Clinical functional MRI for language mapping: Best Practice recommendations from the Organization for Human Brain Mapping

Authors

Natalie L. Voets[1], Manzar Ashtari[21], Christian F. Beckmann[24], Christopher F. Benjamin[26], Tammie Benzinger[33], Jeffrey R. Binder[4], Alberto Bizzi[18], Bruce Bjornson[5], Edward F. Chang[14], Linda Douw[7], Jodie Gawryluk[8], Karsten Geletneky[22], Matthew F. Glasser[15], Sven Haller[17], Mark Jenkinson[1,28,29], Jorge Jovicich[30], Eric Leuthardt[32], Asim Mian[9], Thomas E. Nichols[25,1], Oiwi Parker Jones[19], Cyril Pernet[11], Puneet Plaha[16], Monika Połczyńska[12], Cathy J. Price[20], Geert-Jan Rutten[31], Michael Scheel[23], Joshua S. Shimony[33], Joanna Sierpowska[27], Lynne J. Williams[5], Ghoufran Talib[5], Michael Zeineh[13], Andreas Bartsch*[2], Susan Bookheimer*[3]

* contributed equally

Corresponding author: natalie.voets@ndcn.ox.ac.uk

Affiliations:

[1] Wellcome Centre for Integrative Neuroimaging (FMRIB), University of Oxford, Oxford, UK

[2] Department of Neuroradiology, University of Heidelberg, Heidelberg, and Radiologie Bamberg, Bamberg, Germany

[3] Department of Psychiatry and Biobehavioral Sciences, UCLA School of Medicine, University of California, Los Angeles, CA, USA

[4] Departments of Neurology and Biophysics, Medical College of Wisconsin, WI, USA

[5] BC Children's Hospital, Department of Pediatrics (Neurology), University of British Columbia, Vancouver, Canada

[7] Department of Anatomy and Neurosciences, Amsterdam University Medical Centers, The Netherlands

[8] Department of Psychology, University of Victoria, British Columbia, Canada

[9] Boston University School of Medicine, MA, USA

[11] Neurobiology Research Unit, Rigshospitalet, Copenhagen, Denmark

[12] Department of Psychiatry and Biobehavioural Sciences, University of California, Los Angeles, CA, USA

[13] Department of Radiology, Stanford University, CA, USA

[14] Department of Neurological Surgery, University of California, San Francisco, CA, USA

[15] Departments of Radiology and Neuroscience, Washington University School of Medicine, St Louis, MO, USA

[16] Nuffield Department of Neurosurgery, University of Oxford, Oxford, UK

[17] Centre D'Imagerie Medicale de Cornavin and Faculty of Medicine, University of Geneva, Geneva, Switzerland

[18] Neuroradiology Unit, Fondazione IRCCS Instituto Neurologico Carlo Besta, Milan, Italy

[19] Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford, UK

[20] Wellcome Centre for Human Neuroimaging, University College London, London, UK

[21] Departments of Ophthalmology and Radiology, University of Pennsylvania

[22] Department of Neurosurgery, University of Heidelberg, Heidelberg, and Klinikum Darmstadt, Academic Teaching Hospital of the Universities Frankfurt/Main and Heidelberg-Mannheim, Darmstadt, Germany

[23] Institute of Neuroradiology, Charité Universitätsmedizin Berlin, Berlin, Germany

[24] Department of Cognitive Neuroscience & Donders Institute, Radboud University, Nijmegen, The Netherlands

[25] Big Data Institute, Li Ka Shing Centre for Health Information and Discovery, Nuffield Department of Population Health, University of Oxford

[26] Yale School of Medicine, Connecticut USA

[27] Cognition and Brain Plasticity Unit & Neuroscience Institute, Dept. of Cognition, Development and Educational Psychology, University of Barcelona, Spain

[28] Australian Institute for Machine Learning (AIML), School of Computer and Mathematical Sciences, The University of Adelaide, Australia

[29] South Australian Health and Medical Research Institute (SAHMRI), Adelaide, Australia

[30] Center for Mind/Brain Sciences, University of Trento, Mattarello, Italy

[31] Department of Neurosurgery, Elisabeth-Tweesteden Hospital Tilburg, Tilburg, The Netherlands

[32] Department of Neurosurgery, Washington University School of Medicine, St Louis, MO, USA

[33] Mallinckrodt Institute of Radiology, Washington University School of Medicine, St Louis, MO, USA

Contents

Abstract

Ample reports highlight fMRI’s added value to guide and tailor neurosurgical interventions near brain regions supporting speech and language. However, fMRI’s usefulness for clinical language mapping remains controversial. This controversy is partly fueled by 1) differences between fMRI and the tools it is often compared against, and 2) wide heterogeneity in how clinical fMRI data are acquired, analyzed and interpreted. Both factors limit the objective assessment of the benefits and efficacy of presurgical fMRI.

The OHBM Working Group on clinical fMRI language mapping was formed in 2017. Its scope was to review and propose best practice recommendations addressing the specific challenges posed by applications in patient populations. The first objective was to consider language tasks, and task designs, optimized for specific clinical objectives, and incorporating modifications for patients with language- and broader cognitive impairment. The second objective was to put forward practical guidance, based on high-quality research, for each stage of the workflow from fMRI acquisition and analysis through to the reporting of individual patients' data.

In considering these challenges we focus on implementations that have proven practically feasible based on existing approaches (tasks, software packages) actively in use today. When widely available practices deviate from optimal practices, we highlight emerging developments that merit further evaluation and incorporation into routine clinical use to advance the current practices.

This document was created in collaboration with the OHBM Committee on Best Practices, incorporating community feedback. Its aims are to provide a framework for improved standardization of fMRI to enable much needed evaluations of its ultimate goals; namely, minimization of invasive intraoperative testing and, ultimately, of new post-operative language deficits. In this regard, the single strongest recommendation is for greater transparency and reporting of longitudinal outcomes in patients undergoing clinical fMRI.

Keywords: Functional MRI, language, neurosurgery, presurgical mapping, surgical planning

A high-level summary of the Working Group recommendations is presented in Box 1 for quick reference. Supporting data are detailed throughout the following document and appendices.

Summary working group recommendations for best practices in Clinical Language fMRI mapping — **• BOX 1 •**

Patient selection & preparation

See section 2.1 for further details.

- **Obtain medical history and prior radiological imaging** to identify suitable candidates (Fig 1), identify contraindications to MRI, and anticipate artifacts that may render fMRI futile or require sequence changes (Fig S1).
- **Obtain a minimum cognitive assessment (and a detailed neuropsychological assessment whenever feasible)** before fMRI to ensure patients have (at least) minimum residual abilities and to identify fMRI task adaptations needed to maximize successful performance. The presence and pattern of existing neurocognitive deficits is directly relevant for the selection of patients, choice and adaptation of fMRI tasks, and interpretation of results.
- **Always practice tasks** before fMRI to ensure the patient's full understanding and ability to perform the tasks as desired, including how and when to make any required responses.
- For **pediatric patients, a simulated session** (virtual or in person) is highly recommended.

Clinical goals of fMRI language mapping

See section 2.2 for further details.

- Within the multidisciplinary team, **clearly define the clinical question** to **select the appropriate combination of fMRI tasks** based on intended purpose and known structure-function relationships (Fig 4, Table 1).
- **In multilingual patients,** prioritize the patient's primary language but ideally involve each language needed in their everyday life (e.g., using a fMRI language **'switching' task**).

Design considerations

See section 2.3 for further details.

- **Adjust task parameters** (modality, speed of presentation, difficulty level, etc.) to patient's performance abilities (and/or developmental stage) as much as possible (Fig 6, Box 4).

- **Select language and control conditions** to match the clinical question and anatomy (Table 2).
- **Optimal task designs vary with the intended analysis**.

Acquisition

See section 3.2 for further details.

- **Use 3T whenever possible, a TR of 1-3 seconds and voxels no larger than 3.0mm isotropic**.
- Explain problems caused by head movement and **spend time on positioning and comfort**.
- **Monitor performance** in real-time (through behavioral responses / real-time activation maps) and provide feedback to the patient (e.g., over the intercom) after each task run.
- **Monitor head motion** throughout the task; stop and repeat if necessary.
- **Use multiple short fMRI runs** rather than one long run to help minimize motion and fatigue.
- **Acquire a fieldmap** to correct for fMRI image distortions (Fig 7)**.**

Pre-processing and analysis

See section 3.3 for further details.

- **Quality control each step**: Distortion-correct images and examine the need for and effect of any retrospective motion correction steps.
- **Minimize spatial smoothing to** 5mm (ideally 4) FWHM, as functional zones may be <4mm apart or occupy one bank of a sulcus (Fig 8).
- **Inspect any automated image segmentations** to ensure the area of surgical interest is not accidentally 'extracted' (removed from analysis).
- **Carefully inspect** (and, if needed, adjust) fMRI-to-structural **registrations.**
- **Consider potential delays in task performance, hemodynamic responses** or other sources of temporal variability of the fMRI signal (e.g., through supplementary and/or complementary temporal derivative or independent component analyses)
- Select statistical approaches to **minimize false negative findings**.
- **Inspect results at a range of statistical thresholds** (including unthresholded).

- **If calculating LIs**, **focus on language ROIs**, **inspect a range of thresholds, and** use LI calculation methods that combine a range of thresholds.

Interpretation and reporting

See section 3.4 for further details.

- **Describe data quality, what quality control was performed and any interpretative challenges** (e.g., emphasize areas where signal loss prevents detection of meaningful fMRI signals, especially around a lesion and around the expected location of critical cortex).
- **Images are powerful and potentially misleading**. If the reporting specialist believes the results are not valid, state this is the case and exclude images from the report.
- If intended for surgical planning, provide images showing proximity of lesions to critical areas.
- **Integrate complementary methods** (diffusion tractography, CVMR, perfusion, etc).

1. Introduction

The main clinical applications of functional MRI (fMRI) for individual patients include preoperative *risk assessment* and the mapping of "critical" cortex[a] for presurgical *planning* [1]. *Risk assessment* refers to determining whether surgery will take place in the language dominant hemisphere, which entails greater risk for postoperative language impairments. Consequently, establishing how ‘strongly’ the targeted hemisphere is dominant for language informs patient consent, deciding whether to proceed with surgery, and the surgical plan itself. Presurgical *planning* also includes deciding on a strategy to reach a surgical target while minimizing damage to surrounding functionally important brain tissue, i.e., functionally relevant gray matter structures and important white matter tracts. fMRI - if properly performed - allows the non-invasive visualization of gray matter functions at an individual patient level. But what constitutes ‘high-quality’ clinical fMRI?

The focus of most validation work on clinical fMRI has been mapping speech- and language-related functions, because their localization, with few exceptions (e.g., the ventral premotor cortex), cannot be predicted from anatomy (i.e., cortical folding patterns) alone. However, despite ample reports highlighting fMRI's value in guiding, selecting and tailoring neurosurgical interventions close to speech and language regions, the application of fMRI for clinical language mapping remains controversial [2,3]. While this controversy stems from several sources, one main challenge is the dramatic variation in the approaches used in clinical language fMRI and, likely as a direct result, variability in outcomes [4].

The biggest challenge for developing clinical guidelines to standardize practice is the large variety of tasks and methods [4–7] used to conduct, analyze, and interpret clinical fMRI studies. Clinical language fMRI is more-or-less *equal parts* a language neuroscience experiment, cognitive assessment, radiological exam, image analysis application, statistical determination, and neurosurgical decision-making tool. The requirement for this range of complementary skills has meant that as fMRI has moved into the clinic, the discipline has fragmented among a range of professions seeking to

[a] Critical, often called “eloquent”, which is, strictly speaking a mereological fallacy according to Bennet & Hacker [538], and also used in a much broader sense in the neurosurgical community [539].

characterize and standardize clinical fMRI from their unique perspectives [4,8], which may often mean a lack of interdisciplinary consensus.

In one early attempt at standardizing fMRI, neuropsychologists focused on the *skills* that valid clinical fMRI requires [9]. The main conclusion was that a multidisciplinary approach is necessary, involving professionals with expertise in critical domains including (among others) cognitive and computer science, psychology, neuroradiology, neurology, and neurosurgery. In 2017, the American Society for Functional Neuroradiology (ASFNR) published recommendations for the cognitive *tasks* used to map language functions in surgical patients. Perhaps in part due to the group's radiological focus, recommendations were based on a survey of tasks commonly in use at their members' institutions [10]. In another report, the American Academy of Neurology summarized the *quality of evidence* supporting fMRI for lateralization[b] of language [5]. When representative samples of individuals collecting and analyzing fMRI for epilepsy surgical programs were surveyed in the US [4] and the EU [7], essentially all aspects of the procedure were found to vary across institutions.

The fMRI *tasks* and *analysis* approaches used are the most important aspects of clinical language fMRI to standardize to ensure patient care. It may be of interest to know what tasks are in common use, but it is more critical to consider the *validity* of various practices. Several aspects of fMRI task design (e.g., sensory modality or control conditions) are important to explicitly consider due to their influence on the specificity of the activation patterns. In patients with specific existing language impairments, it may furthermore be necessary to alter task conditions, speed of presentation or difficulty levels, whereas ‘standard’ tasks have traditionally been ‘one-size-fits all’. From an *analysis* perspective, there is a lack of recommendations that address state-of-the-art data collection, preprocessing, analysis, and reporting. For example, in a 2018 clinical survey [4], the most used fMRI data smoothing kernel was 8mm, which happens to be the default setting in a prevalent data analysis package

[b] While it is generally interhemispheric *dominance* for language (and verbal memory) that is sought to be established prior to esp. epilepsy surgery, traditionally by (super-)selective Wada testing, it is the asymmetrical *lateralization* of activations that can be established by fMRI. Although fMRI cannot discriminate indispensable from expendable activations per se (and at the individual patient level), lateralization has been shown to be a useful surrogate to predict postsurgical outcome in sufficiently large patient samples.

often used in neuroscience studies [11]. Yet, default settings in research software are generally chosen with a different outcome objective in mind, namely to maximize *group-level* sensitivity by *minimizing inter-individual variation* in fMRI activation location and extent [12]. **In clinical settings, it is precisely the individual variability that we aim to capture because of its impact on surgical decision-making**. Such technical and methodological considerations have profound implications for the interpretability of clinical fMRI [2,3] and the spatial 'localizability' of findings [13].

Through efforts by the Organization for Human Brain Mapping (OHBM) to improve best practices for neuroscientific imaging data analysis and reproducibility, Best Practices in Data Analysis and Sharing (COBIDAS) have been recommended [14,15]. The OHBM Clinical fMRI Working Group was formed in 2017 in response to the specific challenges posed by clinical fMRI under the COBIDAS principles. This Working Group consists of a multidisciplinary committee spanning all domains of relevant expertise, including neurologists, neuroradiologists, neurosurgeons, imaging statisticians, neuroscientists, and neuropsychologists in both adult and pediatric populations. Bringing together practitioners from North America and Europe, the working group had two main goals:

1. To present the unique challenges of mapping language functions in clinical populations who frequently suffer from neurocognitive impairments; and
2. To put forward practical guidance for the acquisition, pre-processing, analysis, and reporting of individual patient data based on high quality research.

This document details the recommendations of the Clinical fMRI Working Group. Our goal is to offer a comprehensive workflow from patient characterization through to interpretation using evidence-based guidelines. In light of continued progress, especially with respect to technical aspects of fMRI data collection and analysis, it is hoped that these recommendations will improve minimum standards and facilitate meaningful use, enhance comparability between studies, and support large-scale evaluations of fMRI's strengths and limitations in clinical use.

2. Part 1. Clinical indications and approaches for mapping language using fMRI

Because of the complexity of the language system (section 2.2.2.), different considerations arise for tasks, control conditions and analyses depending on the referral question and patient characteristics. Consequently, considerations surrounding presurgical language fMRI begin by establishing the clinical question that fMRI is intended to answer. Additionally, adequate patient selection and characterization - including targeted neuropsychological testing - are crucial in choosing optimal language fMRI paradigms.

2.1. Inclusion / exclusion criteria and additional information required prior to fMRI

Patient selection. Functional MRI is primarily indicated for surgical targets *within* the brain parenchyma, i.e., intra-axially located, *and* where there is a risk of surgery causing language impairment. This risk may arise because the lesion is within or near functional cortex or associated white matter pathways. Alternatively, this risk may exist when the surgical trajectory, e.g., a trans-opercular approach to the insula, could disrupt language networks. In some selected extra-axial surgeries (for example resections of space occupying meningiomas, arteriovenous malformations (AVMs) with a nidus outside the brain parenchyma or aneurysm clipping), language fMRI mapping can also be helpful to inform surgical access and/or the temporal order of multi-stage embolization in relation to language territories at risk of hemorrhagic or ischemic damage [16].

Handedness. Numerous factors influence the risk that neurosurgery poses to language functions. One such factor is the patient's handedness. The overall likelihood of atypical (i.e., 'bilateral' or right-hemisphere) language dominance is higher in non-right-handed (left-handed or ambidextrous) individuals (22-30%) than in right-handed people (4-12%) (e.g., [17–19]), yet the majority of non-right-handed individuals have typical left-hemisphere dominance for speech and language (e.g., [20]). Consequently, handedness alone is useful to *inform*, but not sufficient to *rule out* surgical risks.

Side of surgery. The overall probability of right-hemisphere language dominance, even in left-handed and ambidextrous patients is low. Consequently, some clinicians view it as not strictly necessary to perform fMRI prior to surgery involving the right hemisphere. However, the incidence of atypical (bilateral or right-hemisphere) language dominance is higher in people with chronic epilepsy [21,22] and in some individuals with a long-standing or slow-growing left hemisphere lesion (e.g., [23–25]). Clearly, the consequences to the patient of removing potentially language-critical structures in the right hemisphere may be severe. A case can therefore be made for pre-surgical language fMRI if a surgical target or intended access route approaches possible language structures ***in either hemisphere*** [26].

Cognitive status. Developments in resting state fMRI offer future possibilities to understand the organization of functional networks without active patient participation [27–32] (see Appendix A). However, at present, identifying brain regions involved in *specific* aspects of language requires task-based mapping through the active participation of a motivated patient. Presurgical task-based language mapping is therefore not normally indicated in patients who are unable to cooperate, those unable to tolerate MRI, or in very young children (typically under the age of 6 years) who might need sedation to undergo MRI.

Knowledge of the detailed neurological and cognitive performance status of each patient is essential when planning language fMRI. Use of pre- and post-operative cognitive assessment batteries is well established for some patient groups, such as temporal lobe epilepsy. In these patients, neuropsychological test results predict whether dominant left temporal patients are at low, medium, or high risk of post-surgical naming decline [33]. In other populations, such as high-grade glioma, pre-operative performance is not routinely captured. However, existing impairments may prevent patients from completing a meaningful fMRI study. There is little benefit to be gained, for example, by attempting fMRI language mapping when the patient is too aphasic to comprehend instructions, or task(s) are too difficult for them (see 2.3.6). Neuropsychological assessment can detect **subtle deficits** that are not apparent conversationally, but are nevertheless **specific and predictive for surgical risks**, which informs optimal fMRI task selection. Likewise, comprehensive baseline performance assessments identify patients who cannot complete specific tasks due,

for example, to dyslexia or impaired auditory comprehension. In such cases, fMRI tasks may need to be modified (e.g., by changing the mode of delivery, or changing the speed at which fMRI should be presented, see section 2.3) to enable a successful scan. Finally, certain pathologies may not directly impair language skills but instead cause **wider attention or memory difficulties**. Such difficulties may arise from a lesion impacting on cognitive systems. However, it is also important to consider deviations from age-typical language, cognitive, and motor milestones (whether precocious or delayed) in pediatric patients (section 2.3.7), or in patients with early onset drug resistant epilepsy. Deficits in these domains can impact on the general ability to perform cognitively demanding fMRI tasks and **can directly affect fMRI results**. For example, activation of the right inferior frontal cortex during a language task could reflect essential language functions [34], but could alternatively signal heightened effort not directly related to linguistic processing [35].

For all these reasons, **best practice is to appropriately evaluate neuropsychological performance both before fMRI and surgery, as well as post-operatively** to evaluate outcomes. While we are very sensitive to constraints on clinical time and resources, short bedside testing of primary cognitive and language domains, along with some more targeted in-depth assessment of deficits is generally feasible in virtually every case.

Contraindications. Contraindications include general MRI safety factors (e.g., certain implanted devices, ferromagnetic objects, etc). Vagal nerve stimulators are common in the epilepsy population, and generally necessitate using a transmit-receive coil [36]. Additional confounds that can limit the execution or interpretability of fMRI include ischemia, necrosis, hemorrhage, calcifications, certain tumors, flow void/shunting and/or metallic artifacts from previous surgeries [37–39] (see Appendix B). Their presence is therefore important to establish from prior imaging before attempting fMRI. Claustrophobia is a potential obstacle for fMRI, but incidence rates on modern scanners are very low (<1% [40]). Anxiolytics (e.g., benzodiazepines) for claustrophobia are generally to be avoided, as these may lower task performance and neuronal responses. Very low dose anxiolytics could be cautiously considered if absolutely necessary [40,41].

2.1.1. Working Group recommendations on candidacy

Summary recommendations are displayed in Fig 1. As a minimum, potential **patients should be conscious, able to communicate and tolerate being in a confined space with loud noise while performing fMRI tasks.** They must have **minimum abilities to perform required tasks and** be able **to simultaneously avoid certain behaviors** (e.g., moving or talking out loud during the task if silent responses are required). Detailed description of the fMRI process during appointment planning is therefore recommended. For adults, preparation could consist of online training materials (e.g., stanfordhealthcare.org/fmri) [42], supplemented with essential in-person practice and verification that the subject understands and can perform the task requirements**. For pediatric patients, virtual and in-person MRI simulator sessions** (e.g., https://www.bcchr.ca/3tmri/facilities/about-simulator) **are recommended where possible** to maximize success. These sessions a) familiarize children with performing task(s) supine in a noisy, confining environment and b) verify their ability to perform the task(s) prior to the actual fMRI exam. Children should be able to complete the simulator session successfully to undergo clinical fMRI examination. Multiple simulator training sessions may be necessary to ensure the actual clinical scan success, especially with younger or neurodevelopmentally delayed patients.

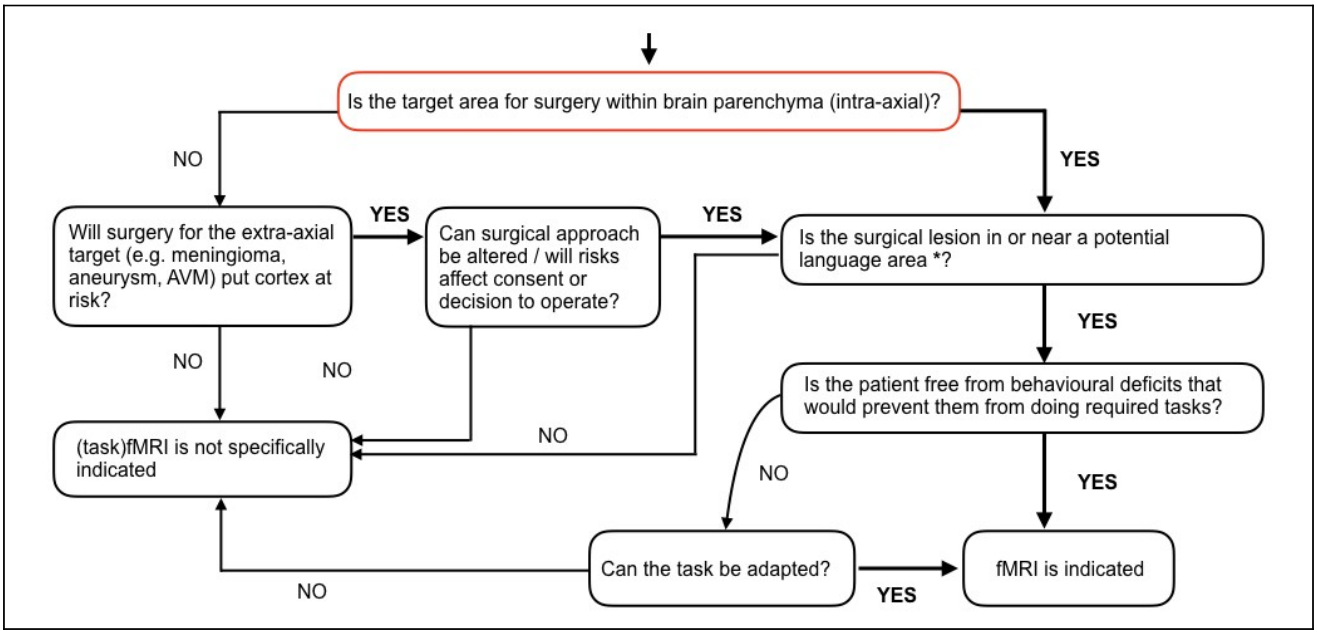


Fig 1. Inclusion / exclusion flowchart for presurgical language fMRI. *Legend. AVM = arteriovenous malformation. * Please see section 2.2.2 for further details describing localization of language regions.*

2.2 Clinical objectives of presurgical language mapping by fMRI

Objectives of speech and language fMRI mapping generally separate into *inter-hemispheric lateralization* versus *localization of functional tissue*. Tasks optimized to *lateralize* language can be less suited to *localizing* functions, and vice versa. **Consequently, precision in formulating the clinical request and surgical goals helps with tailoring tasks to maximize the utility of fMRI**.

