kang, E., Lee, K., Lee, S., Na, D., 2001. Functional magnetic resonance imaging during pantomiming tool-use gestures. Exp. Brain Res. 139 (3), 311–317.

Chumbley, J.I., Balota, D.A., 1984. A word's meaning affects the decision in lexical decision. Mem. Cognit. 12 (6), 590–606.

Cotelli, M., Borroni, B., Manenti, R., Alberici, A., Calabria, M., Agosti, C., Zanetti, O., 2006. Action and object naming in frontotemporal dementia, progressive supranuclear palsy, and corticobasal degeneration. Neuropsychology 20 (5), 558.

Cotelli, M., Borroni, B., Manenti, R., Zanetti, M., Arévalo, A., Cappa, S.F., Padovani, A., 2007. Action and object naming in Parkinson's disease without dementia. Eur. J. Neurol. 14 (6), 632–637.

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.

Damasio, A.R., Tranel, D., 1993. Nouns and verbs are retrieved with differently distributed neural systems. Proc. Natl. Acad. Sci. Unit. States Am. 90 (11), 4957–4960.

Daniele, A., Giustolisi, L., Silveri, M.C., Colosimo, C., Gainotti, G., 1994. Evidence for a possible neuroanatomical basis for lexical processing of nouns and verbs. Neuropsychologia 32 (11), 1325–1341.

Dreyer, F.R., Pulvermüller, F., 2018. Abstract semantics in the motor system?–An event-related fMRI study on passive reading of semantic word categories carrying abstract emotional and mental meaning. Cortex 100, 52–70.

Dreyer, F.R., Frey, D., Arana, S., von Saldern, S., Picht, T., Vajkoczy, P., Pulvermüller, F., 2015. Is the motor system necessary for processing action and abstract emotion words? Evidence from focal brain lesions. Front. Psychol. 6, 1661.

Duffau, H., 2011. Do brain tumours allow valid conclusions on the localisation of human brain functions? Cortex 47 (8), 1016–1017.

Edgington, E.S., 1964. Randomization tests. J. Psychol. 57 (2), 445–449.

Ellis, A.W., Young, A.W., 1988. Human Cognitive Neuropsychology. Lawrence Erlbaum Associates Ltd, Hove, UK.

Fernandino, L., Conant, L L, Binder, J R, Blindauer, K, Hiner, B, Spangler, K, Desai, R H, 2013. Parkinson's disease disrupts both automatic and controlled processing of action verbs. Brain Lang 127 (1), 65–74.

Fodor, J.A., 1983. The modularity of mind. MIT press.

Garcea, F.E., Dombovy, M., Mahon, B.Z., 2013. Preserved tool knowledge in the context of impaired action knowledge: implications for models of semantic memory. Front. Hum. Neurosci. 7, 120.

Garagnani, M., Lucchese, G., Tomasello, R., Wennekers, T., Pulvermüller, F., 2017. A spiking neurocomputational model of high-frequency oscillatory brain responses to words and pseudowords. Front. Comput. Neuroscience 10, 145.

Garcea, F.E., Almeida, J., Sims, M.H., Nunno, A., Meyers, S.P., Li, Y.M., Mahon, B.Z., 2018. Domain-specific diaschisis: lesions to parietal action areas modulate neural responses to tools in the ventral stream. Cerebr. Cortex 1, 14.

Ghio, M., Vaghi, M.M.S., Perani, D., Tettamanti, M., 2016. Decoding the neural representation of fine-grained conceptual categories. Neuroimage 132, 93–103.

Glenberg, A.M., Gallese, V., 2012. Action-based language: a theory of language acquisition, comprehension, and production. Cortex 48 (7), 905–922.

González, J., Barros-Loscertales, A., Pulvermüller, F., Meseguer, V., Sanjuán, A., Belloch, V., Ávila, C., 2006. Reading cinnamon activates olfactory brain regions. Neuroimage 32 (2), 906–912.

Grisoni, L., Dreyer, F.R., Pulvermüller, F., 2016. Somatotopic semantic priming and prediction in the motor system. Cerebr. Cortex 26 (5), 2353–2366.