2.2.1 Lateralizing language functions

Several studies show that the risk of developing postoperative language and verbal memory deficits is partly dependent on language lateralization [33, 5]. Therefore, one important application of language fMRI is determining language lateralization prior to surgery.

Language lateralization using fMRI has been most extensively assessed prior to temporal lobe surgery to treat drug-resistant epilepsy [22,43]. The clinical purpose is usually to establish *graded language lateralization* as an alternative to invasive Wada testing [44]. Laterality results are often used to inform consent discussions and/ or tailor the surgical approach / extent (e.g., selection of laser ablation versus anterior temporal lobectomy) in situations where surgery is considered elective. Consequently, in addition to its use prior to temporal lobectomy, laterality mapping can also be of clinical interest prior to other elective or semi-elective procedures. Examples of the latter include epilepsy surgery outside the temporal lobe, hemispherectomy and microsurgical or embolization treatment of AVMs [45]. Laterality can also be useful to (re-)establish in patients who have undergone previous surgery, in whom the inter-hemispheric localization of language can change as a result of functional reorganization (reviewed in [23]).

Considerations around how to calculate LI are detailed later (section 3.3.8). In terms of the *task approach*, different tasks engage different parts of the language system to variable extents (see section 2.2.2). For example, word *production* tasks reliably engage word selection and verbal working memory processes concentrated in the frontal cortex, while language *comprehension* tasks engage these frontal executive networks as well as sub-regions of the temporal and parietal lobes (see Table 1 and the references therein). The prevalent division of tasks into those probing "receptive" versus "expressive" aspects of speech and language is inherently problematic, since nearly all language tasks engage core phonological, semantic, and verbal executive (selection and working memory) processes (see [46–48] for comprehensive reviews). Within the same patient, the left and the right hemispheres can both participate in and rarely even be "dominant" for different aspects of language (so-called "crossed dominance") [49]. **Consequently, a combination of task contrasts assessing an array of language functions are generally advocated to establish laterality** at the lobe-level [49–53] (Box 2, see also section 2.3.2).

Epilepsy surgery centers typically assess language lateralization with at least two tasks (88-96% of epilepsy programs), and several routinely use at least 3 (24.5%-75%) or more (14-55%) tasks [4,7]. While 91-92% of glioma-specialized centers in two surveys used fMRI to lateralize language functions, details of their technical approaches were not provided [54,55] . However, fMRI scans were reported to take at least 15-30 minutes in 45% of respondents in the survey by Thust [55] (of which almost a quarter (23.5%) take > 30 minutes), implying the use of multiple tasks.

In addition to the type of language task, laterality results are strongly influenced by specific paradigm choices in terms of the task *baseline*, i.e., what the language task is compared against [44,56,57]. For example, comprehension tasks presented in the auditory modality offer poor lateralization if (the strongly bilateral [58]) acoustic processing is not accounted for. Similarly, multiple studies emphasize the importance of controlling for nonlinguistic aspects of language processing (e.g. using non-language visual or auditory decision tasks) to lateralize semantic processing in the temporal lobe [56,59] .

Establishing regional laterality therefore requires careful consideration of both the active and control conditions of fMRI tasks (section 2.3.3).

General approaches to establish language laterality are presented in Box 2. Of clinical note, establishing lateralized representations of language functions does not exclude the possibility that an area of detected activation in the "non-dominant" hemisphere may be indispensable to language in an individual [60]. **The "size" of an area of activation as detected by fMRI strongly depends on statistical analysis and thresholding of the data** (section 3.3) and does not, *per se*, reflect its functional role. Consequently, surgically-oriented applications of fMRI are more often concerned not only with establishing *which hemisphere* harbors language functions, but also with identifying *where they are located* and *what* language functions are most at risk.

Considerations: tasks for lateralizing language functions	• BOX 2 •

Most epilepsy (91-92%) [7,61] and many brain tumor (53-100%) [54,55] centers use fMRI for language lateralization.

Tasks to assess language lateralization:

When considering recommendations for fMRI tasks, many tasks have been used and individually compared to the Wada test. Numerous reviews and meta-analyses emphasize high overall correspondence in language laterality derived from task fMRI compared to Wada [5,43,62,63]. However, each review emphasized extensive variability in fMRI practices. A recent Delphi survey highlighted that different tasks for lateralization may be suitable for different questions and optimal fMRI tasks are still not established [53]. Given the limited data to quantify clinical levels of evidence for any single task, our Working Group approach to creating recommendations was informed by:

a) Dedicated review of neuroscientific studies engaging selected language processes, and how lateralized those processes are (Fig 2 / Table 2, section 2.2.2.1),

b) A published systematic review assessing the lateralization value of different fMRI tasks [57],

c) Published data comparing laterality from fMRI tasks to the Wada test (but see section 4.3. for important limitations with this approach)

The following recommendations are not intended to be prescriptive.

Rather, our aim is to highlight a combination of tasks *most likely* to provide high sensitivity for language lateralization based on current data and our working understanding of the basis of language in the brain.

Recommendation 1. Establishing language lateralization requires a combination of tasks assessing complementary and dissociable language systems. Different language processes both engage different networks of brain regions, and are lateralized to different extents (Fig 2). Our knowledge of these systems is continually evolving and optimal language task development is an area of active research. We advocate adopting established tasks but also further evaluating newly-emerging tasks. It is recommended to **include tasks that, at minimum, assess language production, language comprehension and semantic processing**. Laterality established through a combined analysis of several fMRI tasks has also shown better correspondence to Wada laterality than laterality based on individual tasks analyzed on their own [64].

Recommendation 2: Most everyday 'language' requires a combination of phonological, lexico-semantic and syntactic processing (in addition to non-linguistic operations, such as cognitive control and working memory). **Using at least one sentence-level task is therefore strongly recommended.**

When considering approaches to language lateralization with fMRI, **a historical tendency has been to categorize tasks into those probing the language 'production network'** (aka 'expressive' tasks**), 'comprehension network'** (aka 'receptive' tasks) **and 'semantic network'. Since most fMRI tasks engage several (if not all) of these networks to varying degrees, there are practical and theoretical limits to categorizing tasks as one or the other** [46–48]. Furthermore, it is the **task contrast** to a specific control or baseline condition, that enables to study specific language processes (cf. section 2.3.3.2).

Some specific task approaches to consider include:

1. **Language production***: Generating or producing language, aside from the articulatory system, involves a network of brain regions including the inferior frontal, supramarginal and superior temporal gyri. A systematic review identified **word generation tasks** as **strongly lateralizing** [57], especially when using an appropriate 'active control' (see section 2.3.3.2). Among production studies comparing laterality derived from fMRI versus Wada, letter fluency (9 studies), category fluency (5 studies) and verb generation tasks (5 studies) have been the most evaluated (see [43,62,63]). Letter fluency engages the frontal lobe more than category fluency, which instead engages the temporal lobe more (e.g., [65], noting also [66]). The choice of preferred task is therefore also influenced by the surgical target and by the way laterality will be calculated (see section 3.3.8). Notably, verb generation tasks generally present more cues during each task block (e.g., one item every few seconds) than other verbal fluency tasks (one item per 20 or 30 second block), and patients can have restricted fluency for reasons other than language impairment. Consequently, verb generation tasks are more likely to sustain

patient performance and, thereby, fMRI activation, than fluency tasks. Word generation tasks produce activation that is more confined to the frontal lobes than sentence generation tasks (which additionally engage wider network regions including the temporal lobes [67], and are therefore useful to consider when lateralization is required in the context of a surgical target in the frontal lobe.

2. **Language (particularly sentence) comprehension***: Understanding the meaning of language engages a widespread system including the lateral and ventral temporal lobe and angular gyrus in addition to the inferior frontal cortex [68–70]. Comprehension has been assessed in the visual and auditory domains and using a variety of tasks. A systematic review identified **sentence comprehension tasks,** in particular, to be **strongly lateralizing** [57]. Relatively fewer studies have compared laterality from fMRI tasks assessing sentence comprehension with the Wada test. Three of 5 studies in meta-analyses used visually presented sentences with either reading or a semantic/syntactic decision (both compared to a perceptual control). Auditory versions (2 of 5 studies) generally focused on story listening compared to reversed speech, and tended to emphasize passive listening [63]. Comprehension can also be assessed using a sentence completion task, in which the participant generates a word to complete a sentence stem, or responsive naming**,** in which a phrase is heard or read (e.g., “long yellow fruit” and interpreted to generate a name (‘banana’)), thus combining language ‘comprehension’ and ‘production’. Such tasks are therefore useful to include to establish lateralization of the wider temporo-parietal (and temporo-parieto-frontal) networks. Simple sentence processing such as reading or passive listening are reported to be less lateralizing than sentence comprehension tasks involving a semantic or syntactic decision [57]. These (simple sentence processing tasks) are therefore not recommended for language lateralization, at least when used alone and in standard GLM-type analyses or unless limited patient abilities require it (see below).

3. Tasks requiring a **semantic decision**, when paired with an active non-linguistic control task, engage additional regions of the anterior and inferior temporal cortex compared to word production and passive sentence listening tasks. Since the use-case for language lateralization is often temporal lobectomy, a word comprehension task involving a semantic decision is important to engage the anterior and basal temporal cortices. Various tasks have been developed. The most extensively studied approach contrasts a semantic decision condition, in which the participant decides if an aurally presented word fits a category (e.g., “found in the USA” and “used by people”), with a nonlinguistic tone perception condition [56]. In an alternative approach, pairs of words are presented and patients decide if the words are related (e.g., synonyms). In the visual domain, visually cued naming with action generation (e.g., pen = ‘writing’) can elicit semantic decisions on single words [57]. Most semantic fMRI studies evaluated against Wada have been single word decision studies presented either visually or aurally (5 studies,

see [63]). In a systematic review of fMRI tasks [57], semantic tasks involving a relatedness decision (e.g., word pairs) were more strongly lateralizing than single word decisions. However, the strength of lateralization of semantic tasks is strongly influenced by choices in the fMRI baseline condition, which should be both active and involve a perceptual decision to best isolate linguistic semantic processing (see Fig 5).

Modifications for language lateralization in patients with specific impairments

- **Visual impairments**: Some people have better auditory than visual attention, and vice versa, which may affect whether they find listening comprehension tasks easier or harder than reading comprehension. Importantly, some patients are **blind** or have visual field deficits. When technically feasible (see Appendix C), tasks using the auditory modality are therefore crucial for some patients, and beneficial (though perhaps not essential) for others (e.g., those with stronger auditory attention) [58,59,71]. Common auditory tasks such as auditory sentence completion or auditory responsive naming (the auditory complement to visual responsive naming) generally activate bilateral temporal regions [72,73]. These tasks are therefore not strongly lateralizing unless auditory processing is controlled for using a relevant control condition (e.g., tone discrimination, see Box 3), or through conjunction analyses (e.g., with visual versions of the same task) to identify common (modality independent) areas of activation [61].

- **If sentence level processing is impaired**, one alternative approach is to select word-level tasks‡. Letter fluency is commonly used but is strongly driven by attention and executive functioning, which are both often impaired in patients with focal lesions [74] (often independently of language). Alternatives such as verb generation or single word comprehension (semantic decision) use frequent external cues, which facilitates performance in patients with selective restrictions in executive functioning. Another approach is to adopt passive story listening / sentence reading, depending on the nature of the patient's difficulties, with the caveat that these are less strongly lateralizing in general [57].

- **In patients with a/dyslexia**, object naming can engage widely distributed lexico-semantic and phonological processing. However, lateralization in naming tasks is variable, depending on the task state that picture naming is contrasted against [57]. It is essential therefore to include a perceptual control (e.g., involving a visual detection/decision on complex un-nameable images) to isolate semantic and phonological components of the naming task.

- **In patients with anomia:** Some studies [75,76] have evaluated tasks using image pairs (e.g., an object and a separate written word) for patients to generate short phrases. Further research is needed to evaluate if such "sentence generation" tasks are easier for patients with naming impairments to perform, but one recent study showed this approach [77] may offer better lateralization than single object naming.

Notes

* Of emerging interest (although likely reserved for cognitively high performing patients) are grammar tasks. Tasks assessing syntax or morphology engage brain regions involved in hierarchical sentence-level operations (e.g., posterior middle temporal gyrus [78,79]), potentially more robustly than simpler lexico-semantic tasks [51].

‡ Visual rhyming was recommended for lateralization by the ASFNR [6] but supporting evidence varies [57]. Nonetheless, in patients unable to perform other tasks, or in pediatric populations, rhyming is easier for patients to perform and could be considered.

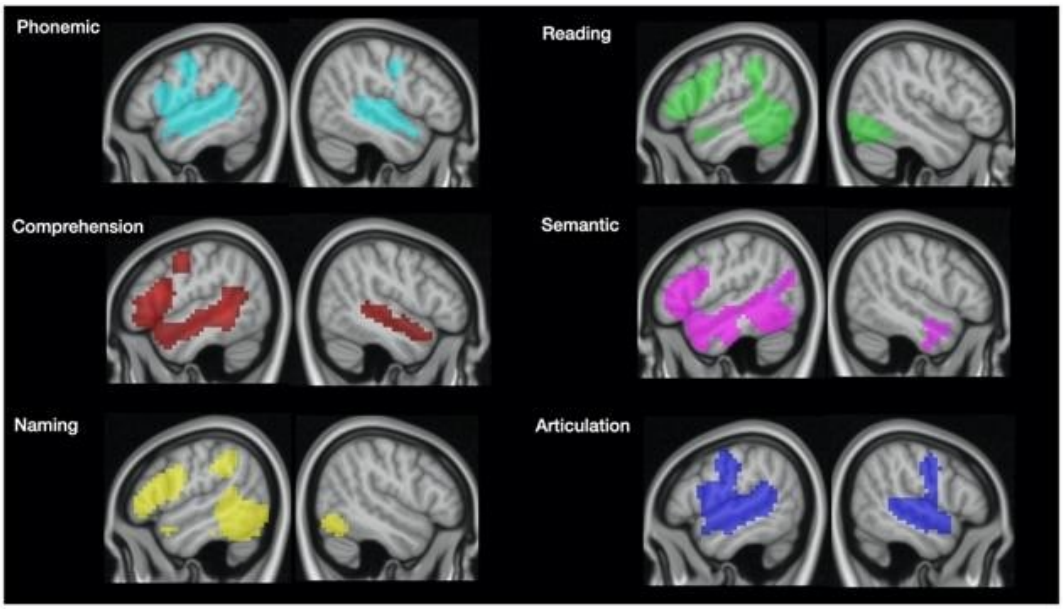


Fig 2. Relative lateralization of language processes based on fMRI. Legend. Results of multivariate prediction based on meta-analysis of approximately 13,500 neuroimaging studies in Neuroquery (http://neuroquery.org), querying the search terms “phonemic”, “comprehension”, “naming”, “reading”, “semantic” and “articulation”. The neuroquery-derived z-score maps are shown thresholded at z=3.1 (corresponding to $p < 0.01$). Different language processes are lateralized to different extents, depending on the language process engaged during a given fMRI task, and – importantly, what control condition the target language process is compared against (see Fig 5).

2.2.2 Localizing language: functions of concern in specific brain areas

Functional MRI efforts to localize language centers in relation to a focal lesion are intended to:

i) inform discussions and consenting of patients

ii) provide a risk map for surgeons to help plan the surgical approach,

iii) provide an estimate of likely resectability / achievable resection extents,

iv) predict the specific risks associated with the surgery and/or

v) guide the need for / use of Wada or intraoperative ESM.

Some 44% of epilepsy centers already cautiously use fMRI to guide surgical margins [61]. Among surveyed neuro-oncology centers, almost 80% of respondents report using fMRI to localize language functions [55]. 88% requested fMRI to plan the extent of resection [54], although not as the sole technique for this purpose, at least in the latter survey. In one survey of European Glioma network centers, 14% of respondents preferentially relied on fMRI for functional (not specifically language) localization [80] even over and above intraoperative assessment. The suitability and limits of fMRI used surgically in this way are presently still difficult to evaluate. There is limited published data systematically evaluating language outcomes. Additionally, there is variability in how language is assessed both with fMRI and intra/peri-operatively. What is clear, however, is that clinical applications of fMRI to help decide ‘is this area safe to ‘cut’?’ generally require several carefully controlled tasks to isolate different aspects of critical language processing according to the structures most at risk in a given surgery.

Optimal fMRI tasks and task designs for language ‘localization’ remain an open question. In selecting an approach, recent reviews of widely used tasks can be found in [4,6,57,81,82]. Users could also consult large-scale neuroimaging database resources such as Neurosynth (https://neurosynth.org/) [83] or its recent extension Neuroquery (https://neuroquery.org/) [84] to observe the typical activation patterns of specific language tasks / processes (e.g., Figs 2, 3). Ultimately, however, fMRI task selection should draw on knowledge of the role of specific brain regions in essential aspects of language processing [46,47,51,85], considered next.

2.2.2.1 Choosing tasks to target linguistic 'processes'

Core elements of the 'language' network(s) are now well established. Data supporting a role for certain structures in aspects of language have been widely replicated over the past 30 years [47]. Based on this understanding, questions surrounding language localization predominantly arise in the context of surgery involving:

a) the inferior frontal gyrus (IFG),
b) parts of the (especially posterior) middle frontal gyrus (MFG),
c) posterior middle and superior temporal gyrus (MTG/pSTG) and sulcus (pSTS),
d) inferior parietal lobule (IPL) and
e) the mid-fusiform gyrus / basal occipitotemporal cortex.

Language-related deficits are also associated with surgery involving the supplementary motor area (SMA) and pre-SMA, but are often transient if the contralateral homologue can support this function [86] and if the corpus callosum remains intact.

Additional brain regions contribute to language in ways that remain incompletely understood and are therefore challenging to 'map'. One such region is the temporal pole. The temporal pole, in addition to naming [87], is proposed to contribute to semantic memory, either as an amodal repository for conceptual (semantic) information [69] or as a multi-modal convergence zone that reactivates conceptual knowledge [87]. Data concerning semantic memory outcomes after unilateral anterior temporal lobectomy in the language dominant hemisphere, however, are inconsistent. Many patients show preserved semantic knowledge (e.g. "oh, in Egypt… where they bury kings") even for test items that they can no longer name (i.e., pyramid) [88]. Differences in the types of tests used across studies and the definitions of what constitutes 'semantic memory' both likely contribute to current uncertainty around the critical role of the language dominant temporal pole. Similarly, the precise functions of both the Graphemic Motor Frontal Area (aka 'Exner's area') [89] and Hopf area 55b [90,91] in the MFG remain to be fully understood, though there is evidence that both are involved in reading and writing [92].

Beyond identifying the relevant cortical areas, a more difficult – yet crucial – challenge is defining the role these areas play in language function. In order to help preserve

function after surgery, the aim of clinical fMRI is not to just activate a *specific brain region,* but – instead – to identify the *network* or assembly of brain regions engaged during a *specific language process*. This distinction is important for two reasons.

Firstly, several **neuronal populations contribute to more than one function**. Some populations are engaged by a wide range of cognitive tasks [93] and/or integrate converging information [94,95], while others seem to 'switch' roles depending on task demands [96]. There also appear to be gradients - rather than clear-cut divisions - in the functions encoded by certain neuronal populations [97,98]. This mixed and overlapping (i.e., redundant) functional coding likely underlies normal flexible behavior [99] and may - at least partly - explain functional network recovery or adaptation after injury. The implications of this complex neuronal processing for clinical fMRI are substantial. 'Activating' language regions with fMRI first requires us to engage language-related processes of surgical concern. Yet, knowledge remains incomplete of a) *what processes* are involved in language, b) how *specific* these processes are for language, and c) *how they map onto the brain's* intricate anatomy.

Secondly, **language is not a single unified behavior**. Many steps are involved in successful language production and comprehension. Brain regions that each perform more or less specialized computations likely combine in specific ways to support particular 'language' requirements [100,101]. Our aim here is not to propose a 'model' of language. However, it is the multiple (incompletely understood) processes underlying language that we measure with fMRI. Therefore, considering these processes offers a helpful basis for recommending tasks best suited to address the questions posed by clinical fMRI.

The '**core' set of processes required for different aspects of language** [102] includes a) *semantic access* (knowledge of concepts and meaning), b) *phonological representations* (the sound of words), c) *lexical* access (our store of learned vocabulary), d) *orthographic and graphemic knowledge* (visual word recognition and spelling), and e) syntax (knowledge of the rules governing word order and the functional roles of words in language). Speech also requires *articulation* (planning, coordination and programming of speech), which is necessary but not specific to

meaningful spoken language [103,104]. At present there remains uncertainty as to the level at which syntactic processing is supported by a neurally-distinct network in the brain [105], and about the neural substrates for temporal /syntactic order predictions (e.g., [106–108]). Other so-called '**peripheral' cognitive processes,** including attention, cognitive control, memory, temporal order prediction and theory of mind, while not specific to language, also directly influence language performance [109].

These 'core' language processes are often described in terms of (at least partially) discrete networks consisting of (at least partially) separate brain regions. Evidence of rapid, parallel neurophysiological signaling [110] paired with the brain's circuitry (e.g., [111]) suggests that many of these processing steps happen simultaneously [102]. The way in which language arises from the brain may not be fully or precisely captured by our concepts of 'how language works'. Nevertheless, the current dominant theory (modified from Wernicke's proposal) suggests two large-scale interacting systems [112,113]; one supporting conceptual (lexico-semantic) aspects of language and the other supporting phonological processing and speech (reviewed in [114,115]). While other models focus on the combinatorial aspect of language (as opposed to single word) processing (e.g., the Memory, Unification and Control model: MUC [116]), the 'dual stream' model focuses on auditory speech comprehension and production [112,113]. Ongoing refinements incorporate the additional processes associated with reading [117–119] and writing [120–122]. Fig 3 illustrates the dual stream model, presented alongside statistical maps of cortical brain regions activated during 'semantic' and 'phonemic' language processing according to predictive modeling of activation results from 13,450 neuroimaging studies in Neuroquery [84].

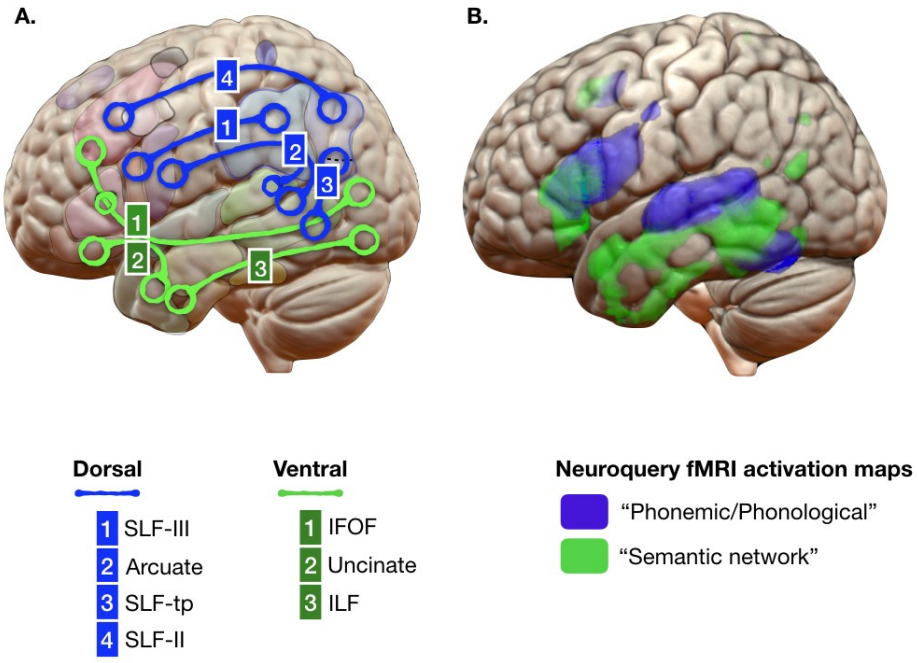


Fig 3. Theoretical 'dual stream' model of language processing. *Legend: A. Current prevalent model depicting at least two interacting streams of information processing that support phonological (blue, 'dorsal pathway') and semantic (green, 'ventral pathway', more bilateral) aspects of language function* [112,113]*). Early studies proposed that information in these streams is conveyed by discrete sets of fiber bundles in the dorsal stream (predominantly the arcuate and branches of the superior longitudinal fasciculus) and in the ventral stream (predominantly the extreme capsule, i.e., inferior fronto-occipital fasciculus and / or uncinate)*[123,124]*. The dorsal and ventral streams largely (but not completely) distinguish between the 'sound' versus the 'meaning' of spoken language, and thus focus on auditory language processing. Nevertheless, there are also ventral parts of the brain involved in phonological processing (e.g., in the superior temporal sulcus) and dorsal parts of the brain involved in semantic processing (e.g., the angular gyrus). Recent extensions propose additional pathways that link visual information to the language system during reading* [119] *and writing* [122]*. B. Results of multivariate prediction based meta-analysis of approximately 13,500 neuroimaging studies in Neuroquery (https://neuroquery.org/)*[84]*. Querying the search terms "phonological"/"phonology"/"phonetic" (blue) and "semantic"/"semantic network" (green) identifies patterns of cortical activation largely consistent with the dual stream model. The neuroquery-derived z-score maps are shown thresholded at z=3.1 (corresponding to p<0.01). For detailed anatomical regions see Table 1.*

Neurosurgical planning, however, demands a higher level of granularity than afforded by the 'semantic' vs 'phonological' system distinction alone. Mapping more specific language processes onto the dual-stream model is complicated by the fact that **most language tasks evoke the functions supported by both streams to varying degrees**. Understanding a spoken sentence, for example, relies not only on our knowledge of the sound of words (phonology) but also their meaning (semantics)[125]; two words individually may sound the same, but we understand the intended meaning in the context of the sentence. Similarly, it is thought that reading can occur either based on 'sounding out' the written letters (sublexical phonological assembly) or – in the case of irregular words (such as *plough* or *yacht*) – by directly recognizing entire words (lexico-semantics) [119,126]. Despite these parallel processes, **some brain regions appear to contribute more critically to language than others** [101,102,127]. The *preferential* contributions of different brain regions to different language processes are summarized in Table 1. These relative specializations are based on reported consequences of neural disruption / injury, summarized next.

[Table 1]

2.2.2.2 Neurological consequences of damage to localized functions and/or networks

"To locate the damage which destroys speech and to localize speech [itself] are two different things" – JH Jackson, 1874 (cited in [128]).

A region may participate in speech or language processing (as shown by fMRI) but not be clinically relevant because damage to this region does not impair speech or language processing when other preserved regions can support the same function. When considering the 'validity' of fMRI language maps, a strong focus is therefore placed either on correspondence to techniques that focally (or, at least, selectively by hemisphere) disrupt brain function (discussed in section 4.3), or on the consequence of irreversible damage through surgery, stroke, neurodegeneration or other injuries. This logic makes sense in the context of surgical goals – i.e., to preserve essential function – but is worth considering further in the bigger picture of brain function.

Many of the **pioneers of language localization theories observed that focal 'deficits' (or stimulation-induced disruptions in the case of Penfield) likely reflect disruption to a wider language network** (reviewed by [129]). Anomia, for example, is among the most frequently observed language deficits, but can result from disruption of several non-overlapping regions in the language-dominant hemisphere [101,102]. That is, the *task* of naming is not 'localizable' *per se* [127,130], because naming combines multiple language *processes.* As an illustration, naming pictures, for example, requires a prelinguistic process of visual object representation, semantic processing for object identification, phonological retrieval for activating an arbitrary label for the concept, and articulatory processing for producing the word form.

Observations in individuals with brain pathology or injury may not be fully representative of language organization in the general population, because injuries generally involve wider brain areas including essential white matter. Nevertheless, converging data from stroke [127,131] and resection outcome studies [132,133] support the existence of critical 'brain hubs' associated with long-term language deficits. The behavioural changes reported after temporary disruption or permanent injury to each of the putatively important cortical language-related areas is detailed in Table 1. Allowing for our still-evolving understanding, in Fig 4 we summarise the general approximations of core function-to-anatomy language mappings that are considered 'reliable' based on converging lesion, brain stimulation and imaging data (noting this is just one of different possible conceptualisations). Accordingly, the language processes most relevant to consider, and tasks commonly used to delineate them, are outlined in Table 2.

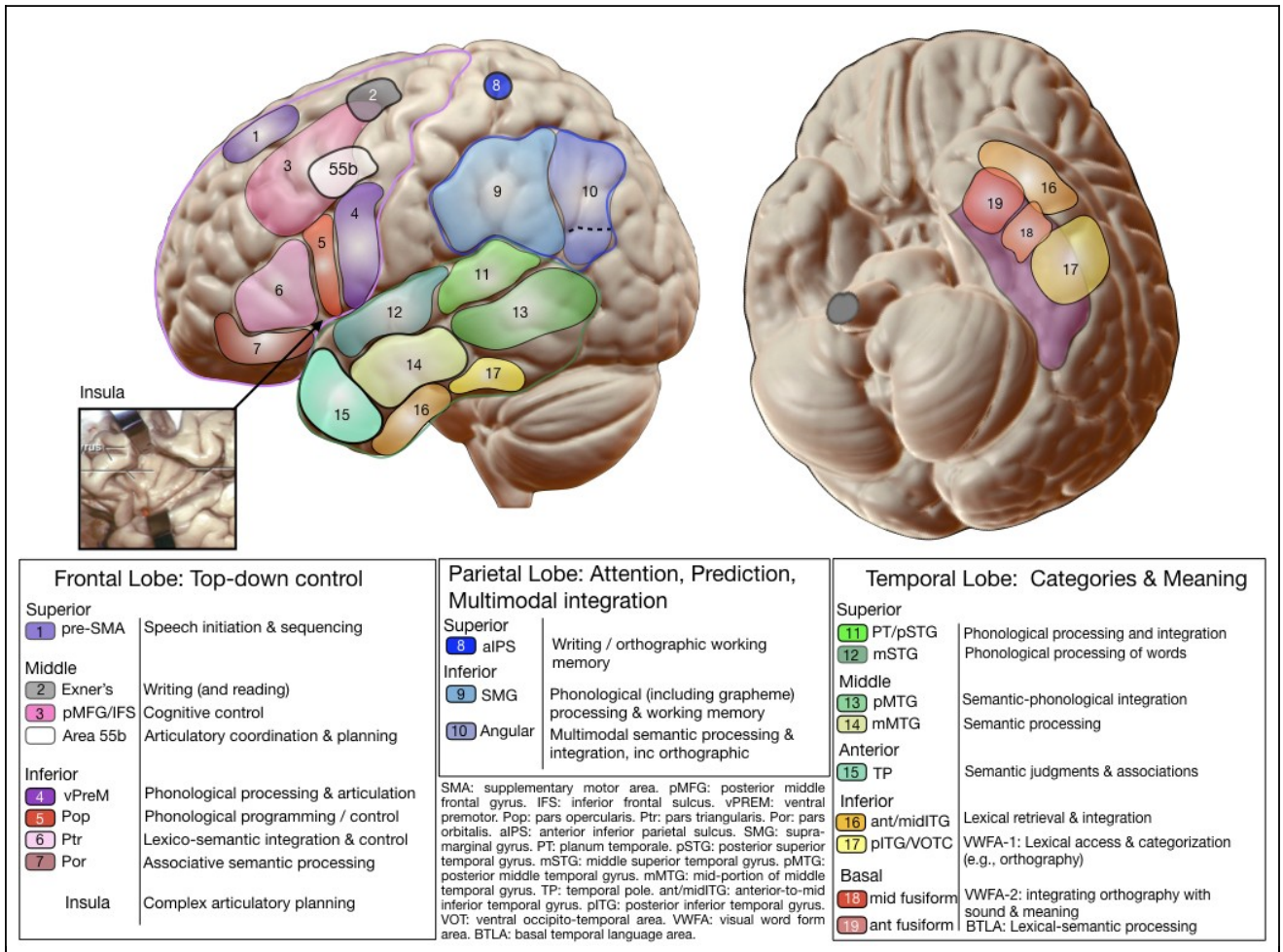


Fig 4. Language-related functional subdivisions of the neocortex. *Legend. Preferential roles of cortical sub-regions in language processing, based on a review of neuroimaging, intraoperative stimulation and lesion data (see details and references in Table 1). These summarized functions are neither absolute nor exclusive; many of these regions contribute to more than one language process, as well as to non-linguistic functions, and anatomical landmarks do not accurately correspond to functional boundaries, especially in patients. Instead, this figure highlights the higher likelihood of a given type of language function being represented in or near a given part of the brain, noting that both imaging and intraoperative findings are highly variable across individuals. Consequently, in a given patient undergoing surgery, it is relevant to consider a range of functions, allowing for the reasonable possibility that functions may have shifted (spatially due to mass effect, or as a result of functional reorganization). Subcortical structures including the basal ganglia and thalamus are known to contribute to language* [47] *but are not considered here due to the typical (resective) neurosurgical focus on functionality at the level of the neocortex (and underlying fiber tracts). Insula inset image adapted from* [134].

[Table 2]

2.3 Design and Paradigm considerations

2.3.1 Maximizing sensitivity

Once the precise clinical question and target language processes have been identified, the sensitivity of fMRI will be influenced by how easily the language-related signal of interest can be distinguished from a comparison baseline signal (measured during 'rest' or active 'control' conditions). When the aim is to capture the neural activity associated with a general task (for example, naming) but the specific responses (e.g., individual objects named) are not crucial, a **'block design' is favored** because this offers the greatest sensitivity to an *average* sustained BOLD response [135]. Using block designs, one or more language functions can be mapped in a relatively short experiment, making this the most prevalent design in clinical language mapping [4]. An added benefit is that it is normally easy to confirm the patient's compliance with task instructions by viewing the measured hemodynamic response, which should co-vary with the timing of each block.

In some tasks, the objective is to isolate correct responses (e.g., only those items named accurately out loud during silent periods introduced in the fMRI sequence). In these cases, capturing the neural response at a precise time is important, and this requires an 'event-related' analysis. The increased precision of event related designs, however, comes at the cost of reduced statistical power [135] and requires longer acquisition times to reliably estimate the overall neural response, especially if there is a low(er) number of successful trials. Event-related designs are therefore not typically employed for clinical language mapping (including among this Working Group). However, further research into the potential benefits of 'mixed' block and event related designs is warranted [136].

2.3.2. Advantages of a task panel approach

Variations in task demands affect within-hemisphere localization results in the same way they affect lateralization (discussed above). For instance, sentence-level fMRI tasks used in pre-operative fMRI (including sentence completion and naming to description) have been associated with greater activation (e.g., in the IFG, MTG and

the temporal pole [67] or in the temporoparietal cortex [137]) than tasks using a single word cue (such as auditory naming, verb generation, category fluency and antonym generation). Likewise, activation in core temporal and frontal language areas is reportedly more consistent when participants name multiple objects in a picture rather than one picture at a time [138]. Połczyńska and colleagues [51] noted a greater volume of activation (e.g., in bilateral angular gyrus) and engagement of additional brain regions (e.g., left anterior and posterior supramarginal gyrus) for a set of grammar tasks (e.g., relativized subject and object clauses) than during standard lexico-semantic tasks (e.g., object naming). It should be emphasized, however, that the location and extent of activation for any language task depends as much on the choice of control condition as on the specific demands of the language task. That is, the same language task can result in markedly different activation patterns when contrasted against different control conditions. Consquently, some task designs are better at revealing language areas than others. It is therefore **highly recommended that task protocols be conceptualized in terms of task contrasts rather than isolated tasks.**

When the clinical objective is to establish language *laterality*, as mentioned, mapping with multiple (e.g., a standard panel of) tasks can generate language maps in the language-dominant hemisphere that are superior (more 'complete') to those generated with a single task [51,137,139,140] and correspond better with Wada lateralization (e.g., [50,141]).

When the clinical objective is to *localize* specific language-related processes around a surgical lesion, a task panel approach is also generally advised, but the constituting task contrasts should be tailored to each surgical target. This is because task contrasts differ regarding their elicited magnitude of activation [56] with potentially a low extent of spatial overlap between tasks [137]. The partial divergence in activation during different language tasks has been confirmed using direct electrical stimulation [142,143].

Therefore, a task panel approach, contrasting different conditions, is typically needed to adequately probe all language functions at surgical risk. Mapping distinct language processes could theoretically be achieved using a single task containing multiple active conditions / task contrasts. In practice, however, there

are advantages (in relation to head motion and to statistical power) to employing short, targeted task contrasts, each separately assessing a given language process at risk. Further empirical data are needed to support the use of specific tasks over others for a given language process/region, but various options are listed in Table 2.

2.3.3 Optimizing specificity

2.3.3.1 Many-to-one processing

A long-recognized [144] criticism of clinical – and especially language – fMRI mapping is that statistical maps derived from fMRI lack *specificity* to individual cognitive processes [3,145]. Localizing specific brain functions implies that there is one-to-one mapping between a cognitive process and a brain region. However, as detailed in section 2.2.2, accumulating neuroscientific and lesion-based evidence indicates that multiple cognitive processes can engage a single brain region; so-called 'many-to-one mapping' [146]. Selecting appropriate tasks – and especially task *contrasts* – that selectively engage individual brain regions is therefore a challenge. Within this constraint, general approaches are considered next.

2.3.3.2 Selecting task 'control' conditions

As mentioned, **different task contrasts can generate very different patterns of 'task' activation**, which differ in their sensitivity (how well they detect language-related activation) and specificity (how strongly the activations relate to language as opposed to other types of task-related processes) (Fig 5a). 'Resting' fixation remains a common 'control' condition in clinical fMRI designs [4] and is implemented in many commercial and clinical task-fMRI protocols that compare blocks when the task is 'on' versus 'off' (i.e., during rest). However, the challenges with using 'rest' as a comparison baseline have been highlighted previously [147]. 'Rest' is an uncontrolled state and elicits notable levels of brain activity [148]. Research suggests that several important functions take place during 'rest', including memory consolidation [148] and internal cognitive and linguistic processes (such as mental imagery, concept recall and mind-wandering) [149], even when the 'rest' periods are very short (e.g., 3 seconds) [148]. Consequently, using 'rest' as the comparison condition in a language task contrast limits sensitivity for detecting activation in some language-related networks, particularly those involved in semantic processing. At the same time, using

uncontrolled ‘rest’ as the comparison conditions limits the specificity with which the underlying cognitive language process can be isolated. The relevance to language mapping is illustrated in Fig 5b, showing the effect on a semantic decision language activation map just by varying the control condition from ‘rest’ to auditory tone decision. Consequently, when increased sensitivity and specificity is required in the interpretation of fMRI, carefully designed baseline conditions should aim to ‘control’ for aspects of task performance that also occur at the same time as the language process of interest [150].

The choice of which control conditions to use is often a balance of *specificity* against *sensitivity* and *clinical feasibility*. Also influencing this choice are: what the results will be used for and how the data will be analyzed. Most commonly, the aim is to map complementary language processes (for example, reading and auditory comprehension) using individual language tasks that will be analyzed separately. In this case, the choice of control condition in each task should reflect the desired level of precision in mapping each function (Table 2), especially if the results are intended to be compared to intraoperative stimulation findings. In this context it is important to note that increasing the complexity of a task may reduce patient compliance or performance unless careful, step-by-step instruction is provided with adequate practice on the tasks prior to scanning. An alternative approach is to analyze multiple tasks together to isolate one target language process (e.g., naming prior to temporal lobectomy) through multiple modalities (i.e., seeing, reading, hearing). In this approach (known as conjunction analysis), the aim is to identify activations shared across naming tasks while controlling for (i.e., removing) the effect of sensory input (e.g., auditory/visual). For this application, a common non-active and/or passive baseline may offer the best sensitivity, at least for the phonological processes engaged during naming (semantic processes engaged during naming are likely to remain invisible due to their engagement during the resting baseline) [151]. Considerations in selecting appropriate baseline conditions are listed in Box 3 and their effects illustrated in Fig 5.

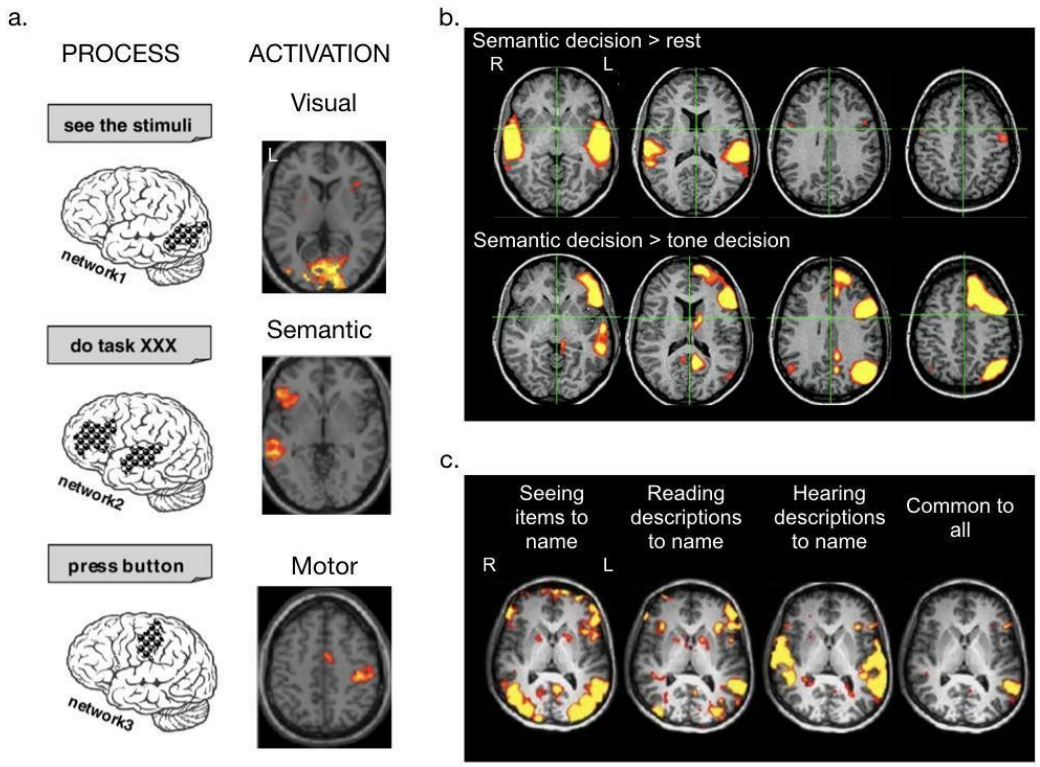


Fig 5. Effect of baseline condition and test modality on fMRI activation patterns. *Legend. a. Language activations maps almost invariably reflect a mixture of processes involved in the performance of a given task, often including sensory (e.g., visual) input unrelated to language, as well as motor (e.g., button press) or cognitive processes (e.g., decision-making) to generate a task-required response (adapted from* [152]*). b. Differential activation patterns during a semantic decision-making task* [153] *(deciding whether aurally presented words are both "found in the USA" and "used by humans") depending on the choice of control condition: resting fixation, tone decision (deciding if auditory tone sequences contain two target tones or not). Note first the subtraction of almost all bilateral auditory sensory cortex activation (yellow-red color in top panel) when the comparison condition includes a control for low-level, non-linguistic auditory processing (bottom panel). Second, note the appearance of a large, left-lateralized frontal-temporal-parietal language network (bottom panel) when an active, non-linguistic control task is incorporated, which suppresses language processes that normally occur during 'rest'. c. – Alternatively, taking a conjunction of different language tasks can identify core language regions. In this approach, a single task (naming) is chosen and repeated several times in different modalities (visual, auditory).*

Considerations when selecting task control conditions • BOX 3 •

- **Active control conditions are generally recommended** over a 'rest' or passive baseline when the objective is to localize activation that is specific to language (rather than activation reflecting low level sensory, motor, or attention processes). Alternatively, a conjunction analysis approach can be used to attempt to isolate a particular function (such as phonological retrieval) through different tasks performed in various sensory domains (e.g., hearing, reading, seeing).
- Active non-linguistic control conditions can also increase sensitivity for detecting many language areas compared to using a 'rest' or passive baseline, due to the fact that many language areas, particularly those involved in semantic processing, are active during 'resting' states (Fig 5b).
- **Active controls that differ significantly from the task condition (e.g., tone discrimination vs word comprehension) can be useful to increase resulting laterality indices by removing non-specific bilateral activation in sensory and attention networks and increasing detection of language-specific activation** (Fig 5b) [56].
- Some types of **control conditions** (e.g., presenting reversed speech or pseudowords) **engage phonological components of the language network** [154] **and potentially heighten attention** [155]; their use can reduce relevant signals related to phonological processing. In such designs, also including resting fixation blocks may be useful to explore the main effect of both active and control conditions separately, to aid interpretation of active > control block comparisons. This may require more powerful / flexible analysis models than allowed by certain commercial fMRI analysis packages.
- **Cognitively demanding** control conditions **make tasks more complex** for patients to understand or remember. These can therefore be impractical for some patients with moderate to severe comprehension, memory, attention or compliance deficits.
- **Detailed practice sessions and monitoring performance** during the scan (in real time / by recording responses) become **essential as task complexity increases**.

2.3.4 Pre-fMRI practice session:

In addition to baseline performance abilities (considered in section 2.1), the success and specificity of fMRI is also influenced by discrepancies between what the patient is *asked to do* and what they *actually do*. Appropriate interpretation of the fMRI results therefore requires good knowledge of patient performance as there is a corpus of data in healthy volunteers suggesting that effort and ability both contribute to activation

magnitude. The importance of pre-scan practice has been previously emphasized [156,157] and is reinforced by results of a recent survey indicating that the most frequent reason for inconclusive language fMRI results was inadequate patient performance (47.3%) [7]. Of note, most participating centers either spent less than 15 minutes instructing and practicing with the patient (55%) or gave no pre-scan instruction at all (14%).

The Working Group strongly recommends that, prior to fMRI, an instruction and practice session on the tasks is performed with the patient to ensure they understand and are able (with appropriate modifications where possible) to perform the task as required. This preparation should include a detailed description of the problematic nature of head movements (section 3.3.2), and specification of when minor movements (such as scratching or clearing the throat) are permissible (i.e., when the scanner is not making noise). **If the fMRI task will require responses (even covert responses) it is advisable to practice example responses with every patient to ensure they understand when and how they are expected to respond**. Comparable but different examples should ideally be presented for the practice session than those shown during the actual fMRI task (e.g., different verbs/letters for a practice word generation task) to minimize unwanted practice effects that could result in a reduced fMRI signal.

2.3.5 Modifications for patients with specific impairments

A few specific considerations arise in populations with selective deficits in vision/hearing or patients presenting with various degrees of language deficit. These are considered next and summarized in Box 4 and Fig 6.

2.3.5.1 Task modality

Language task stimuli are most frequently presented visually [4,10]. Clearly, for patients with difficulties processing certain types of stimuli (e.g., written words in dyslexic or alexic individuals), or any other uncorrectable vision impairment, it may be necessary to change the modality of stimulus delivery (e.g., to auditory cues or picture cues) for language mapping to succeed.

2.3.5.2 Task speed and difficulty

While the target language functions should be at least residually intact, some strategies may enable successful language fMRI mapping in patients with pre-existing language impairments. For a patient with word-finding difficulties, the pre-scan practice session can identify task items they can perform with sufficient consistency to adapt task stimuli to the individual's abilities. Other patients may routinely generate successful responses but require longer to do so. In this case, adjusting the pace of stimulus presentation / response collection may be necessary to elicit meaningful brain activation (see [144]). This is because fMRI activations are influenced by task difficulty in numerous language tasks [154,158–162]. The aim when adapting tasks is to ensure a patient can perform the task over the entire session (e.g., in those with impaired cognition, and especially attention) while keeping the task challenging enough to maintain continuous engagement (and minimize mind-wandering). However, when task performance becomes too difficult, the same regions can show less activity, indicating a relationship (often shaped like an inverted 'U') between task difficulty, cognitive effort, and BOLD response [163,164]. Together, the data suggest that **tasks should be optimized for patient performance** - hard enough to require cognitive effort, not so difficult that the patient disengages altogether. A recommendation ***in patients with existing language impairments*** is to **maintain a balance between accuracy and motivation by targeting around 70-80% performance accuracy.** Conversely, in cognitively high-performing patients with high levels of motivation, reliable activation maps can generally be obtained (i.e., there is not typically a need to make tasks more *difficult*).

Among our Working Group, some employ pediatric versions of a task in adult patients with aphasia or impaired speed of performance. Others have had good results with protocols specifically developed for aphasic patients, such as Adaptive Language Mapping (ALM)[165,166]. ALM adjusts task difficulty on-line based on performance and produces maps of phonological processing relative to a non-language control, semantic processing relative to a non-language control, and semantics versus phonology (https://aphasialab.org/alm/). Proposed adjustments to task parameters are summarized in Box 4 & Fig 6. When task modifications are not feasible on a per-patient basis (for example because of lack of options in certain commercial software

packages), it becomes even more important to establish, through cognitive evaluations and pre-scan practice sessions, that the patient can perform adequately (at least 70% accurate responses). Despite all steps that are taken, there are times when task fMRI cannot be executed successfully, due for example to insufficient residual abilities to perform the task, or inability to maintain focus / inhibit unwanted responses. This eventuality should ideally be identified through appropriate screening before arranging the fMRI scan, but there remains the possibility of occasionally unsuccessful fMRI exams (such as also rarely occurs with diagnostic structural MRI scans).