Grisoni, L., Mohr, B., Pulvermüller, F., 2019. Prediction mechanisms in motor and auditory areas and their role in sound perception and language understanding. Neuroimage. https://doi.org/10.1016/j.neuroimage.2019.05.071. Advanced online publication.

Hauk, O., Pulvermüller, F., 2004. Neurophysiological distinction of action words in the fronto-central cortex. Hum. Brain Mapp. 21 (3), 191–201.

Hauk, O., Pulvermuller, F., 2011. The lateralization of motor cortex activation to action-words. Front. Hum. Neurosci. 5, 149.

Harnad, S., 1990. The symbol grounding problem. Phys. D: Nonlinear Phenom. 42 (1–3), 335–346.

Hauk, O., Johnsrude, I., Pulvermüller, F., 2004. Somatotopic representation of action words in human motor and premotor cortex. Neuron 41 (2), 301–307.

Hebb, D.O., 1949. The Organization of Behavior: A Neuropsychological Approach. John Wiley & Sons.

Heister, J, Würzner, K-M, Bubenzer, J, Pohl, E, Hanneforth, T, Geyken, A, Kliegl, R, 2011. dlexDB – einelexikalische Datenbank für die psychologische undlinguistische Forschung. Psychologische Rundschau 62 (1), 10–20.

Huber, W., 1983. *Aachener Aphasie-Test:(AAT)*. Hogrefe, Verlag für Psychologie.

Ishibashi, R., Pobric, G., Saito, S., Lambon Ralph, M.A., 2016. The neural network for tool-related cognition: an activation likelihood estimation meta-analysis of 70 neuroimaging contrasts. Cogn. Neuropsychol. 33 (3–4), 241–256.

James, C.T., 1975. The role of semantic information in lexical decisions. J. Exp. Psychol. Hum. Percept. Perform. 1 (2), 130.

Johnson-Frey, S.H., 2004. The neural bases of complex tool use in humans. Trends Cognit. Sci. 8 (2), 71–78.

Karnath, H.O., Steinbach, J.P., 2011. Do brain tumours allow valid conclusions on the localisation of human brain functions?–Objections. Cortex 47 (8), 1004–1006.

Kemmerer, D., Castillo, J.G., Talavage, T., Patterson, S., Wiley, C., 2008. Neuroanatomical distribution of five semantic components of verbs: evidence from fMRI. Brain Lang. 107 (1), 16–43.

Kemmerer, D., Rudrauf, D., Manzel, K., Tranel, D., 2012. Behavioral patterns and lesion sites associated with impaired processing of lexical and conceptual knowledge of actions. Cortex 48 (7), 826–848.

Keuleers, E, Brysbaert, M, 2010. Wuggy: A multilingual pseudoword generator. Behav. Res. Meth. 42 (3), 627–633.

Kiefer, M., Pulvermüller, F., 2012. Conceptual representations in mind and brain: theoretical developments, current evidence and future directions. Cortex 48 (7), 805–825.

Kiefer, M., Sim, E.J., Herrnberger, B., Grothe, J., Hoenig, K., 2008. The sound of concepts: four markers for a link between auditory and conceptual brain systems. J. Neurosci. 28 (47), 12224–12230.

Lewis, J.W., 2006. Cortical networks related to human use of tools. Neuroscientist 12 (3), 211–231.

Mahon, B.Z., 2015. What is embodied about cognition? Language, cognition and neuroscience 30 (4), 420–429.

Mahon, B.Z., Caramazza, A., 2008. A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. J. Physiol. Paris 102 (1), 59–70.

Mahon, B.Z., Hickok, G., 2016. Arguments about the nature of concepts: symbols, embodiment, and beyond. Psychon. Bull. Rev. 1–18.

Mårtensson, F., Roll, M., Lindgren, M., Apt, P., Horne, M., 2014. Sensory-specific anomic aphasia following left occipital lesions: data from free oral descriptions of concrete word meanings. Neurocase 20 (2), 192–207.

Martin, A., Wiggs, C.L., Ungerleider, L.G., Haxby, J.V., 1996. Neural correlates of category-specific knowledge. Nature 379 (6566), 649–652.

Moll, J., de Oliveira-Souza, R., Passman, L.J., Cunha, F.C., Souza-Lima, F., Andreiuolo, P. A., 2000. Functional MRI correlates of real and imagined tool-use pantomimes. Neurology 54 (6), 1331–1336.

Moseley, R., Carota, F., Hauk, O., Mohr, B., Pulvermüller, F., 2012. A role for the motor system in binding abstract emotional meaning. Cerebr. Cortex 22 (7), 1634–1647.

Moseley, R.L., Mohr, B., Lombardo, M.V., Baron-Cohen, S., Hauk, O., Pulvermüller, F., 2013. Brain and behavioral correlates of action semantic deficits in autism. Front. Hum. Neurosci. 7, 725.

Negri, G.A., Rumiati, R.I., Zadini, A., Ukmar, M., Mahon, B.Z., Caramazza, A., 2007. What is the role of motor simulation in action and object recognition? *Evidence from apraxia.* Cogn. Neuropsychol. 24 (8), 795–816.

Neininger, B., Pulvermüller, F., 2001. The right hemisphere's role in action word processing: a double case study. Neurocase 7 (4), 303–317.

Neininger, B., Pulvermüller, F., 2003. Word-category specific deficits after lesions in the right hemisphere. Neuropsychologia 41 (1), 53–70.

Nichols, T.E., Holmes, A.P., 2002. Nonparametric permutation tests for functional neuroimaging: a primer with examples. Hum. Brain Mapp. 15 (1), 1–25.

Ohgami, Y., Matsuo, K., Uchida, N., Nakai, T., 2004. An fMRI study of tool-use gestures: body part as object and pantomime. Neuroreport 15 (12), 1903–1906.

Oldfield, R C, 1971. The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia 9 (1), 97–113.

Pulvermüller, F., 1999. Words in the brain's language. Behav. Brain Sci. 22 (2), 253–279.

Pulvermüller, F., 2002. The neuroscience of language: On brain circuits of words and serial order. Cambridge University Press.

Pulvermüller, F., 2005. Brain mechanisms linking language and action. Nat. Rev. Neurosci. 6 (7), 576.

Pulvermüller, F., 2013. Semantic embodiment, disembodiment or misembodiment? In search of meaning in modules and neuron circuits. Brain Lang. 127 (1), 86–103.

Pulvermüller, F., 2018. Neural reuse of action perception circuits for language, concepts and communication. Prog. Neurobiol. 160, 1–44.

Pulvermüller, F., Fadiga, L., 2010. Active perception: sensorimotor circuits as a cortical basis for language. Nat. Rev. Neurosci. 11 (5), 351–360.

Pulvermüller, F, Härle, M, Hummel, F, 2001. Walking or talking?: Behavioral and neurophysiological correlates of action verb processing. Brain Lang 78 (1), 143–168.

Pulvermüller, F., Hauk, O., 2005. Category-specific conceptual processing of color and form in left fronto-temporal cortex. Cerebr. Cortex 16 (8), 1193–1201.

Pulvermüller, F., Hauk, O., Nikulin, V.V., Ilmoniemi, R.J., 2005a. Functional links between motor and language systems. Eur. J. Neurosci. 21 (3), 793–797.

Pulvermüller, F., Shtyrov, Y., Ilmoniemi, R., 2005b. Brain signatures of meaning access in action word recognition. J. Cognit. Neurosci. 17 (6), 884–892.

Pulvermüller, F., Cooper-Pye, E., Dine, C., Hauk, O., Nestor, P.J., Patterson, K., 2010. The word processing deficit in semantic dementia: all categories are equal, but some categories are more equal than others. J. Cognit. Neurosci. 22 (9), 2027–2041.

Reynaud, E., Lesourd, M., Navarro, J., Osiurak, F., 2016. On the neurocognitive origins of human tool use: a critical review of neuroimaging data. Neurosci. Biobehav. Rev. 64, 421–437.

Riccardi, N., Yourganov, G., Rorden, C., Fridriksson, J., Desai, R., 2019. Degradation of praxis brain networks and impaired comprehension of manipulable nouns in stroke. J. Cognit. Neurosci. https://doi.org/10.1162/jocn_a_0149. Advanced online publication.