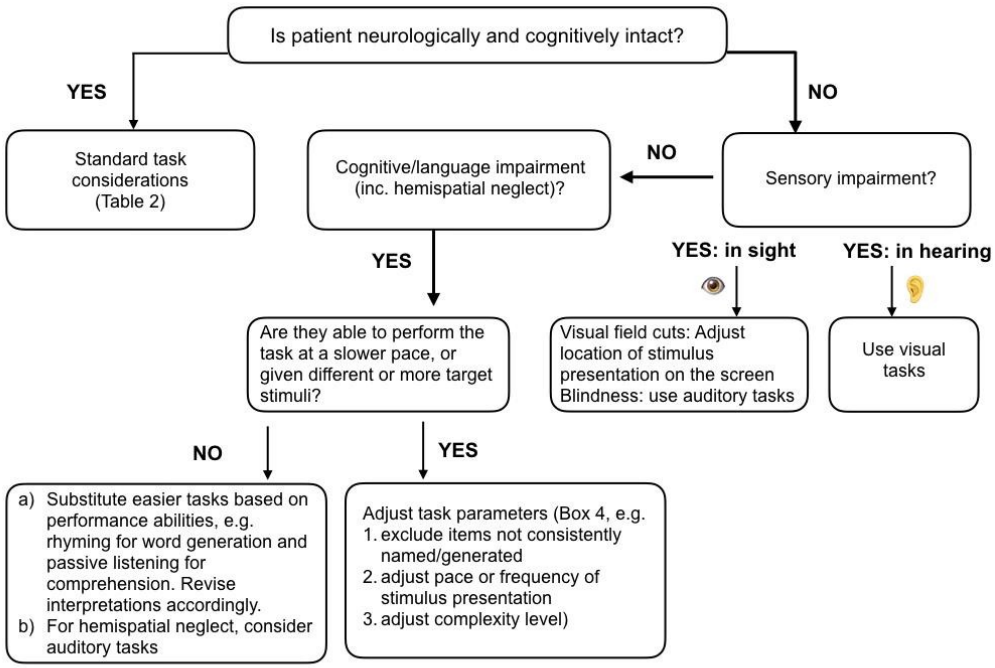


Fig 6. Proposed task adaptations for patients with sensory, cognitive or language impairments. *Legend. Proposed workflow to adapt fMRI language tasks for patients with specific deficits. For patients with concomitant hand weakness, if the task requires manual responses, possible adaptations include using an alternative hand, or capturing responses using alternative approaches (eye tracking if available, or using foot tapping to indicate which stimulus they wish to select, which are manually recorded by the person acquiring the fMRI scan).*

2.3.6 Design and acquisition considerations for pediatric patients

A study of over 400 language fMRI scans in clinical child populations indicated an overall promising success rate of fMRI, but a higher rate of 'failed' scans when compared to typically developing children [167]. Predictably, the success of fMRI varied greatly across different ages. Task 'failures' (uninterpretable or null activation maps: 15.8% of all attempted scans) were driven primarily by head motion, but also attributed to some children falling asleep [167]. These findings emphasize that **children may benefit particularly from task modifications that a) sustain attention and effort** using several (short) tasks, **and b) ensure a minimum number of successful responses** [168]. Additional considerations in pediatric patients include whether they can tolerate a scan without sedation, can lie sufficiently still for the minimum task duration and can reliably make required responses or would, instead, require passive stimulation-based mapping [169–171].

Excellent practical recommendations exist to maximize successful fMRI in pediatric patients [172,173], including a mock or simulated scanning session [174]. Clinical fMRI is generally more successful in children > 7 years old [167] but is feasible from the age of 5 [175] and even as young as 3 when using passive conditions as such story listening [172]. Most clinical fMRI studies to date have involved children with epilepsy, weighting the existing data towards hemispheric language lateralization. In children over 7 years old, **it is generally recommended to use at least 3-4 short tasks** [50,176], partly because head motion increases with every additional minute of acquisition [177].

For meaningful paediatric clinical language fMRI, activation procedures should be **individually tailored** by selecting tasks and stimulus presentation rates not only **according to performance level** (as for adult patients), but also according to **developmental stage** [50,178]. For example, written stimuli need to be selected at the appropriate literacy levels of the child. Around 12 years of age, reading networks are likely to have stabilized [179] but children with a developmental brain lesion or early onset epilepsy may have different developmental trajectories. This consideration is especially important for laterality assessments because greater recruitment of the right hemisphere is observed with greater task difficulty [178]. Non-linguistic 'difficulty'-related contralateral activations likely drive an increase in apparent 'bilaterality' in

some fMRI tasks relative to language lateralization results from Wada testing [176]. Additionally, **selecting an active baseline** is especially beneficial to maintain attention and engagement in younger children who are less able to comply with the requirements of 'resting fixation' [180].

A challenge is that a very wide range of tasks and task contrasts has been described for pediatric language fMRI, and data showing probabilistic activation maps or comparisons between these tasks in terms of lateralization or activation pattern in children remain lacking. When considering an optimal battery of paediatric language fMRI tasks, one **solution is to employ a different panel of tasks best suited to specific age ranges**. The main drawback of such an approach is that the interpretation of results between different age groups, and, importantly, in the same child over time, becomes challenging [181]. **An alternative solution is to adopt a single battery of tasks**, in which each task aims to engage a specific language process (e.g., phonology, semantic retrieval, …), **using parallel versions that vary in difficulty level**.

Language comprehension, for example, can be assessed in children using an auditory description decision task that probes phonological and lexico-semantic networks at three levels of difficulty (tailored to ages 4-6; 7-9 and 10-12+ years). In this approach, sentences (composed of easy, medium or more difficult words) are presented which describe a noun, and the patient indicates if the sentence is correct or not [181]. This active condition is compared against reversed speech containing a target tone. When evaluated against intraoperative stimulation, this task contrast offered higher specificity compared to a simpler auditory category decision task, passive story listening versus reversed speech, or story reading (each contrasted against reversed speech and tone detection for the auditory task and visual detection of white but not black squares in the reading task)[182].

For paediatric *language production*, activation patterns for commonly used word generation tasks (such as covert antonym generation) also vary substantially with age, especially in the left frontal lobe, indicative of task-specific developmental changes [183,184]. Visual rhyme judgment (deciding if two written words or pseudowords rhyme),

which is typically contrasted with a nonlinguistic visual matching task (deciding if pairs of symbols/line arrays match or not), was recommended by the ASFNR, but can only be used in literate children. The difficulty level of language production tasks can also be tailored to provide easy, medium difficult, or more complex words, and can be presented in the auditory or the visual domain. An active unrelated control condition such as finger tapping [183] offers a continuous measure of in-scanner vigilance and performance that can be readily monitored. In preliterate children, language functions can alternatively be mapped with auditory rhyming (or passive story listening) tasks contrasted with tone decisions (or listening to instrumental music or inanimate environmental sounds), for example.

Further paediatric language task recommendations vary depending on the surgical target. For example, for planned resective surgery involving posterior temporal or inferior parietal regions in literate children, a visual sentence completion task (at appropriate literacy levels) contrasted with passively viewing 'gibberish' sentences can be clinically indicated. A semantic task contrast should be considered when the surgical target includes the (especially anterior) temporal lobe. In adult patients, a visual naming task using picture cues to generate either single names (*object naming*) or sentences (*object-in-sentence naming*) (compared to resting fixation) is associated with highly consistent activations of the inferior frontal gyri (especially pars opercularis), superior temporal sulcus and the temporo-parietal junction in healthy volunteers [77] [75] and good correspondence to Wada results in both tumor [141] and epilepsy [85] patients. While naming engages many language-related processes and is therefore less specific for localizing individual processes (section 2.2.2.2), the high sensitivity of naming tasks is of theoretical added benefit in pediatric populations. However, **further development is advocated to establish age-appropriate naming stimuli and norms in pediatric populations** [185], given the ease of object naming, its sensitivity to basal temporal language processing [85] and the robust activation patterns this task provides.

Modifications for specific *performance difficulties* include using age-appropriate story listening (versus e.g., tone detection/discrimination) if the semantic decision task is too difficult or if word generation abilities are limited [186] (Box 4). Adding a tone detection button response during passive listening helps to ensure attention [178].

Similarly, In the absence of auditory stimulus delivery, a written word semantic matching task could be performed in literate children to engage semantic association processes.

Final considerations are that pediatric language mapping may be influenced by higher brain metabolism affecting the measured BOLD signal and the greater likelihood of head motion [187]. Selecting short tasks (around 3 minutes) and repeating these to average the fMRI responses can be advantageous. Active noise cancellation (Appendix C) can be beneficial not only to facilitate auditory stimulus transmission (see below) but also to reduce startle movements irrespective of task modality.

Adaptations for patients with pre-existing impairments and pediatric patients	• BOX 4 •

Sensory impairments:

- **Select tasks based on modality**; if patients are hard of hearing, use visual tasks, and if vision is severely impaired, then use auditory tasks. For dyslexic/alexic patients, adapt the rate of presentation of words if possible (to map residual reading abilities) or select auditory tasks.

1. Adult patients

Language impairments:

In patients with existing (mild or mild-to-moderate) language deficits, it is advised to either select overall tasks according to the patient's abilities, or adjust the difficulty of task items according to patient performance levels. Tried-and-tested adjustments include to:

- **Reduce the rate of stimulus delivery** (e.g., presenting items at a slower pace or for a longer duration [76,85,138,141] and / or **increasing the frequency of stimulus delivery to increase response chances** (e.g., presenting 6 target letters / categories during a 20 or 30-second fluency block instead of only 1 or 2).
- **Adjust to reading/comprehension level** for story reading/auditory comprehension/semantic decision.
- **Remove items that cannot be consistently named** for naming tasks (as for intraoperative testing) or using only high frequency words [188].
- **Use an adaptive mapping approach** that adjusts task difficulty on-line

based on performance (e.g., https://aphasialab.org/alm/).

In extreme cases of patients with severe language production impairments, a very simple sequencing task (e.g., days of the week / counting) may produce results when self-directed tasks fail, [178] though the resulting activation may reflect mainly motor articulatory processing rather than language *per se* [189]. In any case, patients with severe comprehension impairments are generally not suitable for fMRI (or for other language mapping procedures, e.g., cortical stimulation mapping).

Cognitive impairments:

- **Reduce complexity and load** of attentional demands (e.g., switch from a more 'difficult' fluency task to an 'easier' verb generation or picture naming task).
- If the patient has a learning disability, **consider active noise cancellation** / modified sequences to reduce startle effects.

2. Pediatric patients

- If possible, practice in a **mock scanner / virtual MRI simulator** and, if necessary or helpful to ensure compliance, have a parent present in the scan room during fMRI.
- **Sustain motivation and effort** by adapting stimuli to abilities and use short tasks, repeating these for statistical power. Using **2 to 4 tasks** is advocated in young patients to maximize chances of at least 1 interpretable result [167,176].
- Select reading or comprehension stimuli based on ability level [181,190]. **In preliterate children,** consider comic-style pictures to generate sentences and auditory tasks to assess comprehension with responsive naming ("where do people live?" …).
- If word generation or reading abilities are limited, **passive story listening** may be useful with an auditory control condition given careful interpretation [191].
- **If highly anxious: consider** substituting active tasks with **passive tasks**.

2.3.7 Bi- and multilingual patients

Census reports indicate that more than half of European and US populations are bi-or multi-lingual. A recent systematic review of evidence in bilingual neurosurgical patients suggests that, in addition to brain regions that are shared, separate areas of

cortex are often found that uniquely subserve the primary versus additional languages [192]. Patients who acquired a second language early in life are more likely to have less robust left-language dominance and recruit their right hemisphere more than matched monolingual patients, both for their first and second languages. Conversely, co-localization of the first and second languages was less likely when the second language was acquired later in life and when the languages were dissimilar (e.g., German and Tongan versus German and English). Consequently, laterality can be the same for both the first and additional spoken languages [193]. Because of the potential differences in the organization of languages in bilingual and multilingual individuals, **language fMRI mapping should prioritize the primary language, but should ideally include all languages used by a patient in their everyday life**. The number and order of languages feasible to map may need to be evaluated on a case-by-case basis, considering the additional burden on the patient for a longer fMRI scan.

In practical terms, certain tasks, such as silent object naming, can be administered in multiple languages without the need to modify stimuli. However, cultural differences in familiarity with certain stimuli may affect performance. Similarly, the difficulty of some tasks, such as phonemic fluency, differs depending on the frequency of individual letters in different languages. For such tasks, the task stimuli should be chosen according to available normative data (i.e., appropriate 'difficulty' level) in the language to be used. Several common paradigms are available in multiple languages as part of some commercial packages, or freely from research groups (www.cogneuro.net/hbm2017 [85]). Among participants of the Working Group, several employ **parallel language versions of tasks** (i.e., the same task in multiple languages), especially for reading and comprehension tasks.

Over and above possible differences in the brain regions supporting language production, speakers of multiple languages make use of a distributed 'control' network [194–196]. Damage to this network can result in uncontrolled fixation to a single language [197,198] or mixing/switching of languages without aphasic symptoms [199,200]. This 'language control' network is formed by several brain regions, including the anterior cingulate cortex, basal ganglia (caudate), IFG, dorsolateral prefrontal cortex (including Brodmann area 9), superior temporal sulcus, and inferior parietal lobe [196,201–211]. Among these, the middle frontal gyrus (MFG) has been implicated as a hub for language

switching by several neurosurgical language mapping studies [204,208]. Specifically, impaired language switching has been reported in bilingual patients following stimulation during resection in the region of the inferior frontal sulcus and MFG [209,210]. However, no single 'core' area for language switching has been identified [212]. Stimulation-related language switching has also been reported in the superior posterior temporal lobe [206,212] and damage to the dominant hemisphere supramarginal area may sometimes result in inappropriate mixing of languages within the same statement (e.g., [213]. Therefore, **an additional language switching task is useful to consider when undertaking language fMRI mapping in bilingual patients**.

2.3.8 Working Group recommendations on task selection

In summary, key considerations in localizing language functions include **selecting tasks and tailoring task contrasts (1) to the anatomical location of the pathology, and (2) to the targeted type(s) of language processes. Consequently, our Working Group advocates the implementation of tasks firstly according to intended clinical purposes and, secondly, in line with the known contribution of anatomical regions to language** (Fig 4).

Consistent with previous reports**, we recommend administering multiple tasks, probing a range of speech and language processes, for clinical fMRI**. The optimal selection of both the language task condition, and the comparison 'control' condition, vary depending on the purpose of the scan and the abilities of each patient. When the purpose is **to establish hemispheric or lobe-level language laterality**, **a standardized panel** of well-characterized tasks can help facilitate the expert's interpretation of laterality profiles based on extensive prior experience using those tasks (Box 2). In terms of task design, **for language lateralization, an inactive (resting) baseline is generally not advised**. **When the clinical question involves evaluating surgical risks in and around specific brain regions**, it is of higher clinical relevance to select task contrasts that more precisely engage the most relevant language processes at risk depending on the **location and extent of the surgical target.** For localization attempts, it is important that each chosen task contains appropriate contrast conditions to help maximize the specificity of activations by reducing the impact of non-language related brain processes engaged by the

language task (Table 2). Where first-line recommended tasks are not feasible due to existing sensory or language impairment, suggested alterations based on published literature are summarized in Fig 6/Box 4.

3. Part 2. Acquisition and analysis of clinical fMRI data

3.1. Skills required for optimal clinical language fMRI

The goal of presurgical clinical language fMRI is to 'map' language's biological correlates. Clinical language fMRI is therefore a *form of cognitive assessment*. Executing clinical language fMRI without a good understanding of cognition, task design, and assessment introduces a high risk of occasional clinically relevant errors.

Similarly, the *result* of clinical language fMRI is typically a form of probabilistic map of the brain regions involved in selected language processes. The analytic steps taken to create this map involve decisions that fundamentally change the results, revealing or obscuring task-related signal. **Clinical language fMRI results are therefore critically influenced by the operator's expertise in image analysis and statistics**. Further, the knowledge of what questions a surgical team is asking of fMRI – explicitly and implicitly – and of how statements about fMRI results are interpreted by a clinical team, will determine if data are used appropriately in surgical decision making.

Therefore, while fMRI is sometimes considered a standard radiological exam, it is not. **Clinical fMRI is inherently interdisciplinary and requires ongoing (case-by-case) input from experts in multiple domains** [9]. Standard doctoral training in any one discipline on its own – e.g., neuroradiology, neuropsychology, neurology, neuroscience – may not be sufficient to provide reliable, valid results. **Therefore, a single individual with extensive multidisciplinary training in clinical fMRI, or a pair of individuals with different training** – one centering on neuroimaging/neuroradiology/physics and the other neuropsychology/cognitive neurology/cognitive neuroscience, with each having additional training in fMRI – **are needed for optimal patient care**. These individuals should; 1. be embedded in the neurosurgical team to understand what information the surgeon finds most important

for each specific case, 2. have input in designing a site's protocol; 3. be available on a case-by-case basis to assist in acquisition, analysis and interpretation of fMRI data; 4. have access to post-surgical language outcomes and potentially interact with neurorehabilitation teams. Given that both the surgery and the preparation for it could result in patient distress, it is worth considering if the same specialist practitioner can accompany the patient through the entire (or multiple stages of the) process (i.e., pre-op evaluation, fMRI and surgery).

3.2. **Data acquisition**

In addition to the unique nature of *what* is being measured, factors over and above task selection influence *whether and how precisely* the desired signal will be detected. Technical aspects of MRI data acquisition and processing have a key impact on the success and utility of clinical fMRI in the *individual patient*. Among the most critical recommendations is to **monitor the success and quality of fMRI exams in real time and seek feedback from patients after every fMRI run**, so that any apparent problems with data acquisition (such as movement during the scan or inadequate task performance) can be resolved before the patient leaves the scanner.

3.2.1 *Generic fMRI acquisition parameters*

Technical considerations in the acquisition of fMRI data for language mapping are largely comparable to those arising in other pre-surgical applications (e.g., motor / visual function mapping). Recommendations are summarized in Box 5 and further detailed in Appendix C.

Clinical fMRI acquisition parameters	**• BOX 5 •**

For greater in-depth discussion, see Appendix C.

- **3 Tesla is the recommended field strength** for clinical fMRI given its improved contrast-to-noise compared to 1.5T [214].
- Multi-channel head coils, **ideally 32** (but alternatively 64) **channel coils**, are preferred due to higher signal-to-noise ratio and ability to enable accelerated pulse sequences.

1. Temporal resolution

The rule of thumb is that, within clinically achievable parameters, *faster is better.*

- **It is not advised to use long TRs (much above 3s),** because of lower statistical sensitivity accompanying slow image sampling rates [215].
- Where accelerated sequences are available, shorter TRs (between 1-3s, and ideally closer to 1s) offer substantial statistical benefits.
- Optimal flip angles have to be adjusted to the TR, e.g. to about 60° for a TR around 1s (depending on the assumed T1).
- Shorter TRs generally improve the ability to remove structured artifacts from fMRI data [216–218]. However, very short TRs (e.g., substantially < 1s) using simultaneous slice (SMS) acquisitions and parallel imaging acceleration techniques (PAT, such as GRAPPA or SENSE) carry a high computational load. These may be incompatible with some current hospital picture archiving systems, generate much larger file sizes and require longer imaging computation times. These may also produce artifacts on some scanners and with some head coils. There are therefore diminishing returns for high accelerations to obtain TRs <1s versus using the additional acceleration to pursue alternative strategies such as multi-echo time fMRI [219].

2. Spatial resolution

- **In-plane voxel-size of EPI should not be larger than 3mm,** in order to balance spatial sensitivity and signal to noise.
- **Choose slice thickness to provide near isotropic voxels** [220,221].
- Higher spatial resolution fMRI (e.g., less than the mean cortical thickness of 2.6mm) offers the important benefit to more specifically localize signal to one side or the other of a sulcus [218,222]. However, smaller voxels of higher spatial resolution come at the cost of less within-voxel signal [223].

- ***A currently optimal sequence with single echo simultaneous multi-slice fMRI might currently be ~2.4mm isotropic resolution, TR close to 1s, using a SMS factor of 4 or 5.***
- At 3T, a TE of ~30ms is often used as a trade-off between signal loss and BOLD sensitivity.
- Further gains might be possible with further acceleration and obtaining multiple echos but maintaining the spatial resolution and TR [224]. Multi-echo approaches have the benefit of sampling both shorter and longer echoes and optimally weighting their contribution to each voxel based on the voxel's T2*, where each echo represents an independent sampling of the data, allowing multiple echos to be combined to reduce random noise [224].
- Vendor B1- normalization techniques (e.g., Siemens pre-scan normalize) are recommended to reduce image intensity inhomogeneities and have the added benefit of reducing motion-related intensity changes.

3.2.2. *Additional hardware requirements for fMRI*

Performing task fMRI requires hardware over and above the MRI scanner itself. A combination of both visual and auditory tasks together allows a more comprehensive mapping of language networks than tasks delivered in only one modality (e.g., [58,59,71,123]), as is also reported for intraoperative stimulation mapping [225]. Dedicated equipment needed for language mapping in visual and auditory domains is considered in Appendix C.

3.2.3. *Complementary sequences*

Specific guidelines exist for diagnostic imaging, for example in gliomas [226] and epilepsy [227]. Optimal radiological sequences are not considered here, except to highlight that functional scans are generally acquired separately from diagnostic MRIs. The most diagnostically appropriate anatomical scans would normally be acquired within the same scan session as fMRI for co-registration of the fMRI results and/or fusing in the neuronavigation system. During all fMRI exams, however, it is recommended that **additional field mapping sequences** are acquired for subsequent correction of EPI distortions; a step that is essential **to ensure accurate within-subject alignment between EPI and** non-EPI (e.g., **T1) data** (see section 3.3.1). Because of their effects on T2*, it is preferable to acquire fMRI prior to administration of T1 relaxivity agents containing Gadolinium; however, successful fMRI results have been obtained after contrast administration [228].

3.2.4. Integration with diffusion tractography

Tractography based on diffusion weighted imaging (dMRI) provides complementary clinically-relevant information to fMRI because lesions that are not directly in cortical language regions may still affect language-relevant subcortical fiber tracts. **Diffusion** (tensor, or ideally more advanced) **imaging is therefore recommended as part of any functional study**, in order to be able to reconstruct the key tracts (Appendix D). Most commonly, the tracts connecting language regions (i.e., arcuate fasciculus, fronto-parietal and temporo-parietal branches of the superior longitudinal fasciculus, the inferior fronto-occipital fasciculus, inferior longitudinal fasciculus, uncinate and frontal aslant tract) are of interest. Especially in the presence of a space-occupying lesion, tracts may be displaced and/or infiltrated in ways that are not easily predicted.

Indeed, thanks to diffusion tractography it is possible to explore if white matter fibers have been displaced, deformed or interrupted by a space occupying lesion (e.g., [229,230], and thus ensure better presurgical planning and intraoperative neuronavigation. There are many ways to collect and process diffusion MRI data for tractography [231], contributing (alongside lack of consensus regarding tract nomenclature) to variability in tract segmentations (e.g., [232]). However, the importance of integrating subcortical connectivity in attempts to 'map' distributed language networks has been repeatedly highlighted [48] and key surgically relevant tracts are summarized elsewhere [115,233–236].

3.3. Analysis

Unlike structural MRI, it is not the acquired fMRI-BOLD images themselves, but the statistical results generated from them upon which clinical interpretations are based. Choices in processing and statistical analysis have substantial impact on the resulting activation maps. This issue is well-recognized, and several comprehensive reviews have detailed analysis 'best practices' for the general neuroimaging community [14,237,238]. These predominantly aim to promote reproducibility across group studies of – typically – healthy volunteers. Lack of standardization is a critical problem for clinical fMRI as well [239]. However, **the decisions relevant to analyzing individual patient data differ in important ways from guidelines that apply to group analyses**. Pitfalls arise at various stages in the fMRI processing pipeline that are not specific to language mapping (Appendix E). Below, we highlight just a few select processing choices that particularly impact on the interpretability of *single-subject clinical fMRI results*. Overall pros and cons for various analysis choices are summarized in Table 3. The single most important recommendation is to **undertake quality control at every step**.

3.3.1. Distortions and distortion correction

Interfaces between tissue and air (e.g., in sinuses) cause magnetic susceptibility changes in common fMRI sequences (i.e., gradient-echo echo planar imaging). These susceptibility changes (or 'field inhomogeneities') cause both geometric distortions and signal loss in the EPI (fMRI) images. Aligning distorted fMRI images to (undistorted) anatomical scans results in a mismatch, displacing the apparent location of fMRI activity with respect to the patient's anatomy (Fig 7). Because spatial image

distortions are of great clinical concern and directly influence interpretations of the proximity of critical cortex**, corrections for geometric EPI distortions should be performed wherever possible** (see Appendix E). Quality control should be performed to ascertain the adequacy of registration and any areas of relevant signal loss (e.g., basal temporal cortex), should be clearly emphasized in the final fMRI report (see section 3.4). If distortion correction is not possible, an alternative, better than nothing, approach is to use in-plane acceleration such as iPAT in the image acquisition to reduce the magnitude of the distortions.

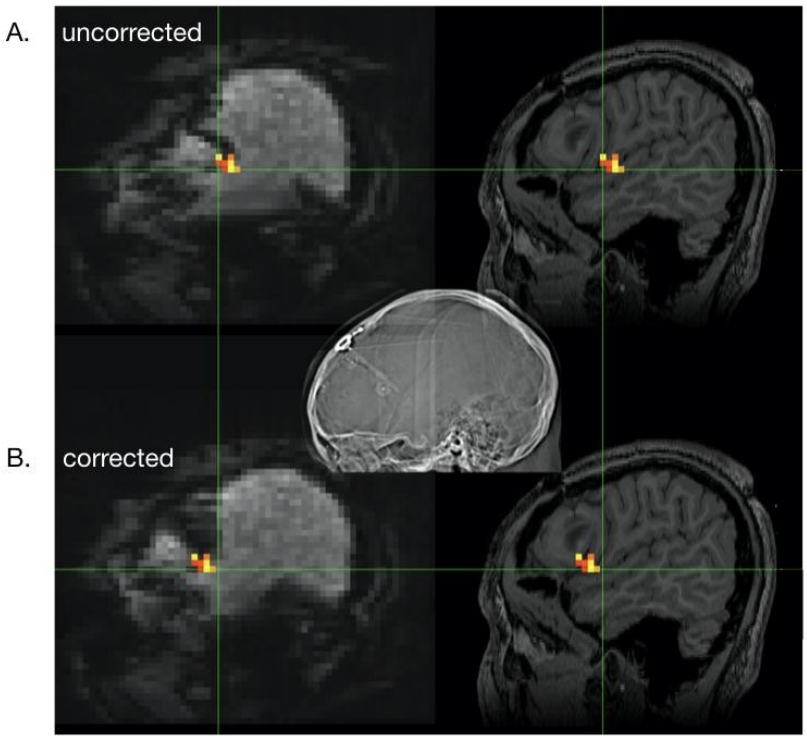


Fig 7: Displacement of fMRI task-detected activations due to echo planar imaging distortions. *Legend: A. Gradient-echo echo-planar images (GE-EPI) - the most common fMRI sequence - are most often acquired with a phase encode direction from anterior to posterior (as illustrated here), resulting in a typical anatomical deflection of the brain inwards (i.e., in the posterior direction, top left image, see distorted image anatomy at the anterior frontal and temporal lobes). Areas of signal loss (e.g., around the paranasal sinuses and temporal bones, or from metallic abrasions/craniofix of previous trepanations, central X-ray inset revealing the previous osteoplastic trepanation and Rickham catheter with Ommaya reservoir) cannot be recovered, but several methods exist to 'correct' for anatomical distortions. If fMRI task activations are overlaid onto the (non-distorted) anatomical scan for surgical neuronavigation, these EPI distortions, if not corrected for, can create large mismatches in where the activations appear to*

be located (compare location of the fMRI task activations and crosshair in the distortion-uncorrected top right image to location of the fMRI activations / crosshair in the distortion-corrected bottom right image). B. Illustration of the effect of distortion correction on the localization of language task-related fMRI activation results. Note the location of task-related activation relative to the lesion is substantially different: the uncorrected images imply activation is located at least 1 gyrus backward from the lesion, whereas the appropriately corrected images show intraoperative stimulation confirmed language task-related activation directly inferior to the lesion, in this case a left frontal ganglioglioma. Note that distortions were particularly pronounced due to prior surgery with craniofix and Ommaya reservoir implantation.

3.3.2. Head motion and motion correction

Clinical populations tend to move more during scanning than healthy research participants. Language tasks, especially those involving overt speech, are particularly prone to stimulus or task correlated motion. Objective criteria for how much head motion is “too much” cannot be easily established, but in the worst case, head motion can make brain activation uninterpretable. The direction, frequency and temporal correlation of head motion with the task can be more important for the usability of fMRI data than the absolute amount of movement, and various strategies to ‘correct’ head motion exist [240].

Correction methods for head motion can be prospective, i.e., applied by the scanner at the time of acquisition, or retrospective, i.e., applied in post-acquisition analysis. It is also possible to apply both. Prospective correction realigns the scanning planes to reduce motion at acquisition [240–242], and some vendors have integrated this into their EPI pulse sequence and image reconstruction algorithms. Aside from specific MR physics effects (such as motion-by-susceptibility interactions, spin history effects), prospective motion correction is not detrimental to the data per se. **Using prospective motion correction, however, does not remove the need to consider additional retrospective motion correction** to account for residual head motion in the data. Retrospective motion correction is an off-line data preprocessing step, which realigns the acquired images to minimize the changes due to motion. Additional analysis methods/options also exist to reduce motion changes that are not

based on geometric realignment, but on modifying the intensities in certain voxels, such as motion parameter regression, or at certain time-points, motion outlier removal. These methods tend to reduce false positives at the expense of increasing false negatives, and so should be used with caution.

The best and most effective approach for dealing with head motion is prevention. Prevention of problematic levels / types of head motion is often readily achievable through careful preparation of the patient and performing real-time inspection of the images while these are being acquired so that scans can be repeated, if required, before the patient leaves the scanner. **Motion correction practices vary widely.** The consensus of this group is that **neither prospective nor retrospective motion correction should be assumed to fix all issues arising from head motion.** Since motion always has some impact on individual activation maps [243], it is recommended that **subject head motion should be assessed in each fMRI scan through a variety of means**. This includes inspection of the raw images and retrospective motion correction / realignment plots, as well as performing multiple analyses to compare the effects of different motion correction choices on the resulting activation maps (see Appendix E for more in-depth consideration).

3.3.3 Brain extraction and statistical mask generation

Some data pre-processing pipelines remove non-brain tissue (eyes, orbits, skull and dura) from the images. Problematically, this step can have the unintended effect of also removing low-signal lesions (such as intra-/peritumoral hemorrhages, Zabramski type IV cavernomas or flow-voids in AVMs) and their perilesional areas (due to “blooming” of signal loss in T2*-weighted EPI), with the important consequence of excluding these areas of interest from later statistical analysis [244]. Therefore, if brain extraction/mask generation is part of the initial processing pathway, it is crucial to **verify that brain extraction has not also ‘removed’ the pathological lesion and the area around it from the analysis** and, if necessary, edit or replace the brain mask.

3.3.4 Spatial Smoothing

Spatial smoothing is often beneficial in fMRI, where SNR tends to be low; smoothing leads to noticeable benefits in statistical power, or the ability to distinguish real activations from noise. However, this is only true if the extent of the smoothing is less than the size of the activations; large amounts of smoothing reduce the ability to detect smaller activations (Fig 8) and can spatially displace the focus of detected activations [245].

All fMRI data possess an intrinsic smoothness. A rule of thumb is to additionally smooth fMRI data by a Gaussian kernel with size (measured at the full-width at half maximum, FWHM) of 1.5 to 2 times the acquired voxel dimension [246]. To retain spatial specificity in activation maps for presurgical clinical fMRI applications, the consensus of this Working Group is to **avoid or minimize smoothing, using no more than 1 to 2 times the voxel dimensions** [247], **up to a maximum of 5mm** isotropic FWHM (for a 2 to 3mm voxel dimension), **but ideally 4mm**. This recommendation minimizes the chances of blurring together noncontiguous cortical speech processing areas (e.g., for word production versus word hearing) identified 4mm apart using high density subdural electrode grids [248]. This recommendation is in keeping with the 4mm optimal smoothing filter identified by Pajula & Tohka [246] (for data acquired with 2mm isotropic voxels). This approach furthermore aims to balance smoothing as little as possible while retaining conspicuity of activations of interest (which can be more difficult to identify in unsmoothed data).

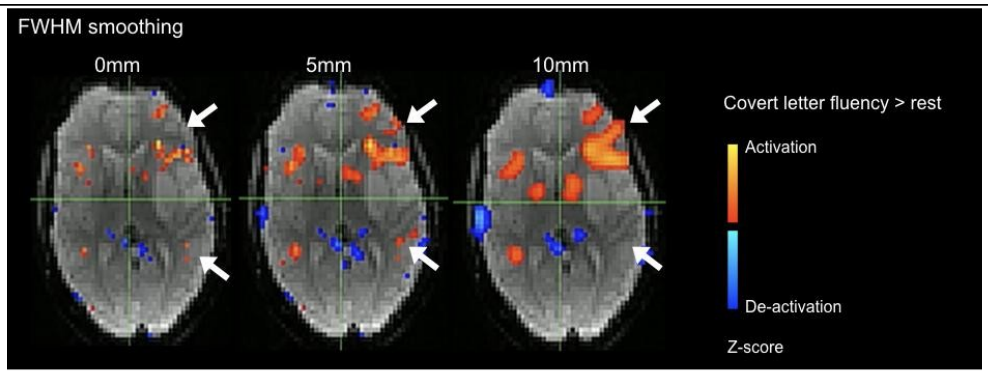


Fig 8. Effect of spatial soothing on statistical activation maps. *Legend. a. Effect of varying spatial smoothing on an example word generation language activation map, acquired at a*

voxel resolution of 2x2x2mm, illustrated in a patient with a left fronto-insular glioma. The amount of smoothing was varied from 0 to 10mm while keeping all other analysis steps constant. The resulting spatial maps (each presented at the same statistical threshold) show substantial influence of the choice of smoothing on the spatial extent and foci of activation (e.g., white arrows). FWHM: Full width at half maximum.

3.3.5 Registration of clinical FMRI and high-resolution structural images

When the clinical question is to determine proximity of language-critical sites to a surgical target, fMRI results need to be fused to the most appropriate anatomical scan to visualize the lesion and important anatomical landmarks. A variety of registration methods are available (Appendix E) and, as mentioned, should include EPI distortion correction whenever available. Among these methods, **boundary-based registration is advocated** when available/accessible (see Appendix E, Fig S3). Instead, when the objective of language fMRI mapping is to predict interhemispheric dominance, laterality calculations are best performed on the untransformed fMRI data. This is because functional-to-structural registration involves aligning and up-sampling the low-resolution fMRI data.

Transfer of final corrected and registered fMRI results into neuronavigation systems can be tricky because these systems usually possess limited data import functionalities. Because of rapidly evolving software, recommendations for specific neuronavigation workflows cannot be offered. Key is that, **whenever fMRI results are overlaid onto an anatomical scan, it is important to be aware of (and communicate) any limitations in the underlying fMRI data**. Specifically, areas of susceptibility-related signal loss in the fMRI images (often affecting the inferolateral temporal lobe) can lead to false-negative interpretations (for example in relation to a visual naming or reading task) when only the statistical results overlaid onto anatomical images are reviewed (highlighted in Appendix E Fig S3). Such misinterpretations can usually be avoided by **first inspecting results in the original fMRI data space,** where areas of distortion and signal loss are directly visible.

3.3.6 Data analysis and statistical inference

Finally, and crucially, meaningful clinical interpretation of fMRI data requires a good understanding of the practical application and associated pitfalls of any chosen processing pathway used for generating fMRI results. This does not mean that a full understanding of all the mathematical details is necessary. However, pragmatic knowledge of the advantages and pitfalls associated with the different steps is needed.

Several commercial fMRI analysis packages exist specifically for clinical fMRI analysis. These typically do not offer the user (m)any options to vary (or verify) parameters relating to statistical thresholding and multiple comparison correction. Therefore, in the wider community and among this Working Group, freely available research packages are commonly used, alongside clinically-licensed ones, because of the added benefits these tools provide [4,7]. However, research tools are typically not specifically approved for clinical use, and **fMRI analysis packages developed for research use (on which many commercial packages are modeled) are optimized for a different use**. An appreciation of the rationale behind analysis settings is therefore important to understand why it **is suboptimal to apply the default recommended settings in some existing *research* fMRI analyses to *clinical* applications.**

3.3.6.1 Hypothesis testing: model-based and data-driven approaches

The most common approach to analyze fMRI data - including clinical fMRI - is hypothesis-driven general linear model (GLM) fitting. A simple GLM implementation is available on many scanners and can be used to monitor fMRI scans in real-time. Real-time monitoring of activation maps while the patient is on the table can ensure that the collected data will be high yield and contain useful information for producing final fMRI activation maps, as well as allowing the scan to be repeated with renewed instructions, if necessary, before the patient leaves the scanner.

Normally, the GLM assumes that the hemodynamic response function (to a given stimulus) will be the same across the entire brain. However, we know that even in healthy individuals, the hemodynamic response function varies both between people and between brain regions [249–252]. This variability is even more pronounced in patients,

for example those with epilepsy [253], stroke [254] and arterial stenoses [255]. For this reason, some members of this Working Group routinely supplement GLM-based analyses with data-driven spatial independent component analysis (ICA), which does not impose the same assumptions (see Appendix E). These are available as plug-ins to some clinically approved software packages. In data-driven ICA analyses, activations and de-activations generally separate into different components, whereas for the GLM, additional steps are needed to model both positive and negative task contrasts (especially during resting blocks) and to explore a range of hemodynamic response functions (or extra parameters such as dispersion derivatives or HRF basis sets). A limitation of data-driven analysis is that the meaning (i.e., relevance to language functions) of the resulting spatial components, which typically are numerous, must be interpreted post hoc, either by subjective inspection or by matching to prior probabilistic templates of uncertain validity. In addition to testing for 'activations', both 'de-activations' and/or temporally delayed activations should also be considered, especially in the context of pathologies that can affect blood flow (e.g., high flow AVMs or neoangiogenesis in or around a tumor).

3.3.6.2. Effect size maps, thresholding and inference

Statistical thresholding and inference constitute a major challenge for clinical fMRI. Thresholding turns fMRI results into maps of activation (or no activation), which guides the neurosurgical decision as to what tissue is potentially resectable (subject to intra-operative confirmation [256]) or should be avoided. However, **currently, there is no standard or even common approach that guides how to best threshold statistical fMRI maps for clinical purposes**.

The first key step is to decide–among all the signals detected during the fMRI scan–which of the brain voxels or clusters of voxels show a signal that we are confident is related to the task (see Appendix E for a more in-depth discussion). Selecting the appropriate statistical criteria to answer this question depends on assumptions about the (temporal onset, magnitude and extent of) fMRI signal measured during a task, as well as choices about when we have statistical confidence in the detected activations (i.e., are willing to reject the null hypothesis that there is 'no activation' in this patient's brain). False negative (FN) results – where actual areas of language activity

are erroneously ‘missed’ due to statistical choices – generally pose the highest concern in pre-surgical applications of fMRI. **A false negative in an fMRI map that is not interpreted with appropriate caution might result in surgical removal of an area of cortex that is crucial for language abilities**. Conversely, a false positive (FP) result – indicating more areas of activation than are truly there – could unnecessarily prevent surgery or maximal resection. The first scenario (FNs) on balance carries a higher risk of direct patient harm. Consequently, choices in the statistical analysis of clinical fMRI data *should* be balanced towards **avoiding FN** (i.e., type II) errors. However, most of the widely used fMRI analysis packages are based on classical statistical inference, which focuses on controlling FP rates, and are therefore not optimized (nor fully adequate) for clinical fMRI. This assumption, i.e., null hypothesis testing in general, disregards prior clinical knowledge about patient performance (and deficits) and appears to ask the wrong statistical question (“Where can the null hypothesis of no activation not be maintained, at sufficient protection from false-positives”?) from the outset in the clinical context. Several alternative statistical approaches are available that aim to control the balance between FNs and FPs (e.g., [257–261]). However, these alternative methods still require broader validation on a range of datasets from multiple patient groups, preferably where intraoperative ESM data have been obtained and especially where the post-surgical clinical performance outcome is known. **This working group particularly encourages the collection and pooling of such data across sites.**

Once the task activation map (identifying which voxels/clusters show task-related changes) has been generated, a second decision arises *whether* and *how* to threshold the resulting image. The magnitude of fMRI activation can vary substantially across individuals. This problem is amplified in clinical practice, because numerous factors affect activation magnitude, including the effects of the lesion itself on blood flow; levels of performance/impairment [262]; head motion and the presence of sedating medications [263], just to name a few. Therefore, appropriate thresholds differ for different individuals [264] and **choices need to be tailored to each patient**. Consequently, definite guidance as to how to choose the ‘appropriate’ threshold cannot provided. **Our group recommends inspecting a range of statistical thresholds as well as the un-thresholded results** both for every activation map, as well as task de-activations (where the same thresholding considerations apply) in and around areas of interest for presurgical mapping. As a general principle, we

recommend that established “institutional approaches” are reported clearly and evaluated rigorously against all available clinical data, most especially post-operative language outcomes.

3.3.7 Cerebrovascular reactivity mapping

One concern raised about clinical fMRI is that pathologies may disrupt cerebrovascular reactivity, and, consequently, the BOLD signal. For example, non-negligible methodological considerations arise in the context of pathological lesions (including high-grade [265] as well as low-grade [266] brain tumors and AVMs [267] that may reduce BOLD signal and therefore the sensitivity of fMRI to localize areas with language functions (Appendix A).

Cerebrovascular reactivity mapping (CVRM) has been promoted as an emerging standard of clinical care for presurgical assessment of peri- and intralesional BOLD responsiveness prior, or in addition, to conventional fMRI [268]. CVRM promises to identify areas at high risk for false negative findings by conventional fMRI and, thereby, patients requiring awake surgery with intra-operative ESM. However, CVRM suffers some interpretative limitations. For example, CVRM aims to detect neurovascular uncoupling, but is itself based upon uncoupling cerebral blood flow (CBF) from the cerebral metabolic rate for oxygen ($CMRO_2$) by vasodilatory changes (such as breath holding or carbogen inhalation) and does not assess neurovascular coupling directly. Furthermore, a “lack of CVRM activation” is difficult to ascertain, both from the signal as well as statistical perspective. At present, it is not known under which conditions, and to what extent, neuronal BOLD responses may dissociate from hypercapnia induced CVRM. It seems of crucial interest to systematically investigate whether fMRI responses evoked by neuronal activity may be preserved while the response to hypercapnia is lacking or attenuated. If cases with such dissociation exist, absent cerebrovascular reactivity in and around a lesion should not preclude the patient from speech and language fMRI prior to resective surgery; rather, it should guide the interpretation, to give appropriate caution regarding the lack of sensitivity. Complementarity to fMRI speech and language mapping is further discussed in Appendix F.

3.3.8 Calculating laterality

Theoretical challenges to determining language laterality were discussed in section *2.2.1/ Box 1*. Asymmetry in brain structure and function is a topic of much research, and best practices in calculating laterality both in general and for clinical practice are still undefined [53]. From a practical (analysis) perspective, the outcome of language laterality calculations with fMRI can change depending on the statistical threshold employed [269,270]. Using a single fixed statistical threshold (i.e., not tailored to the individual) increases a risk of suboptimal or inaccurate assessment of fMRI language lateralization [271–273], especially when cognitive deficits or pathological features may be present that can reduce the overall magnitude (or height) of BOLD fMRI signals [269]. Since language laterality is a continuum [274,275], the question of how to determine laterality based on fMRI is an important methodological challenge.

Many methodological variations to analyze fMRI language laterality have been proposed and evaluated against the Wada test (e.g., [22,50,272,276]). However, the Wada test is itself not infallible (see [43,277,278] and the references therein). Consequently, the "ground truth" of language organization in the individual patient may not be absolutely known, with the result that the superiority of any one method of calculating a fMRI laterality index (LI) over the others remains undetermined. There are therefore still no agreed standards as to the optimal way to calculate LI [57]. Among this Working Group, several groups do not formally calculate LIs. Other practitioners feel that clinical language lateralization should be based on a numerically computed LI, or a combination of visual inspection and a LI [53]. **If LIs are generated, it is generally undisputed that LIs calculated at a single standard statistical threshold are not adequate** [53], especially for clinical use. Additionally, a single 'global' LI is likely to be misleading in patients who have mixed dominance (e.g., anterior and posterior language areas in different hemispheres). In such cases, at any given threshold, laterality indices may be biased towards the non-affected hemisphere, limiting their utility. Consequently, when the clinical question is purely to establish relative lateralization (i.e., a presumed functional 'reserve') to inform discussions around surgery, our recommendations are:

1. To **include an appropriate range of tasks (and carefully consider task contrasts)** to establish lateralization for the components of speech and language of greatest surgical relevance.

2. If LIs are calculated, **visualize a range of LIs tailored to the patient's** levels (i.e., signal magnitude) of **activation, or** employ an LI calculation method that produces **a weighted average across a range of thresholds** (e.g., [279]**,** see other approaches described in [53])**.**
3. **Consider LIs from different language-related regions (not just the frontal lobe),** but excluding regions involved in sensory (e.g., visual or auditory) or cognitive processes not specific to language (e.g., frontal eye fields, dorsolateral prefrontal cortex, anterior cingulate cortex). This is because task-specific regions of interest can produce very different results from large hemispheric LIs [52].
4. **If the fMRI results are categorized** (i.e., into 'left-lateralized', 'right-lateralized' or 'mixed/bilateral'), **indicate how these categories were defined** (e.g., based on arbitrary LI cut-offs, or by comparison to specific population norms). Caution is advised in how atypical laterality is interpreted, noting high variations among experts in the meaning of bilateral/mixed and atypical dominance [53]. Some clinicians might conclude from a non-nuanced fMRI report of a 'mixed' or 'bilateral' LI that surgery on the left hemisphere is without risk (for example, interpreting this result as showing that both hemispheres are functionally equal), whereas such an interpretation is likely neither intended nor indicated, especially in patients with existing language impairments.

[Table 3]

3.4 Reporting fMRI results

Writing a qualified report is crucial for translating clinical fMRI exams into medically relevant conclusions. Central to clinical fMRI report generation is the concept that **images are powerful**. Great care should be taken to minimize misinterpretation, especially in terms of an apparent absence of activation near a susceptibility artifact, or when superimposing fMRI data onto structural images in which areas of artifact may be less prominent or invisible. Specific emphasis should be placed on limitations of a particular exam (e.g., regarding patient head motion, impairments, etc). It is the opinion of this Working Group **that if the person interpreting the data does not believe the report to be valid, then no images should be generated at all**.

Minimum fMRI reporting checklists and best practices have been proposed for the neuroscience community [18] and complement existing guidelines for BOLD-fMRI dictation issued by the ASFNR (https://www.asfnr.org/wp-content/uploads/BOLD-fMRI-Dictation-Guidelines.pdf). Our Working Group recommends specific additions relating to the description of tasks and analysis methods used. In general, the clinical fMRI report should be concise and contain the following: 1. indication, 2. patient history and condition, including any known performance limitations 3. paradigm selection along with special adaptations for the patient, 4. Data assessment, including quality control (e.g., EPI and anatomical scan co-registrations, head motion statistics) and analysis technique, 5. findings & conclusions (including any interpretational confounds such as due to performance limitations), 6. Disclaimer (Appendix G). Results of complementary data including tractography or perfusion studies should ideally be integrated with the fMRI report.

In addition to the written report, **oral presentation of the fMRI findings, i.e., at the regular surgical team meeting**, offers valuable opportunities to highlight confidence and specific limitations of the results in an individual patient's scan. This is recommended wherever possible and routinely done in epilepsy surgery programs (~80% of centers [4]). The individual presenting results must be able to clearly and simply balance: 1. the strengths and limitations of the tasks and approach chosen to address the clinical question (e.g., known sensitivity / specificity for lateralization and localization); 2. the strengths and limitations of the specific study presented (e.g., domains assessed vs not assessed), 3. the confidence a surgical team can have in a given set of fMRI results (e.g., due to data quality, modifications based on patient performance, etc.). **Repeatedly failing to manage this balance can result in a surgical team losing confidence in fMRI** and (e.g.) over-using Wada testing, with its attendant risks, or conversely over-interpreting data to, for example, draw pseudo-precise surgical margins in a non-evidence-based manner. While evidence on the use of fMRI to guide surgical margins is still accruing, 44% of epilepsy programs already use fMRI in this way, not removing cortex within a given distance of language fMRI activation [61]. Surveyed programs reported using precise but different distance limits as a guide, which varied from 3mm to 50mm (mode 10mm). While some programs use fMRI in this manner given skilled clinicians, extensive experience, and cautious

interpretation, this is not likely to be accurate in most (if not all) cases, because **distance calculations in fMRI are intrinsically related to preprocessing steps** (especially smoothing) and statistical choices made by the user, rather than reflecting true functional boundaries.

4. Discussion

4.1. Summary

This document presents the summary best practice recommendations from the OHBM Working Group on clinical language fMRI mapping. Our approach to generating these recommendations was based not only on the consensus and experience of this multidisciplinary committee spanning the range of relevant expertise, but was also broadened to consider the weight of evidence in the literature available behind the clinical decision-making. Previous groups have reviewed a different range of clinical fMRI language applications and found different levels of validation for them. For example, the level of validation was found to be higher for predicting language outcomes based on fMRI laterality assessments than for preventing deficits with MRI localization and tailored surgery, at least in epilepsy [5]. Here, we further consider language tasks and task designs optimized according to specific clinical objectives. In doing so, we focus on specific patient populations (e.g., adult/pediatric, with/without cognitive impairment) and possible modifications for their particular needs. Putting forward recommendations for optimal tasks, task conditions and modifications based on the strongest available data remains challenging. A major hindrance is the absence of randomized controlled trials of language fMRI applied to surgical populations. A second obstacle is the sparsity of studies conducting head-to-head comparisons of different tasks and their performance when predicting post-operative language outcomes. There remains, therefore, a long way to go. Given these challenges, we place emphasis on *language processes* that should inform the selection of task fMRI, focusing on data that survives meta-analysis (predominantly for language lateralization) and converging lines of evidence from research neuroscience studies and surgical lesion outcome data (in the case of localization). We further expand on these considerations to put forward practical guidance, based on state-of-the-art in neuroimaging science, for all aspects of clinical fMRI in relation to acquisition and analysis of individual patient data. Of course, optimal language task selections and fMRI approaches remain subject to ongoing refinements, as our understanding of the

basis of language in the brain, as well as technical approaches, continue to evolve. These recommendations therefore do not aim to be final or prescriptive. Instead, our objective is to provide practical steps towards reducing the wide heterogeneity in how fMRI language mapping is performed. In this way, this document aims to improve minimum standards and facilitate the objective assessment and quantification of the benefits, efficacy and limits of high-quality clinical fMRI. To achieve the latter requires thoughtful application to ensure that fMRI is:

a) performed in a multidisciplinary manner, based on careful definition of individual patient characteristics and performance abilities

b) backed by converging neuroscientific and clinical data

c) follows best practices in data acquisition, processing and analysis

d) is interpreted and used by the operating neurosurgeon based on interdisciplinary consultation (i.e., in according with / as intended by those who conducted and reported the fMRI)

e) systematically evaluated against clinical outcomes to determine added value for predicting outcomes and minimizing language-related risks

In this process, a clear need was identified for the wider reporting of individual groups' experience in relation to fMRI, including relevant factors such as amount of ESM mapping required and duration of awake surgery based on fMRI predictions, and crucially, language outcomes. A wide variability in practice of fMRI for clinical language mapping was evident amongst this Working Group. Naturally, it is not possible (nor desirable) to be overly prescriptive in recommendations, as the ultimate aim is to move towards agreement in the community. Towards this goal, some identified avenues for targeted developments and research, needed for language fMRI to substantiate and improve its clinical utility, propagation and availability, are detailed in Box 6.