Rojkova, K., Volle, E., Urbanski, M., Humbert, F., Dell'Acqua, F., De Schotten, M.T., 2016. Atlasing the frontal lobe connections and their variability due to age and education: a spherical deconvolution tractography study. Brain Struct. Funct. 221 (3), 1751–1766.

Rorden, C., Karnath, H.O., Bonilha, L., 2007. Improving lesion-symptom mapping. J. Cognit. Neurosci. 19 (7), 1081–1088.

Rorden, C., Fridriksson, J., Karnath, H.O., 2009. An evaluation of traditional and novel tools for lesion behavior mapping. Neuroimage 44 (4), 1355–1362.

Rorden, C., Bonilha, L., Fridriksson, J., Bender, B., Karnath, H.O., 2012. Age-specific CT and MRI templates for spatial normalization. Neuroimage 61 (4), 957–965.

Rosci, C., Chiesa, V., Laiacona, M., Capitani, E., 2003. Apraxia is not associated to a disproportionate naming impairment for manipulable objects. Brain Cognit. 53 (2), 412–415.

Rumiati, R.I., Zanini, S., Vorano, L., Shallice, T., 2001. A form of ideational apraxia as a delective deficit of contention scheduling. Cogn. Neuropsychol. 18 (7), 617–642.

Shallice, T., Skrap, M., 2011. Localisation through operation for brain tumour: a reply to Karnath and Steinbach. Cortex 47 (8), 1007–1009.

Shebani, Z., Patterson, K., Nestor, P.J., Diaz-de-Grenu, L.Z., Dawson, K., Pulvermueller, F., 2017. Semantic word category processing in semantic dementia and posterior cortical atrophy. Cortex 93, 92–106.

Shtyrov, Y., Butorina, A., Nikolaeva, A., Stroganova, T., 2014. Automatic ultrarapid activation and inhibition of cortical motor systems in spoken word comprehension. Proc. Natl. Acad. Sci. Unit. States Am. 111 (18), E1918–E1923.

Strijkers, K., Costa, A., 2016. The cortical dynamics of speaking: present shortcomings and future avenues. Language, Cognition and Neuroscience 31 (4), 484–503.

Tomasello, R., Garagnani, M., Wennekers, T., Pulvermüller, F., 2017. Brain connections of words, perceptions and actions: a neurobiological model of spatio-temporal semantic activation in the human cortex. Neuropsychologia 98, 111–129.

Tomasello, R., Garagnani, M., Wennekers, T., Pulvermüller, F., 2018. A neurobiologically constrained cortex model of semantic grounding with spiking neurons and brain-like connectivity. Front. Comput. Neurosci. 12, 88.

Tomasello, R., Wennekers, T., Garagnani, M., Pulvermüller, F., 2019. Visual cortex recruitment during language processing in blind individuals is explained by Hebbian learning. Sci. Rep. 9 (1), 3579.

Tranel, D., Kemmerer, D., Adolphs, R., Damasio, H., Damasio, A.R., 2003. Neural correlates of conceptual knowledge for actions. Cogn. Neuropsychol. 20 (3–6), 409–432.

Vannuscorps, G., Andres, M., Pillon, A., 2014. Is motor knowledge part and parcel of the concepts of manipulable artifacts? Clues from a case of upper limb aplasia. Brain Cognit. 84 (1), 132–140.

Vannuscorps, G., Dricot, L., Pillon, A., 2016. Persistent sparing of action conceptual processing in spite of increasing disorders of action production: a case against motor embodiment of action concepts. Cogn. Neuropsychol. 33 (3–4), 191–219.

Vukovic, N., Feurra, M., Shpektor, A., Myachykov, A., Shtyrov, Y., 2017. Primary motor cortex functionally contributes to language comprehension: an online rTMS study. Neuropsychologia 96, 222–229.

Willems, R.M., Labruna, L., D'Esposito, M., Ivry, R., Casasanto, D., 2011. A functional role for the motor system in language understanding: evidence from theta-burst transcranial magnetic stimulation. Psychol. Sci. 22 (7), 849–854.

Yee, E., Thompson-Schill, S.L., 2016. Putting concepts into context. Psychon. Bull. Rev. 23 (4), 1015–1027.