Select areas for further evaluation and research	**• BOX 6 •**
1. For optimal choices in tasks and task designs	
➢ In-depth **comparisons of activation maps derived from different well tested tasks** and task designs that aim to engage similar language	

processes

- **Evaluation** of the benefits **of newly emerging tasks** probing higher complexity in language processing compared to well-established simpler tasks for the merits in more comprehensively engaging wider language networks
- Greater **clinical evidence** supporting the **specificity** of particular 'activated' neuronal populations for a given language process
- Patient preferences and useability of different **task adaptations** (such as different approaches to varying task demands, difficulty levels and timing)
- Development of **naming stimuli** that consider **different ages and cultural backgrounds**, including **appropriate norms in pediatric populations**

2. For improvements in acquisition, analysis and interpretation of fMRI

- Objective evaluation of different approaches to visualise fMRI results (e.g. levels of statistical thresholding). **Additional development of alternative thresholding approaches** that offer variable control not only over false positive, but importantly also over potential false negative findings.
- Transparent reporting of clinical evidence testing the predictive value of specific fMRI task designs for language outcomes, especially in patients undergoing tailored lesionectomies. Crucial for such evidence gathering is the implementation and systematic **collection of standardised and longitudinal (pre- and post-operative) language and neurocognitive testing batteries** across all patient populations routinely undergoing clinical fMRI. Specific value would be gained from pooling of data, especially across sites using fMRI alongside focal disruption techniques such as navigated TMS and / or ESM. Such endeavours are currently hindered by restrictions surrounding sharing of data collected in clinical settings, and by resources. Resource considerations arise in relation to both direct costs, and that of staff needed to collect standardised (potentially longer or different) datasets in clinical settings. Despite these hurdles, multi-centre trials remain needed to objectively characterise the relative impact of different techniques in use today on comprehensively characterised language outcomes, including patient reported quality of life.

4.2. Working Group's position on the utility of clinical fMRI

An estimated 30-50% of European neurooncological centers [55,80] and most (>90%) epilepsy surgical programs worldwide [61] employ fMRI. Yet, justified uncertainty persists around fMRI's ability to localize functions with the precision required for surgical planning. There are inherent constraints on the spatial precision that fMRI can achieve

based on the limitations of the underlying BOLD contrast [280]. Frequently cited confounds include the difficulty dissociating task-associated from language-essential neural activity with fMRI, and consequent variable sensitivity and specificity of fMRI relative to 'virtual lesion' techniques [281].

The existing evidence base in epilepsy surgery candidates demonstrates the suitability of using fMRI as a surrogate for Wada testing to establish language dominance. As long as care was taken with task selection and design, fMRI has shown ability to predict language outcome after surgery to some degree. Measuring the effectiveness of language mapping in these scenarios is currently difficult, perhaps impossible, in clinical studies in which the maps are used to adjust the surgical procedure, since there is then no comparison with how the surgery would have been performed without the fMRI information or what the outcome would have been. As a result, the literature consists of mainly uncontrolled observations and the occasional comparison to an historical control group.

Cases of failure to prevent language declines are likely to exist according to a recent clinical survey reporting the relatively common practice of resection of fMRI 'activations' [61]. Here, 17% of epilepsy surgical programs reported one or more cases where all language fMRI-positive activation was preserved, but a patient still suffered post-operative language decline (noting a possible contribution of subcortical tract damage). None of these had been published. According to a recent survey on post-surgical care practices in Europe, not all neurosurgical centers refer each of their patients for neuropsychological assessments to evaluate language performance / outcomes. Furthermore, some patient groups are particularly unlikely to get a referral for language evaluation. For example, only 3% of all patients with high-grade gliomas were offered language assessments after brain surgery, in comparison to 30% of individuals with low-grade gliomas [282]. Numerous practical challenges clearly arise in the neuropsychological scheduling and follow-up of patients who are on a rapid treatment pathway or due to undergo post-surgical adjuvant treatment. This appears to be less problematic in an epilepsy setting, where 70% of patients receive follow up [61]. However, **more outcome reporting is much needed to help inform what clinical questions fMRI is suited, or not suited to answer, and potential reasons for fMRI failures, if these are encountered.**

In the opinion of this Working Group and growing evidence of its predictive power, language fMRI mapping has the potential to add substantial value in the neurosurgical selection, consenting and planning of appropriately chosen patients. However, clinically meaningful fMRI mapping hinges on the precise identification of the surgical questions that fMRI is asked to inform, alongside careful characterization of patients to minimize fMRI studies in patients who are unlikely to benefit. In this Group's experience, requests for fMRI from inexperienced referrers often consist of open requests to "please map eloquent cortex" without further precision in what is to be mapped or what the results will be used for. Yet, it adds another layer of challenge for fMRI to provide 'results' when the question is poorly (or even not at all) defined. Additionally, practitioners who unintentionally focus on different brain regions in analysis (e.g., frontal vs temporal) can generate different language maps from the same data [61]. Detailed knowledge and training are needed [283] to evaluate fMRI results in the absence of "ground truth", and meaningful clinical fMRI requires a highly interdisciplinary approach. When performed with due diligence and expertise, language fMRI mapping offers potential direct patient benefit. Examples of fMRI advantages include reducing the need for Wada testing [5], potentially enabling more aggressive or extensive surgical intervention when combined with tractography and ESM [284–286], and guiding optimal use of intraoperative stimulation [284,287] and/or intracranial electrode placement [182,288], especially, but not exclusively, in pediatric patients.

4.3. Requirements for the validation and standardization of fMRI language mapping

Reservations surrounding the clinical use of fMRI language mapping are derived largely from comparisons with the Wada procedure (for language lateralization) and ESM (for localization). Previous reviews have summarized the highly variable rates of sensitivity and specificity reported when fMRI results have been compared to these gold standard techniques [22,44,289,290]. There are reasons to be cautious about interpreting such direct evaluations, however, not only because of factors such as brain shift upon opening the skull, but also because of differences in what each technique reflects and how results from each tool are derived.

ESM is considered a gold standard tool for the preservation of brain function during surgery. However, like all techniques, ESM has limitations. Not all patients are able to tolerate [145,291] or complete it [292], and there remains little standardization in both how ESM is executed [293] and how stimulation-induced language errors are interpreted [293,294]. Furthermore, trial-to-trial reproducibility can be highly variable [295]. For example, a single occurrence of anomia following stimulation does not indicate essential cortex (i.e., is not predictive of post-surgical language deterioration [296]. Consequently, stimulation sites must produce an error on three non-consecutive attempts to be considered reliable (i.e., identify critical cortex) [297]. Whether ESM effects are truly focal or potentially reflect activation of a wider network [298] is also under investigation [299]. Importantly, 40% of epilepsy programs report cases where the preservation of all ESM-positive language cortex has failed to prevent enduring language deficits [300]. Similarly, although the use of ESM reduces the rate of long-term neurological complications after tumor surgery, a small percentage of permanent deficits remains (generally around 3.4% overall [301], but higher in critical locations such as the inferior parietal cortex (16.7% in one study [132]). **Alternative methods, able to *pre*-operatively predict surgical risk and the likely safe(st) surgical approach, therefore retain an important role in neurosurgical planning**. Among such methods, fMRI is most used, both in epilepsy [61] and - to a lesser extent - in neuro-oncology surgery settings [54,80], likely because it offers the greatest amount of spatial detail and does so non-invasively. fMRI can capture the entire brain and may have the potential to predict resilience of the language network to permanent deficits prior to the actual surgery which would be a significant advantage over ESM. However, fMRI specificity is per se limited compared to ESM. As mentioned, practitioners adopt statistical choices that aim to reduce the chance of false negative findings in task activation maps often acquired without knowledge of what will be tested intra-operatively. These choices directly influence comparability to ESM; more areas are typically activated in fMRI than are associated with language errors during ESM, resulting in only modest overall specificity (approximately 55%-71%) when grouping across populations and approaches [290,302]. Specificity increased to >80%, however, when adequately statistically powered and when visual ‘expressive’ language tasks (presumably matching intraoperative procedures) were used and analyzed using higher statistical cut-offs. Conversely, specificity decreased to under 30% using short, auditory, ‘receptive’ language fMRI tasks [290].

It is important to contextualize these results: by temporarily evoking or inhibiting behavioral responses, **Wada testing and ESM assess language function in fundamentally different ways from fMRI**. Often, pre-operative fMRI paradigms do not probe the same components of language, or do so to different degrees than Wada and intra-operative ESM. For example, silent word generation tasks optimized to control for non-linguistic motoric aspects of word generation will by design *not* predict sites that cause speech arrest when stimulated intra-operatively [81]. Instead, fMRI offers distinct advantages to evaluate crossed dominance for different aspects of language that are difficult to evaluate through Wada testing, and to explore language processing at the systems-level, including contributions of contralateral brain structures that are inaccessible to ESM as ESM is only performed in the operated hemisphere. Importantly, the use of Wada or (positive, i.e., error-inducing) ESM as mentioned does not preclude (long-term) postoperative deficits (e.g., [132,281,303,304]). Still, the expectation – wrongly – prevails that correspondence between these techniques must first achieve 100% for fMRI to be useful as a clinical tool. Like others before us, we advocate that **the clinical usefulness of fMRI should be evaluated in terms of its ability to anticipate and minimize (further) post-operative language decline, especially for patients unable to tolerate or complete awake surgery** [145].

Preoperative language fMRI mapping predicts postoperative performance on tests of naming [5,305] and verbal memory [5,305,306] in patients following dominant temporal lobe surgery for epilepsy. The predictive value for post-surgical outcomes appears superior for language laterality as determined by fMRI than by Wada [5,306,307], although these studies are hampered by varying approaches for interpreting the Wada, typically ignoring mixed dominance possibilities. In small case series, post-operative language deficits arose in patients whose resection area spatially overlapped with pre-operative fMRI activations [308–310] even when no overlap occurred with positive ESM sites [304]. However, diverging results have also been found [311]. A recent study reported better 3-year survival rates in high grade (but not low grade) glioma and metastasis patients operated with fMRI than without fMRI, irrespective of the use of ESM [312]. This result was ascribed to greater surgeon confidence in undertaking extensive resections when guided by fMRI, as also highlighted in previous reports [281]. The only existing prospective randomized trial found no difference in AVM microsurgical complications whether using fMRI or not [313], though this study was limited by not pre-selecting patients with an AVM in proximity to critical cortex. Finally, fMRI language activations

in the left auditory cortex and right cerebellum before pediatric cochlear implantation (CI) predicted language outcome at 2 years with 93.8% accuracy, surpassing the accuracy of established factors including age at treatment and pre-implant hearing thresholds [314].

Thus, studies examining the predictive value of fMRI for avoiding postoperative language decline (or effectiveness of CI) generally support fMRI's complementary role in surgical planning. They reinforce the notion that **no technique on its own is infallible** and **the best patient outcomes likely result from a combined approach** [304]. Nevertheless, reliability of speech and language fMRI mapping depends on the nature of the pathology and is likely to differ between epileptic lesions, brain tumors and vascular malformations. In high-grade brain tumors and high-flow AVMs, pathological vessels *per se* may only contribute within different limits to the oxygen supply of the parenchyma, somewhat restricting the use of fMRI (as well as Wada testing). Consequently, more comprehensive longitudinal evaluations of the applications that fMRI is particularly valuable for, and where its prediction failed, remain much needed.

4.4. Conclusions and vision(s) for improving clinical fMRI

In the experience of this multi-national and multidisciplinary team, fMRI is a clinically useful and informative exam for presurgical planning and intra-operative use to minimize debilitating (and cost intensive) language deficits in carefully selected patients, as an alternative to Wada, and to guide ESM. However, **international standardization of practice and reporting is needed in order to direct large-scale studies** to objectively evaluate the long-term benefits brought by clinical language fMRI across the expanse of neurosurgical indications. In this context, both emerging technological advances that offer higher resolution acquisitions and the adoption of sophisticated preprocessing and analysis methods from the research realm into clinical practice hold promise for more precise localization of clinical fMRI results. Wider evaluations are needed in different patient populations using validation methods adapted to the clinical question asked of fMRI. Additional research is advocated to comparing the performance of individual tasks head-to-head in terms of the strength and robustness of their associated activation patterns, and reliability in the context of patient performance difficulties. **A particular limitation in validating language**

fMRI mapping is the sparsity of data in patients with atypical language organization. Due to the rarity of atypical language dominance, reports to date have largely consisted of case studies or very small case-series from which it is difficult to draw firm conclusions as to the accuracy and limitations of fMRI. Thirdly, very few studies have investigated the impact of common medications on the BOLD signal. Yet, preliminary studies of antiepileptic and antiedematous drugs (e.g., [315–317]) highlight the need to better understand how common pharmacotherapies and interactions may adversely affect BOLD fMRI. Finally, since all clinical tests carry some level of error, further consideration should be given to the level of evidence required to evaluate fMRI as a quasi “gold standard” for select indications. We propose this discussion should move towards detailed evaluation of pre- vs post-operative speech and language outcomes as determined through patient-reported and validated, standardized neuropsychological outcome measures.

Tables

Table 1. Preferential roles and lesion consequences of proposed core language regions

Gyrus/ Area	Subregion	Prevalent* language process	Predicted** consequence of damage
Inferior frontal (IFG)	Pars opercularis	**Phonological processing** & articulatory control [117,318] ● cognitive aspects of speech [47,319] ● higher-order predictive processing [320,321] for syntax [110,322] and semantic processing [323]	Intraoperative stimulation: ● Most likely **anomia** [324–327] **& phonological paraphasias** [325,328–330] ● More rarely: Speech arrest [296,327,330–335], dysarthria [330], syntax errors [336], semantic paraphasias [325,329,337,338], sign language errors [339] Excision / infarct ♯ ● Reduced (phonemic) fluency [127,340–342] ● Impaired speech production, aka 'Broca's aphasia' (especially with posterior STG damage) [343,344] ● Transient speech deficits [345], anomia / dysnomia [346]
	Pars triangularis	**Semantic executive control** i.e., ● processing **when semantic demands are high** (e.g., goal-directed selection [69,105,347,348], disambiguation [349]) ● lexico-semantic integration [107] ● syntactic working memory [350]	Intraoperative stimulation, ● Most likely **Anomia** [325,327,346,351] ● Occasionally: semantic paraphasias [325,329,330,337,338], speech arrest [327,334,352], syntax errors [353] Excision / infarct ♯: ● Disrupted speech [354] and possibly comprehension [340,355,356], possibly due to disrupted semantic executive control [342,357] ● Transient language deficits [345] ● (Semantic) errors during naming [328,340] or reading [358] ● Agrammatic speech [359]
	Pars orbitalis	**Semantic executive control** [318] ● especially **controlled semantic retrieval** [347,360–362], combinatorial or associative processing [105,363,364]	Intraoperative stimulation: ● Most likely: no behavioral change [334,365] ● Uncommon: semantic paraphasias [366], anomia[367], arrest [327] Excision / infarct ♯: ● Comprehension of syntactically complex sentences [78], likely due to demands on working memory [368] ● Naming errors (with extended network damage [369]

			● Semantic errors during reading [358]
	Ventral premotor cortex (inferior precentral)	**Phonological processing** ● **Speech output / articulation** [115,325,332] (not unique to language [370])	Intraoperative stimulation: ● Most likely**: dys/anarthria** [325,327,333–335,352,365,371,372] (see [117]) ● Very rarely: anomia [325,327] Excision: ● Transiently reduced fluency [341] ● Speech production deficits [373]
Middle frontal (MFG)	Mid-to-posterior MFG, including inferior frontal junction	**Domain-general semantic processing** ● (e.g., semantic / lexical decision making) and cognitive control, including sentence level **working memory** [360,374] ● **task switching** [375] & language switching [210]	Intraoperative stimulation[†]: ● semantic errors & paraphasias [325,330,338] ● anomia [325,330,332,352,365] / action naming [326] ● phonologic errors [325,330] ● articulation [330] Excision / infarct: ● transient dysarthria [376], ● transient naming and speech difficulties [377]
	“Area 55b”; junction with precentral gyrus	**Articulation, speech fluency** [90]; integration of phonological and semantic processing streams [91]	Intraoperatively: ● Speech arrest [327] ● semantic paraphasias and anomia [91] Excision / infarct: ● apraxia of speech [90]
	“Exner’s area”; junction of SFS and precentral gyrus	**Graphemic representations / control** (transforming phonological representations into writing) [89,378]	Intraoperatively: ● **Agraphia** (writing arrest/errors) [379,380] Excision / infarct: ● Agraphia [381] / transient writing deficit [382] (see [122] for review) Phonological errors during reading [383]

Superior frontal (SFG)	pre-Supplementary Motor Area	**Initiation and sequencing of spontaneous speech** [384,385]; lexical access [385–387] and domain-general control during more executively demanding tasks [364] and internally guided tasks.	Intraoperative stimulation: ● Slowing / arrest of speech [327,362,388,389], naming/fluency errors [390] (see [391]), verb generation errors [392] Excision: ● Ataxic speech, ranging from word finding difficulties to reduced fluency and akinetic mutism [308,311,393–398], verb generation errors [392]
Insula		**Complex articulatory control** (e.g., [47,399], see [47] for review) ● general language processing [400] (cognitive control & attention [401–403]); ● possible interface between cognitive / phonetic aspects of speech and motoric aspects of articulation [404]	Intraoperative stimulation: ● Speech arrest/anarthria, anomia [405,406] (see [407]) Excision / infarct: ● (sometimes persistent [408]) **aphasia** [409–412]; mild anomia and dysarthria [412–414] , apraxia of speech [415] ● reduced fluency [127,358,399,408,416,417] ● phonological errors [369], repetition errors [418] ● comprehension deficits [419]
Inferior Parietal Lobe; TemporoParietal junction (IPL/ TPJ)		**Cognitive processes required for writing** [378,420], such as ● **Integrative / attention-driven cognitive control** [421] and goal-based action planning [422,423] ● Phonological working memory [374] ● General (non-language-specific) semantic processing [424,425]	Intraoperative stimulation: ● **Writing errors/agraphia** [426] Excision / infarct: ● Pure or apraxic **agraphia** (see [378])
	Supramarginal gyrus	**Phonological processing** [364,427,428]. Functional sub-regions for: ● Articulatory sequencing [429], including for writing [430] ● Verbal short-term memory [431–433] ● Auditory feedback monitoring / executive demands [429,434,435] / phonological temporal order prediction [108,421,436]	Intraoperative stimulation: ● Most often: **anomia** [324,325,327,330,334,365,438,439] and **phonological paraphasias / articulation** errors [330,337,352,439,440] ● Variably: semantic paraphasias [324,329], speech arrest [327], repetition errors [441], alexia [365,440,442] agraphia [120], verbal working memory errors [443] Excision / infarct: ● Word finding difficulties [132,444], mild speech apraxia / phonological errors, “Broca’s aphasia” [328,343,346,358,369,439,445] ● Verbal short-term memory deficit [446–448], complex (sentence/auditory) comprehension deficits [437]

		● Sentence or event-level concept processing / phonological-semantic integration [429,437]	● (transient) agraphia [449,450] (see [122,437] ● Occasionally transient alexia [449] ● Repetition errors [340,451]
	Angular gyrus	**Multimodal executive semantic processing & integration** [68,360,364,432,452] with subdivisions [453,454] ● Temporary 'dynamic buffer' for combinatorial-semantic processing [455–459] (e.g., integrating words into sentence-level meaning [460] / learning phonological-orthographic-meaning relationships [461]) ● Concept similarity analysis / assignment [462–464]	**Intraoperative stimulation**: ● Most often, anomia [324,334] ● Reading errors [440], verbal working memory errors [443] Excision / infarct: ● (Transient) language problems (apraxia, dysnomia, phonological errors) [465], dysnomia [132], anomic aphasia [466] ● Agraphia with alexia [467]; pure agraphia / dystypia [369,447,468] / alexia [127] / orthographic working memory impairment [469] ● Sentence /auditory comprehension deficits [78,340,343,437,470] ● Semantic paraphasias [445] / errors [127,451] ● Repetition errors [451] ● Syntax errors [359]
Lateral temporal lobe	Posterior superior temporal gyrus/STS (lateral to the 'buried' planum temporale)	**Continuous phonological processing** [471,472] (see [46,112,473])**,** including: ● During speech: auditory feedback monitoring [362] ● During comprehension (especially sentence [110]): auditory short-term memory [432,471,474] representations ● Integration of phonological, semantic and syntactic processing [360] for syntactic or semantic predictions [107,360], aka lexical-syntactic attention [463]	Intraoperative stimulation: ● **Anomia** [325–327,334,365] ● Alexia [365,440] ● **Phonological errors**/ paraphasias [330,362,440,475] / repetition errors [441,475], sometimes speech arrest [327] ● **Comprehension deficits** (receptive aphasia)[362,476] ● Semantic paraphasias [324,325,337] ● Agraphia [120] Excision / ischemia: ● **Wernicke's aphasia** (comprehension deficits [127,340,474,477]) or reduced auditory short-term memory [78,474] ● Repetition errors [127,340,341,417,418,451] ● **Broca's aphasia** (speech production/fluency/phonological errors)[117,127,340,343,417,437] ● **Naming errors** [127,340]

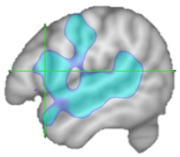
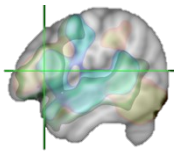
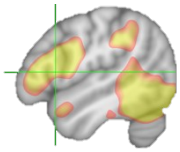
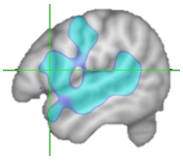
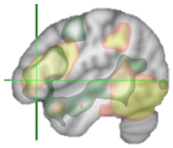
			● Syntactic difficulties [369,477,478]
	Middle STG/STS (lateral to Heschl's gyrus)	**Acoustic processing** of attended / perceived [479] meaningful [480] word sounds [481]	Intraoperative stimulation: multiple error types ● Anomia [325–327,482] ● Phonological errors [324,325,483] ● Semantic errors [484] / paraphasias [325,330,337] ● Sentence / word [476] comprehension deficits (see [483] ● Agraphia [120] Excision / ischemia: ● Comprehension deficits, especially sentence-level [78,369] ● Reduced semantic fluency [342]
	Posterior MTG	**Multimodal semantic processing and integration** ● Lexical-semantic integration, [112,445,485] ● Semantic-phonological integration [117,362,432,458], esp. when a) high levels of semantic [110,486] control are needed (e.g., auditory / sentence-level comprehension [364,437]), or b) mapping concepts when converting from orthography onto phonology [459]	Intraoperative stimulation: ● Most often: anomia/dysnomia [117,324,325,330,352,362,365,440,482] ● Additional frequent errors: semantic errors [117,324,337,476,487], phonological paraphasias [330,337], reading errors [330,440,488] ● Less reported: receptive aphasia [362], syntax errors [489] Excision / ischemia: ● Comprehension [78,416] ● Syntactic abilities [369], e.g., picture-matching sentences [490] ● Naming errors [325,369] ● Semantic paraphasias [342,445] ● Phonological paraphasias [445] / errors [358]
	Middle MTG	**Word-level semantic processing** [360] ● lexical-semantic processing [491,492] ● syntactic-semantic interface [485]	Intraoperative stimulation: ● Semantic paraphasias [324] ● Phonemic paraphasias during reading [330,440] ● Auditory single word comprehension [493] Excision / ischemia: ● Auditory comprehension [78,494] ● Naming errors [341,492] ● More rarely: semantic paraphasias and anomia [117]
	Posterior)	**Item categorization and**	Intraoperative stimulation:

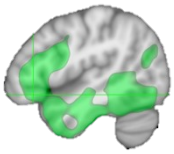
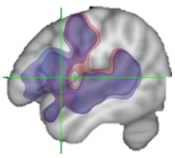
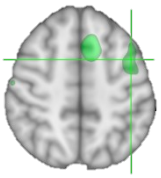
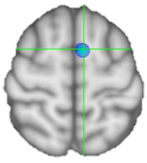
	lateral ITG	**integration** ● Early **lexical access** for words in reading & spelling (lexical-orthography) [118,122,495] ● Auditory & visual interface [360], **integrating sound** (lexical-phonology interface [362]) **and meaning** (lexical-semantic interface [117,360,496] **of words**	● **Anomia** [225,330,352,362,497–499] ● **Alexia** [330,500,501] (meaning-based (morphogram) and sound-based (phonogram) reading [501]) ● Phonological paraphasias [337] Excision / ischemia: ● **Naming** errors [341] (see [502]) ● (long-term) **alexia** (if white matter involved [440,450,503])
	Temporal pole (ant. STG + MTG, excluding anterior ITG & basal entorhinal / fusiform cortex)	**Semantic combinations / judgements** (see [504,505]), needed to process complex sentence structure [506], including (or perhaps especially [505]) integration of emotional valence and social concepts of words [95]	Intraoperative stimulation: ● **Naming** errors [225,325,326,330,507] ● Semantic errors [330] ● Sentence comprehension deficits [483,508] Excision / ischemia: ● Word finding difficulties [117,340,466], esp. visual **naming** [509,510] ● Word and sentence comprehension deficits [78,343,369,418,437] ● Language production deficits [437] ● Occasionally (less often after unilateral damage [511]) semantic errors or paraphasias [328,445]
Medial inferior temporal	Ventral occipitotemporal cortex (VOTC), (mid-post. fusiform)	**Orthographic processing** (reading and spelling [512,513] (see [514])**.** Subdivisions [515–517]: ● Post. (VOTC-1): visual categorization (e.g., of text) [518]/'VWFA' ● Ant. VOTC-2: orthographic-semantic interface [118,516,519]; integrating text with word meaning & sound [520]	Intraoperative stimulation: ● Reading errors /disruption [330,500,503,521–525], especially meaning-based (morphogram) reading [501] Excision / ischemia: ● Long-term naming deficits [341] (see [502]) ● Impaired reading [117,450,503,511,521,524,526] (see [517]), i.e., pure dys/alexia without hemianopsia and agraphia ● Spelling errors at VOTC-2 [118]
	Basal temporal language area (BTLA) (ant. fusiform, between pole and	Multimodal **semantic processing** [360,363,432] ● lexical orthographic processing ● interfacing orthographic and semantic information (see [118]	Intraoperative stimulation: various errors (e.g., [522] see [527,528]) ● Most consistently **naming** errors [498,499,529–534] ● Occasionally: Speech arrest / slowing during (meaning-based [501,522,532]) reading [529], auditory naming [529,530], comprehension, [522,527], semantic association errors [533]

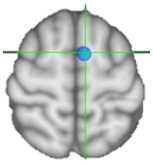
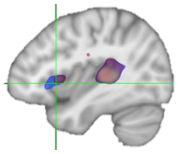
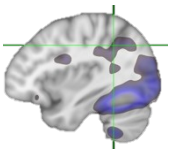
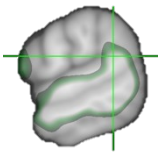
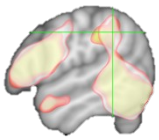
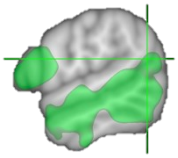
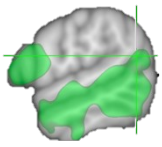
[redacted]	VOTC)		Excision / ischemia: ● Persistent naming decline [133,341,497,499,528,535] ● Alexia [497]

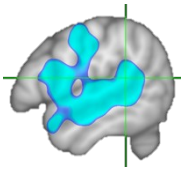
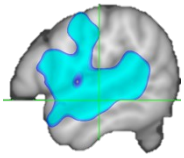
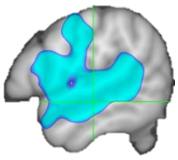
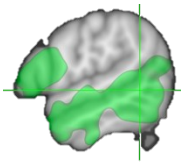
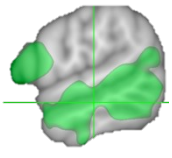
*Table 1 Legend. *Some regions perform multiple computations, either specific to language or not. This table is not intended to suggest one-to-one mapping of function, but rather reflects current evidence favoring relative graded representations of certain language-relevant computations across sub-regions of the brain. † The localization of stimulation-induced language errors, especially in the region of the posterior middle frontal gyrus, are difficult to ascribe to specific anatomical sub-regions due to the ambiguity of these functional boundaries and deformation of the surface anatomy by tumors. ‡Similarly, infarcts generally cannot be precisely ascribed to specific cortical subregions due to their involvement of vascular territories that usually span more than one functional subregion. NOS = not otherwise specified. **The true incidence / likelihood of subsequent deficits is very variable due to obvious variations in the nature of injuries/resections and individual patient factors; so 'predicted consequences' are intended to give an idea of the range of deficits that may arise, in the context of the language-related process most commonly associated with dissociable brain regions.*

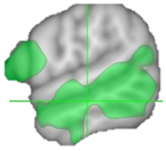
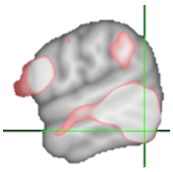
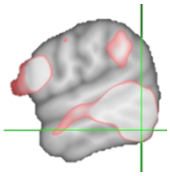
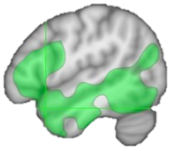
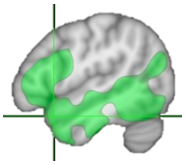
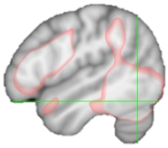
Table 2. A selection language localization tasks, according to surgical target.

target AREA	active condition	main targetted process(es)	control condition	Pros / Cons	adaptations	Neuroquery (with cross-hairs at target area)
Pars Opercularis (Pop)	(Covert) verb generation*	Phonological processing	Silent rehearsal (e.g., numbers) or repetition	Easier than fluency tasks but no language response measure	● Auditory / visual ● Rate of stimulus delivery (e.g. from 3s (hard) to 6s (easy)	"Phonemic" 
	Sentence completion	All, including syntax	False fonts (rearranged letter parts)	Relatively easy, more lateralizing than word tasks [67], but no language response measure	● Auditory / visual ● Rate & duration of sentence presentation	"Phonemic" & "naming" & "comprehension" 
	(Covert) naming, e.g., object-sentence generation [77]	Phonological & semantic processing	Silent counting (to control for motor planning) or rest	Simple, similar to intra-operative tests, but difficult for patients with anomia	● Remove items not consistently named ● Substitute rhyming or repetition task if severe anomia	"naming" 
	(Covert) phonemic / semantic fluency	Phonological, Executive control	Silent rehearsal or repetition	Cognitively demanding, no language response measure	● Auditory / visual ● Frequency of letter targets (e.g. 1 (hard) to 5-6 (easy) targets per block)	"phonemic" 
Pars triangularis (Ptr)	Sentence completion	Sentence level semantic processing, phonological processing	Visual scrambled letters / auditory reversed speech	No response measure, difficult for alexic patients or needs aural delivery system	● Auditory / visual ● Rate of stimulus presentation (fewer sentences per block)	"comprehension" & "naming" 
Pars	Semantic	Complex	Visual line	Button box	● Auditory /	

orbitalis (Por)	description decision task	associative semantic processing	drawings / auditory reversed speech	response, but default version requires auditory system	visual ● Adapt to performance level (e.g., children)	"semantic" 
ventral premotor cortex	Overt articulation or covert word generation/naming	Articulation and phonemic	Rest	Direct comparator to intraoperative speech but increases head motion confounds	● Auditory / visual	"speaking" & "phonemic" & "articulation" 
Mid-to-posterior Middle Frontal Gyrus	Sentence completion	Semantic and phonological processing	Visual scrambled letters / auditory reversed speech	No language response measure, difficult for alexic patients / need auditory system	● Auditory / visual ● Rate of stimulus presentation (fewer sentences per block)	"semantic" ** 
	If bilingual: language switching (naming) task [210]	Cognitive control	Rest	No language response measure	● Remove items not consistently named ● Rate of picture presentation (fewer/slower)	N/A
Area 55b	Articulation/ fluency	Phonemic/ articulation	Rest	No performance measure	● Choose task according to abilites	N/A
Exner's area	Writing dictated words [380]	Graphemic	Insufficient data, possibly drawing non-word shapes	Technically challenging to implement in MRI	● None identified	N/A
Pre-Supplementary Motor Aarea‡	Verb generation	Phonological processing	Resting fixation	No behavioral measure	● As for Pop	"phonemic" ** 

Insula	Fluency	Phonological, Executive control	Resting fixation	No behavioral measure	● As for Pop	"phonemic" ** 
	Two object naming/ sentence generation (e.g., [77])	Articulatory control	Rest	No behavioral measure (or needs a noise-cancelling microphone)	● Visual / auditory ● Adjust pace or rate of stimulus presentation	"articulation" & "phonemic" ** 
Intra Parietal Sulcus	Writing (or a motor planning task might suffice)	Graphemic/ goal-direction action planning	As for Exner's	As for Exner's	● As for Exner's	"graphemic" 
Supramarginal Gyrus	Auditory or Written sentence comprehension	Phonological processing and semantic integration	Perceptual control, e.g., reversed speech/ false font	No behavioral measure	● Speed, complexity or length of written word presentation	"comprehension" 
	Auditory or Written sentence completion	Sentence level semantic processing and phonological integration	False font / reversed speech	Relatively easy but no language response measure	● Auditory / visual ● Rate & duration of sentence presentation	"naming" &  "reading"
Angular gyrus	Visual semantic relatedness decision (e.g., [536])	Semantic processing	Decision on non-linguistic items (e.g., tones/lines)	Behavioral response; stronger semantic response for visual than aural stimuli[499]	● Present aurally for deaf patients	 "semantic"
	Written sentence completion	Semantic processing (integrating written words into sentence level meaning)	False font sentences	Relatively easy but no language response measure	● Auditory / visual ● Rate & duration of sentence presentation	"semantic" 

posterior Superior Temporal Gyrus	(auditory) Sentence completion	Phonological processing and integration (phonological-semantic-syntactic processing),	Perceptual control (e.g., reversed speech / false font/symbol strings) AND resting blocks†.	Generally easy, but no response measure	● Visual / auditory ● Adapt stimulus rate / length to (e.g., reading) abilities ● For severe comprehension difficulties, use object sentence generation or story listening (vs reversed speech)	"Phonemic" 
Middle Superior Temporal Gyrus	Auditory naming	Semantic and Phonemic / phonological	Acoustic control (e.g., tones or reversed speech)	Requires dedicated hardware, no response measure	● Adjust stimulus rate and pace ● Use object or written naming in deaf patients	"phonemic" 
	Auditory sentence completion	Semantic and Phonemic / phonological	Acoustic control (e.g., reversed speech)	Relatively easy but no language response measure	● Use visual stimuli for deaf patients ● Vary rate & duration of sentence presentation	"phonemic" 
posterior Middle Temporal Gyrus	(written) Sentence completion	Semantic integration (sentence level)	Visual control (e.g., scrambled letters)	No language response measure and difficult for patients with alexia/dyslexia	● Auditory version for patients with reading difficulties ● Adjust to ability (reading level / pace)	"semantic" 
Middle Middle Temporal Gyrus	Sentence completion	Semantic processing (word level)	Perceptual control (auditory/visual)	No language response measure	● Auditory/visual ● Adjust to ability (reading level / pace)	"Semantic" 

	Semantic relatedness decision [536,537] or categorization [153]	Semantic processing (word level)	Auditory: tone decision; visual: non-semantic decision (e.g., font)	Semantic categorization is most validated, but relatedness tasks more lateralizing [57]	● Auditory or visual	"Semantic" 
posterior Inferior Temporal Gyrus	Written sentence completion	Lexical-semantic/ orthographic processing	Visual control (e.g., scrambled sentences)	No language response measure and difficult for patients with alexia/dyslexia	● Adjust to ability (reading level / pace) ● Use single words instead of sentences	"Reading" 
	Word or sentence reading	Lexical-semantic/ orthographic	Visual control (e.g., scrambled letters)	No language response measure and difficult for patients with alexia/dyslexia	● Adjust to ability (reading level / pace) ● Use single words instead of sentences	"Reading" 
Temporal Pole	Semantic relatedness decision [536,537] or categorization task [153]	Semantic	Perceptual control: tone (auditory) or non-linguistic visual decision (e.g. upper case/lower case font) [153]	Auditory version needs hardware; sentence decisions more lateralizing than word decisions	● Visual / auditory ● Adjust stimulus presentation (slower or fewer per block)	"Semantic" 
	Auditory description decision task	Semantic	Tone decision	Button box response, but default version requires auditory system	● Auditory / visual ● Adapt content to comprehension level ● Slow down stimulus pace	"Semantic" 
Ventral Occipito Temporal Cortex	Sentence reading	Reading	Rest and consonant, false font or symbol strings	Difficult for patients with visual field deficits/dyslexia/alexia; pseudo- or non-word letter strings can reduce	● Vary complexity (short or longer) sentences, adjust to reading age / performance level	"Reading" 

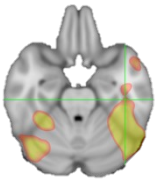
				relevant activation		
Basal Temporal Language Area	Naming (read-response and / or picture)	Semantic processing and phonological retrieval	Scrambled images or false-font text scanning	Generally easy but challenging for patients with visual impairments	● Adjust stimuli to performance (slower rate or easier stimuli) ● remove pictures not reliably named	"Naming" 

*Table 2 Legend. A range of published fMRI language tasks that probe the predominant function(s) of anatomical regions of surgical concern (Table 1). Tasks are not intended to be prescriptive, but offer examples of tasks that engage specific language processes most commonly affected by damage or disruption to a given brain region. Similarly, other control conditions are, of course, possible, depending on the clinical question. Sagittal views depict the left hemisphere. Axial slices are presented in radiological convention (left hemisphere on the right side of the image). Cross-hairs indicate the "target" brain region. *Covert generation is recommended to minimize head motion; if mapping of articulation-related sites (e.g., ventral premotor cortex) is clinically relevant, overt speech compared to resting fixation should be considered. † Some control conditions engage or de-activate language-related areas, or heighten attention/cognitive processing; inclusion of additional rest blocks to examine the (de)activations of both active and control conditions is strongly encouraged. ‡ Inclusion of a motor task may be helpful to define boundaries between the supplementary motor area (SMA) and pre-SMA. **Statistical maps derived from multivariate prediction-based meta-analysis of approximately 13,500 neuroimaging studies in Neuroquery (https://neuroquery.org/)[84] using the search terms listed in quote marks are shown thresholded at z=3.1 (corresponding to p<0.01) for all target regions (indicated by cross-hairs) except the Insula, preSMA and MFG (shown at a z-score threshold of 2.3 (corresponding to p < 0.05) for enhanced visibility).*

Table 3. Variations and steps in fMRI analysis: Pros and Cons

	Pros	**Cons**	**Recommendation**
Commercially automated fMRI analysis	Easy to use; work with neuronavigation suites; certified for clinical use (FDA/CE-approved)	Limited support for complex designs and advanced processing options; blind to errors such as excessive head motion or low magnitude responses; do not (easily) accommodate non-standard task designs; can impose constraints that block durations must correspond to fMRI volumes (multiple of the scanning TR); do not allow for variations in patient HRF/time-course	**fMRI results should be evaluated by a trained and experienced user** to confirm accuracy of results and identify potential errors introduced by inflexible automated pipelines
Distortions and distortion correction	Shifts signal back into the correct anatomical position	Cannot recover signal loss; requires (short) extra scans	**Should always be corrected** (using a fieldmap / comparable technique)
Head motion correction	Improves sensitivity (to detect true activations) and specificity (can reduce false positives)	Cannot correct for all motion effects; may give false impression of data quality	The best approach is prevention. Motion correction can fail, therefore **head motion should be assessed for each fMRI task**, including inspection of the raw images and performing multiple analyses to compare the effects of different motion correction choices on the resulting activation map(s).
Physiological noise correction	Can improve sensitivity and specificity, especially in inferior regions	Requires extra scans or extra processing	**Not generally used** for clinical fMRI by this working group
Brain extraction and mask generation	Removes non-brain tissues from the analysis, which can improve the statistics (irrelevant concern if individualized thresholds are used)	Can remove brain tissue if there are errors (e.g., due to bleeds/pathological signal change)	If automatically performed during pre-processing, the brain mask **needs careful inspection and, if necessary, correction** to ensure the surgical lesion is covered by the brain mask used for statistical analysis
Spatial smoothing	Improves sensitivity in general; especially for surface-based smoothing	Can reduce sensitivity, blur small areas, bias location of activation, and make activation localization less specific, particularly with larger smoothing; surface smoothing requires a lot of extra processing, and accurate surface reconstructions are intrinsically difficult in the	Perform no or minimal smoothing only; 1-2 x the acquired voxel size, up to a maximum of 5mm FWHM (full width at half maximum) but ideally 4mm.

	Pros	Cons	Recommendation
		presence of brain lesions, tumors in particular, and perifocal edema	
Data Analysis: ICA and GLM	ICA does not make assumptions about the time-series (like the GLM does), and instead provides maps with associated time series that can be post-hoc correlated with the time-series of the fMRI task model.	The output of ICA can produce several maps that are each correlated with the fMRI task time-course of interest, from which either the most appropriate map needs to be selected, or the maps need to be merged into one for surgical guidance. Additionally, different task conditions cannot be directly contrasted against each other.	Consider supplementing the standard GLM approach with ICA, especially when delays in task performance or hemodynamic responses are suspected based on patient performance / feedback or the nature of the surgical pathology. Further research exploring the utility of currently research-based data driven analysis techniques is needed.
Statistical inference	Provides an objective assessment of trustworthiness and a way to differentiate noise from true activations	Choosing a threshold is tricky as there is significant variability in activation magnitude across subjects; need to control false negatives rather than standard methods that control false positives; institutional approaches usually used	Consider the potential for delayed or even inverted hemodynamic response; complementary exploration of both general linear model and independent component analyses is encouraged. Thresholding of statistical maps should be tailored to each individual. Inspecting - and possibly presenting - unthresholded effect size maps can also be important on occasion.
Registration	Important for accurate visualization and localization	Simple methods can be inaccurate; can be difficult to transfer to neuronavigation system	Not essential (nor desirable) for laterality calculations, but crucial for accurate localization of fMRI activations. Inspect registration accuracy carefully, with particular attention to areas where signal loss or distortion in EPI may affect the interpretation of fMRI results displayed onto the high resolution anatomical.
Cerebrovascular reactivity mapping	Promises to identify regions with reduced BOLD responsiveness (“neurovascular uncoupling” = NVU); complementary to task fMRI	Requires induction of hypercapnia (e.g., by breath-holds or similar) but temporal shifts between the assumed and real response may lead to false-negative detections and mimic NVU.	Not specifically recommended but can be considered to interrogate the sensitivity of being able to evoke and detect BOLD fMRI signal changes

Appendix A: Resting state fMRI for language mapping

Resting-state fMRI (rsFMRI) has been hailed as a promising technique in language mapping, particularly since fluctuating task performance, patient motivation, and cognitive and/or neurological deficits do not impact the resting-state to the same extent as they do task fMRI. Moreover, current acquisition of resting-state fMRI is generally proposed to be faster than task-based fMRI, by virtue of collecting one single resting scan in place of several different tasks. However, the 'optimal' scanning length for resting-state data has not been determined (Hacker et al., 2013), and may well be longer than for task fMRI due to the fluctuating nature of resting cognitive states as well the analysis framework (Termenon et al., 2016).

The localizational power of correlations between BOLD signals in a no-task, 'resting' condition was first described by Bharat Biswal and colleagues in 1995 (Biswal et al., 1995). Functional connectivity as a statistical measure representing communication or functional connection between brain areas was introduced in 1989 (Aertsen et al., 1989), and was then applied to PET data by Karl Friston (Friston et al., 1993). Subsequently, Biswal and colleagues showed that rsfMRI time series emanating from the left and right motor cortex are always highly correlated, even when not performing any movement or thinking about doing so. Thresholding the correlations of either motor cortex with every other voxel in the brain therefore leads to a map of the resting state motor network.

Since rsfMRI became an often-used tool to investigate brain function in the mid 00's (Fox et al., 2005; Matthews et al., 2006; Shimony et al., 2009), a growing body of literature has investigated the relationship between the resting-state and task conditions (Cole et al., 2014, 2016; Gratton et al., 2018). Taken together, these studies show that task activations generally take place on top of a resting-state backbone of communication, while individual fingerprints of activity and connectivity are maintained throughout both conditions. These findings support the possible use of rsfMRI in localizing cognitive functions.

There are roughly two analysis strategies when using rsfMRI for language mapping. The first approach involves using prior information, e.g., setting a seed region for the

analysis, such as Broca's area, from which highly correlated regions are mapped, similar to the original work by Biswal and colleagues. More recent versions of this approach can be described as supervised classification techniques using machine learning methods (Hacker et al., 2013; Luckett et al., 2020). The second, data-driven strategy involves extracting resting-state networks or independent components from the entire scan without *a priori* seeding in particular regions (Kiviniemi et al., 2004; Van De Ven et al., 2004), which was also first used for motor mapping (Kokkonen et al., 2009).

Initial studies using both methods are promising: in healthy and smaller patient cohorts, seed-based rsfMRI seems able to extract language areas (Cordes et al., 2000; Hampson et al., 2002; Tomasi & Volkow, 2012; Zhu et al., 2014). However, it is difficult to accurately place a seed region in a 'language' area in patients, for several reasons. Firstly, a mass lesion (such as high grade brain tumor or metastasis) and surrounding edema can profoundly distort, compress, displace and obscure the underlying anatomy. Secondly, considerable plasticity and reorganization may have taken place already (Sanai et al., 2008) even in patients with no visible lesions (such as 'MRI-negative' temporal lobe epilepsy). Consequently, the need to anatomically define a seed region for analysis creates a circular problem, since 'language' regions generally do not correspond well to anatomical boundaries. More modern supervised classification approaches partially overcome these difficulties with the benefit of prior information on the expected connectivity of different networks across the brain (Dierker et al., 2017). Indeed, data-driven approaches have shown potential in language lateralization and localization within patient populations (Branco et al., 2016; DeSalvo et al., 2016; Mitchell et al., 2013; Parker Jones et al., 2017; Tie et al., 2014), although greater interindividual variation in resting-state mapping as compared to task fMRI has been reported (Sair et al., 2016).

Crucially, uncertainty in what (language or other) functional process is performed by a given resting network have hampered enthusiasm for the use of rsfMRI in the clinical setting. Many centers currently do not use the technique – outside of research – despite its initial promise (Sair et al., 2017; Tanaka & Stufflebeam, 2016). However, there is at least one on-going clinical trial aiming to evaluate resting fMRI (and CVRM) for preoperative language mapping in adult glioma patients

(https://classic.clinicaltrials.gov/ct2/show/NCT03964909), currently due to complete recruitment at the end of 2023.

Specific methodological issues should be kept in mind when interpreting currently still ambiguous resting fMRI-derived results. Firstly, many studies use the task-based fMRI as a gold standard, which potentially obscures the independent benefits of rsfMRI in mapping critical cortex. Further complicating this issue is that there are expected differences between rsfMRI and task fMRI (such as activated task-general areas not seen in rsfMRI, and right hemisphere language areas not seen in task fMRI) (see (Luckett et al., 2020) for more detail on this issue). Secondly, both approaches towards analysis of rsfMRI demand specific expertise: placing a seed region in surgical patients necessitates extensive knowledge of anatomy and functional reorganization, possibly combined with task fMRI to establish the source of activity. In the data-driven methodologies, spatial language components need to be compared to known maps and classified as such. Furthermore, both approaches may depend also on advanced computational expertise that is not always available in clinical teams. Thirdly, similar to task fMRI, it is generally necessary to threshold connectivity values in order to obtain a dichotomous map to be of use in the surgical theatre, for which an optimum has not been defined (either for task or for rsfMRI). Finally, there are currently no studies investigating the (added) value of rsfMRI in limiting postoperative language deficits, which renders the clinical relevance of the method unknown. Instead, the few studies that explore correspondence between resting fMRI-derived ‘language’ networks report high sensitivity and specificity (both 96%) when compared to Wada (DeSalvo et al., 2016), but low sensitivity against ESM when ‘validated’ templates (Tie et al., 2014) were selected in 31 patients undergoing awake surgery for low grade glioma (Cochereau et al., 2016).

Despite this lack of general consensus on the specificity and sensitivity of rsfMRI in language mapping, at some of the Working Group institutions, resting-state (language) mapping has been investigated throughout the years (Lee et al., 2016; Roland et al., 2019; Shimony et al., 2009) and is currently part of the presurgical work-up in all patients (Leuthardt et al., 2018). An emerging intermediate approach is movie watching during fMRI (Tie et al., 2015; Yao et al., 2022). Relative to rsfMRI, active movie watching may – pending further evaluation – enhance both patient compliance

(i.e., remain still without falling asleep) and the engagement of language networks in patients with reduced abilities to perform task batteries is fMRI.

Appendix B: A compendium of fMRI confounds and artifacts relevant to language mapping

Structural imaging evidence of pathological confounds – especially ischemia, necrosis, gliosis, cyst formation, hemorrhage, calcifications or flow void / shunting – poses unique challenges for the interpretation of the fMRI blood oxygen level dependent (BOLD) signal in their proximity.

Pathologies prone to present with such confounds are brain tumors, vascular malformations both with (especially arterio-venous malformations = AVMs, pial arteriovenous fistulas = pAVFs) and without shunting (cavernomas / cerebral cavernous malformations = CCMs or slow flow venous malformations according to the International Society for the Study of vascular anomalies, ISSVA, developmental venous anomalies = DVAs, capillary teleangiectasias), and strokes. Tumors can present with hemorrhages, both space-occupying intracerebral hemorrhages as well as intra- and peritumoral microbleeds, but also intratumoral ischemias, calcifications and vessel flow-voids feeding (arterial) or draining (venous) the tumor. Flow void / shunting is the hallmark of AVMs which, however, may also present with hemorrhage, ischemic strokes from stealing, gliosis and calcifications. CCMs present with MRI signal alterations depending on their Zabramski type, with blooming signal loss of variable degrees on T2*-weighted BOLD-sensitive GE-EPI being paradigmatic, but can also exhibit calcifications. Most commonly, CCMs are encapsulated and hardly contain neural tissue. Consequently, an uncomplicated atraumatic surgical removal of a CCM would not be expected to cause functional deficits, because no significant neuronal tissue is resected. However, when associated with DVAs, resection is dangerous due to pending venous congestive hemorrhage. DVAs present venous flow voids which can be quite large and typically form a caput medusae. Capillary teleangiectasias, on the other hand, present with T2* signal loss and some blooming, similar to – but usually somewhat less than – in CCMs.

Previous surgeries (and embolizations) can also result in imaging artefacts that may impact on the reliability of fMRI results (Bartsch et al., 2006; Peck et al., 2009), while prior treatment may also reduce detectability of fMRI signal due to lesser understood effects of radiotherapy or edema (e.g., (M. M. Połczyńska et al., 2021)).

The impact of some of these confounds are discussed and illustrated below. Other factors that are not unique to clinical populations – such as head motion, image distortions and physiological noise – are discussed separately in Appendix E.

a) Ischemia: Cytotoxic and vasogenic edema significantly disrupt T2* of affected tissue (Grüne et al., 1999), which can result in abnormal, absent, protracted or shifted BOLD responses (e.g., (Amemiya et al., 2012; Murata et al., 2006)). The interpretation of fMRI on its own may therefore pose challenges in neurosurgical patients with a previous infarct because these regions will likely not elicit much, or any, measurable fMRI activation. This concern arises even in the context of a small subcortical infarct remote from the current (or re-current) surgical focus (Pineiro et al., 2002), such as resulting from injury to perforating arteries during neurosurgery. It is therefore good practice to obtain at least a fast standard stroke diffusion scan in or around the time of the fMRI session. When diffusion tractography data are recorded, as most centers do along with language fMRI, ischemic lesions can be detected in these.

b) Necrosis: Strictly speaking, necrotic tissue is dead and dead neuronal tissue should not evoke a BOLD response. Tissue necrosis can be accompanied by, result from and essentially exert a similar confounding impact as intra-/perilesional hemorrhages or infarctions. Islands of necrotic and viable tissue can be intermingled with each other in, for example, high grade primary brain tumors. Again, the degree to which such viable tissue remains functional and coupled to detectable BOLD responses is not known but it has been proposed to account for “function within tumor” (J. G. Ojemann et al., 1996; Skirboll et al., 1996).

c) Gliosis: Gliosis is scar formation by glial cells within the CNS. Gliotic tissue usually presents with prolongations of their T2-relaxation time. The extent to which gliosis affects the BOLD signal and neurovascular coupling is not easily established but the BOLD response can certainly be attenuated by it.

d) Hemorrhage: High-grade gliomas and metastases, for example, are not uncommonly accompanied by intra- and peritumoral microbleeds. Accumulation of paramagnetic iron-containing substances in blood deposits after a bleed profoundly affects magnetic susceptibility and can dramatically reduce the T2*-

weighted BOLD signal not only in but also around the bleeding (“blooming effect”). The consequence can be a more or less dramatic signal void especially in and around areas of hemosiderin staining and associated false negative fMRI findings (i.e., missing ‘real’ activation) (Thickbroom et al., 2004). This is illustrated in Fig S1a. Similar to flow void / shunting (see below) signal loss / blooming from hemorrhages requires careful checking that the lesion of interest is included in the analysis mask (see main document, section 3.3.3). Notably, early subacute hemorrhages with intracellular methemoglobin will present T1-hyper- but T2*-hypointense, i.e., the MR signal will dissociate into different directions on the fMRI (signal loss) and commonly used anatomical T1-weighted images (signal increase). Such imaging characteristics should be appreciated and properly accounted for as potential fMRI confounds. This is primarily a duty of the neuroradiologist involved in the multidisciplinary team involved with clinical language fMRI.

e) Calcifications: Calcifications in brain tumors such as oligodendrogliomas or other intra-axial lesions such as CCMs inevitably result in flat BOLD signals, which are per se true negative findings. However, when partial volume effects and blooming sets in, these can also lead to false negative detections.

f) Flow voids/shunting and stealing: Flow void and shunting are the hallmark of AVMs and pAVFs. Both pose special challenges in terms of fMRI considerations. AVMs and pAVFs are vascular malformations with shunting. Because within an AVM (and pAVF) blood shunts from the arterial to the venous side of the vascular bed, it has less time to deliver oxygen to surrounding neurons. Thus, shortening of the circulation time within, at least, high-flow AVMs and pAVFs substantially affects the effectiveness of oxygen supply of the neighboring cortex by pathological AVM /pAVF vessels. This corresponds to the clinical experience that, usually, no deficit will ensure if a compact AVM / pAVF can be interventionally treated by embolization or surgically removed without damage to the surrounding brain parenchyma. En-passant AVM feeders of arteries supplying distal language-critical areas and a diffuse / infiltrative nidus can make this difficult (because the main artery is at risk to embolization with subsequent infarction, and en-passant feeders require meticulous surgical skeletonization) and can lead to partial embolization / removal with high risk for regrowth of the AVM. Another phenomenon is that the high flow of shunting with an AVM (or pAVF) “steals” blood from other vessels supplying close-by or even distal brain

regions. Clinical language fMRI primarly aims to clarify the “eloquence score” of AVMs according to the Spetzler-Martin or further supplementary gradings prior to interventional and / or surgical treatment attempts. However, flow voids and shunting can obviously cause false negative fMRI results while ‘stealing’ may, at least in theory, also lead to false positives. Thus, both activations and a lack of activations near or within an AVM are extremely difficult to interpret. Importantly, the signal loss of flow voids, shunting and stealing on T2*-weighted BOLD-EPI requires careful checking of the analysis inclusion masks as with signal loss due to blooming from hemorrhages such as in CCMs (see above and main manuscript section 3.3.3).

g) Metal implants / craniotomy defects: in addition to normal air-tissue boundaries, surgically created pockets of air, metalwork (e.g., drilling abrasions) and other implanted hardware (trepanation fixes, inserted catheters or ventricular drainages) are all potentially important sources of EPI distortions and signal loss. In some cases, signal loss and distortions caused by implants, craniofix (or even dental work such as fixed metal braces or retainers in the maxillary teeth) are so large as to prevent sensible fMRI mapping near to these areas (Fig S1b). A related consideration arises when a biopsy has been performed prior to fMRI referral, even when, for example, a borehole craniotomy is not covered osteoplastically and no metal screws or plates were inserted. This is because EPI distortions and signal loss arise at air-tissue boundaries (including drilling abrasions). Therefore, a biopsy of a tumor reaching the surface of the brain can result in substantial EPI distortions and signal black-out immediately next to the biopsy site. An example illustration is presented in Fig S1c. Prior radiological imaging will help to decide whether such implants or other artefacts would render attempts at fMRI futile.

Signal confounds, if known in advance, can sometimes be partially mitigated by adapting scanning sequences, for example by adjusting the phase-encoding direction and possibly slice acquisition orientation, and by spin-echo (SE) BOLD and/or perfusion-based sequences. SE-EPI also suffers from distortions, which can be corrected, but not from T2*-related signal loss/ blooming which cannot be recovered. Similarly, careful timing of fMRI, primarily prior to embolization procedures for AVMs, may be of advantage to avoid related artifacts.

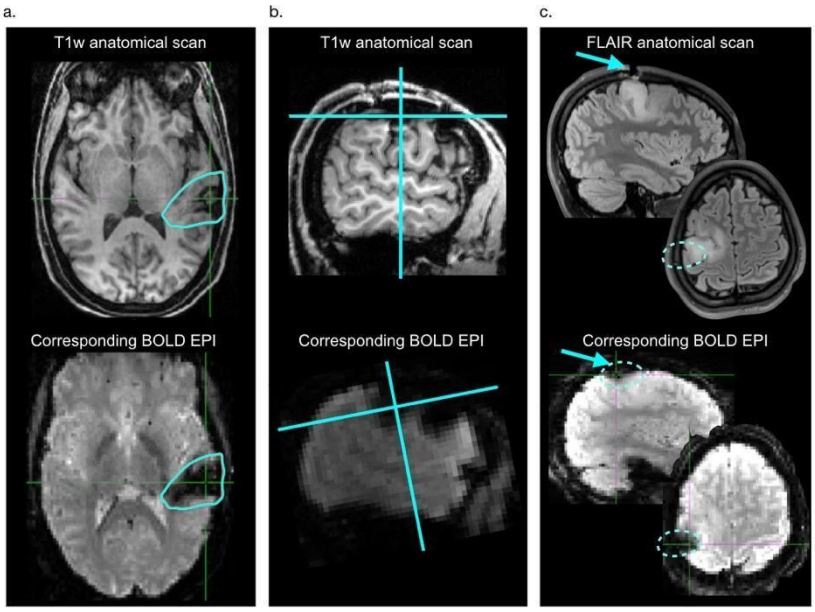


Fig S1. A compendium of artifacts in clinical fMRI. *Legend. Illustration of the effects of paramagnetic metal artifacts on BOLD-sensitive GE-EPI most commonly used for fMRI. A. hemorrhage residuals - in this case from a cavernoma (CCM) but also relevant in tumor- or AVM-related hemorrhage - can induce an extensive ("blooming") signal loss (outlined in cyan) disproportionate to the size of the CCM (identified by the crosshairs on the T1-weighted scan). Examining a diagnostic T2*- or susceptibility-weighted scan shortly prior to planning fMRI is highly valuable to decide if signal voids in the context of hemorrhage / hemosiderin staining may hinder language fMRI. (In other words, when there is virtually no signal, it will be difficult to detect a language-related signal change). B. Miniplates and other metal implants from prior surgery, even when non-ferrous and MRI-compatible and 'safe', generally also produce substantial signal drop-out and image distortions in EPI. C. Burr holes can result in similar signal loss and image distortions affecting the directly underlying tissue.*

Appendix C. fMRI technical and hardware considerations

In addition to task selection, the clinical utility of fMRI is strongly influenced by technical factors in the acquisition and processing of clinical fMRI data. These technical considerations

are independent of the specific pre-surgical functions (language, motor, vision, etc.) that are being mapped with fMRI. Protocols for high quality fMRI data collected as part of the Human Connectome Project (http://protocols.humanconnectome.org/HCP/3T/imaging-protocols.html) and UK Biobank [538] offer useful guides. Here we consider choices in acquisition parameters and factors that impact specifically on clinical fMRI mapping performed in the *individual patient*. The primary recommendations in relation to technical aspects of fMRI acquisition were indicated in the main manuscript. This Appendix aims to provide additional background to these recommendations.

MRI field strength

The vast majority of modern MRI scanners in clinical use in 'high income' countries today operate at a static magnetic field strength of either 1.5 or 3 Tesla (T). Aside from potential artefacts (considered in Appendix B), the crucial consideration for fMRI data quality is how well the pulse sequence, usually T2*-weighted gradient-echo echo-planar imaging (GE-EPI), detects small fluctuations of task-related BOLD signal changes (typically around 1%) on a given scanner at chosen parameter settings. Because BOLD contrast-to-noise ratio approximately doubles at 3T compared to 1.5T (Kruger et al., 2001), fMRI substantially benefits from higher field strength. 7T may in future offer an additional benefit, with FDA-approved scanners now in operation and emerging recommendations for applications in the clinical domain (Opheim et al., 2021). However, doubling the static field strength similarly increases artifacts, especially from magnetic susceptibility-related signal loss, and also quadruples the specific absorption rate (SAR). Higher SAR may raise safety restrictions at ultrahigh field strength, while susceptibility-related signal drop-outs and gradient non-linearities substantially increase at 7T. **Therefore, 3T scanners remain the recommended choice for routine clinical fMRI.**

Temporal fMRI resolution

A key factor influencing the statistical detection of “activation” in task-fMRI is how well the task-induced hemodynamic response is captured. The latter is influenced by the temporal resolution, i.e., how quickly images are sampled (Kruger et al., 2001). In order to assess the lateralization of language functions, it is important to image the entire brain. With standard (i.e., unaccelerated) clinical fMRI sequences, it takes a minimum of 2 to 3 seconds (depending on other choices, such as spatial resolution) to acquire one complete volume of the brain, referred to as the EPI repetition time (TR). A rule of thumb in choosing the optimal TR is that, within clinically achievable parameters, *faster is better*. For this reason, faster temporal sampling often boosts statistical power and increases the chances of fully sampling the hemodynamic BOLD response (Constable & Spencer, 2001). **Long TRs (i.e., much above 3s** (Constable & Spencer, 2001)**)** are not advisable, because additional task blocks (beyond the standard 4.5 – 8 minute tasks) would be needed to compensate for the low statistical sensitivity that accompanies slow image sampling rates. At shorter TRs, it is important also to adjust the flip angle for GE-EPI below 90° to equal the Ernst angle, to maximize the signal-to-noise ratio (SNR). Emerging state of the art acquisition sequences enable rapid acquisition of multiple slices simultaneously, allowing sub-second TRs. These simultaneous multislice (also called Multiband) sequences boost statistical power in fMRI (Fig S2, Box S1) but currently still raise challenges with real-time data reconstruction, storage and processing. Therefore, when available (e.g., on recent generation scanners), limited amounts of acceleration (e.g., by a factor of 2 or 3) are currently used in clinical settings, usually to acquire the same images in shorter scan times. With progressive modernization of MRI and data management systems, fMRI acquisitions with a TR close to 1s are becoming more readily achievable.

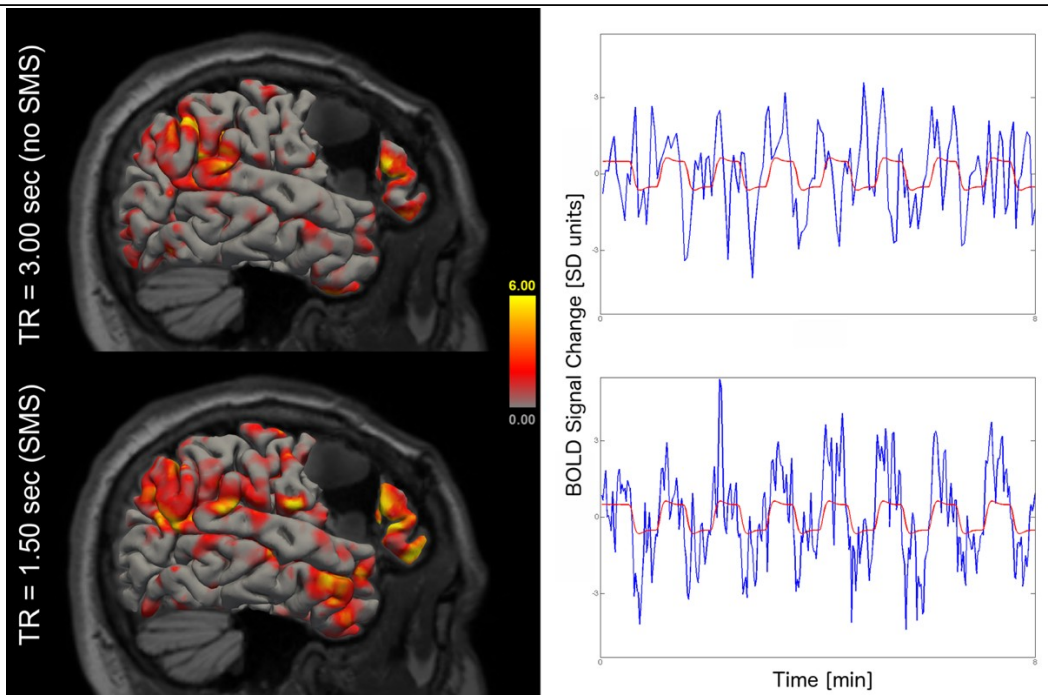


Fig S2. Statistical boost in fMRI from simultaneous multislice acceleration schemes. *Legend. Benefits of temporal acceleration for clinical fMRI to map speech and language functions in a left-hander with a recurrent, right frontal low-grade glioma prior to second surgery scanned twice with and without SMS. Doubling the temporal fMRI resolution, i.e., from 160 to 320 volumes in the recorded time series by using SMS increased the statistical confidence (red-to-yellow Z-statistics) of language activations and improved the temporal correlation (Pearson's r = 0.2 vs. 0.7) of the respective time-courses (blue) with the model (red) of the block design convolved with a hemodynamic response function (HRF). Dominant right-brain speech had already been confirmed by intra-operative electrical stimulation mapping (ESM) during the primary, partial tumor resection but ESM was, at the time, complicated by a series of intra-operative seizures. The current fMRI examination prior to secondary resection was performed to support a sufficient safety margin between the recurrent, low-grade glioma and the cortical fMRI activations.*

Achieving fast temporal resolution • BOX S1 •

Recent advances in accelerated imaging are based on simultaneous multi-slice (SMS), also called "multi-band" EPI acquisitions. SMS acquisitions collect multiple slices of the brain at once. In this way, these acquisitions speed up the temporal sampling of whole-brain volumes to half the time or less, compared to standard

EPI (Moeller et al., 2010; Smith et al., 2013) and boost statistical sensitivity to the BOLD signal [206]. Currently, however, some practical considerations limit the widespread adoption of high acceleration sequences.† The added specificity of accelerated sequences for pre-surgical mapping applications has not been widely evaluated against standard acquisitions. However, in our experience SMS can improve the correlation of time-courses extracted from the data with the model of the language paradigm and increase the sensitivity of clinical fMRI (Fig S2).

Of note, parallel acquisition techniques (PAT) have earlier been used to speed up EPI recordings in-plane (rather than through-plane, as done with SMS). In addition to allowing for more slices within a TR period (or to reduce the echo time, TE), in-plane PAT reduces geometric EPI distortions (by reducing the "effective" echo spacing), and can be used when an adequate signal-to-noise ratio exists as with 3T.

† Multiband / SMS capabilities depend on having a head coil with at least two rungs of coil elements (e.g. a 32-channel coil) and dedicated image reconstruction. Importantly for clinical applications, at high acceleration factors (6 and above), the acquisition rate of data can exceed the image reconstruction capabilities of many but the latest generation scanner computer interfaces. This risk can be mitigated on older systems by using a modest or low acceleration factor (such as 2 to 4). High SMS acceleration factors enabling low TRs (< 1s) also come at the expense of SNR and, particularly, CNR (contrast-to-noise ratio) between gray and white matter which is why then single-band reference scans without SMS are often recorded for registration purposes.

Spatial fMRI resolution

In order to image the whole brain in a short TR, clinical fMRI data are typically measured at a spatial resolution between 2x2x2 and 3x3x3mm. This lower resolution, compared to structural scans (usually 1mm or less), reflects the trade-off needed to maintain an adequate SNR (Bodurka et al., 2007; Triantafyllou et al., 2011). The importance of spatial resolution is in the spatial accuracy with which fMRI activations can be localized and would ideally be informed by the neuroanatomical structure of

interest being imaged (i.e., the cerebral cortex with mean cortical thickness of 2.6mm (Glasser et al., 2016)). Spatial resolution also directly informs statistical image processing choices including data smoothing (section 3.3.4). Ultra-high resolution fMRI (less than 1.5mm voxel edges) is more typically reserved for 'mapping' non-language related activity in very small structures such as the lateral geniculate nucleus, amygdala or sub-regions of the hippocampus. Importantly, higher spatial resolution decreases the signal measured for each voxel. **For language fMRI mapping, this Working Group advocates for an isotropic (or as close as possible) voxel size of 2.4mm^3** (and not larger than 3 mm^3 (Triantafyllou et al., 2011)).

Additional fMRI hardware requirements

Comprehensive fMRI language mapping requires additional hardware over and above diagnostic MRI. These are detailed in Box S2. In short, task-based language fMRI typically requires:

- A mode of delivering visual or auditory stimuli to the patient in the scanner,
- A system to record patient responses (such as a button box or microphone/verbal response capture system), depending on task requirements,
- Computer hardware with dedicated stimulus presentation software,
- A means to synchronize the functional paradigm with the scanning, commonly by the scanner triggering the stimulus presentation or, less commonly, vice versa (requiring the appropriate and usually additional cable linking the scanner to the stimulus delivery computer).

Real-time mapping of function

Some vendors offer real-time generation of functional activation maps. These are not recommended as the primary or sole method for analyzing data, but do offer the advantage of providing an early indication as to whether the expected activation pattern is appearing during the task (Chwang et al., 2017). If this pattern is not present, and there is additional cause to suspect the patient is not performing the task (e.g., no expected behavioral responses are being recorded) the task can be stopped,

and an assessment be made as to why and how to fix the problem, such as repeating a quick demonstration or adjusting stimulus presentation volume/rate.

Hardware considerations for fMRI • BOX S2 •

Dedicated computer to run fMRI tasks

- Delivery of task paradigms requires software installed and run on a permanently available computer, dedicated tablet/laptop (or on the console controlling the scanner). Numerous free as well as proprietary software packages for fMRI task coding and presentation are available. Software choices are usually based on site preference and budget considerations.

Dedicated in-scanner hardware for visual and / or auditory stimulus delivery

- **Visual presentation** of task instructions and stimuli requires MR-compatible goggles, a screen or projecting device to transmit the stimuli to the patient in the scanner. Visibility is facilitated usually through a mirror mounted to the head coil or using specialist goggles worn inside the coil. Dedicated MRI-compatible frames with a range of corrective lenses will allow most patients who normally wear glasses to adequately perform visually-presented fMRI tasks. Some goggles allow for direct correction of visual acuity. Clear perception of visual stimuli in the scanner should always be tested before acquiring fMRI data (both to ensure good task performance and to minimize head motion).
- **Auditory stimulus transmission** is needed for fMRI mapping of certain language functions including **auditory** comprehension (Saur et al., 2008) and processing auditory feedback during speech (Scott et al., 2000). However, auditory tasks are challenging due to high acoustic noise levels (easily approaching 100dB) of EPI conventionally used for BOLD-fMRI, requiring means to ensure task stimuli are perceived while the overall noise levels of the scan remain tolerable. Furthermore, the fundamental auditory frequency peak of GE-EPI (given by the inverse of twice the echo spacing of

the EPI read-out) and its harmonics overlap with the main frequency spectrum of human speech between 0.25 to 4 kHz. Modifications to enable auditory fMRI generally divide into altering fMRI sequence parameters or using specialist audio delivery equipment. Reflecting the complexity involved, **the use and mode of auditory stimulus presentation varied among this working group**. Approaches include, among others, active noise cancellation (ANC), and 'sparse sampling' approaches to acquire fMRI data. ANC headphones offer the advantage of ear-selective auditory stimulus presentation. Given that auditory fibers are thought to mainly (yet not completely) cross the midline from ear level to the primary auditory cortex, auditory stimuli presented to the ear ipsilateral of the hemisphere to be operated upon will be primarily processed at the lower, non-linguistic level by the contralateral hemisphere which can be of advantage for clinical language fMRI. Good practice, regardless of the specifics of auditory stimulus delivery, is to test the auditory paradigm while a test-scan is generating EPI noise as during the real fMRI task. The sound volume can then be adjusted to an appropriate level matching the patient's hearing abilities. Auditory delivery of task stimuli often requires custom sequence or implementations, at extra investment, explaining why visual presentation remains the dominant mode of language fMRI mapping.

Hardware to monitor and record patient task performance (optional but preferred)

- **Patient performance.** Tasks requiring a response are typically favored over covert word generation, so that compliance can be monitored, and behavior used to inform the interpretation of results. Response monitoring helps to ensure, for example, that the patient did not fall or asleep or become exhausted or less responsive during the task. Task responses are typically monitored using a **button box response** (e.g., selecting between task stimuli) to minimize head motion artifacts associated with overt speech. Spoken responses, however, may be important for certain tasks including sentence completion, and can be captured using a **dedicated microphone** within the MRI scanner bore, typically at non-negligible extra

cost.

Hardware to temporally synchronize the onset of tasks with the start of fMRI scans

- Stimulus presentation and fMRI scanning, i.e., the EPI time-series, must proceed in precise **temporal synchronization** – either by the scanner triggering the stimulus presentation (most common) or vice versa (less commonly used; e.g., by using an auditory click as an external trigger). For optimal task analysis, the onset of the fMRI paradigm should be time-locked to the fMRI volumes acquired after magnetization equilibrium has been reached. Some MRI systems automatically exclude ‘dummy scans’ at the start of each fMRI scan from the saved EPI time-series. On other scanners, it is up to the practitioner to start the fMRI task after an appropriate delay. The number of “dummy scans” to be discarded depends on the repetition time (TR) of the EPI sequences. Automating this process requires vendor-specific trigger mechanisms to read the scanner-emitted ‘trigger’ pulses into the computer delivering the task stimuli (when the scanner “drives” the paradigm) or external dummy triggers (when the paradigm is used to “drive” the scanner). Care has to be taken to ensure that any equipment transferred into the Faraday cage of the scanner room does not introduce any artefacts into the fMRI images.

Appendix D. Diffusion Tractography

Acquisition of complementary multidirectional diffusion data is highly recommended as part of every language fMRI scan prior to resective lesion surgery. Even if cortical language structures are not placed at risk by a given surgery, knowledge of the location of subcortical axonal fiber bundles that connect distributed language systems is crucial to avoid disconnection aphasia / alexia. These fibers may be displaced by pathology in unpredictable ways and are as important (if not potentially more) to preserve as cortical language sites (Herbet et al., 2016).

There many ways to collect and process diffusion MRI data for DTI and tractography (D. K. Jones & Cercignani, 2010). However, the importance of integrating subcortical connectivity in attempts to 'map' distributed language networks has been repeatedly highlighted (Tremblay & Dick, 2016). A key decision is whether to analyze DTI and tractography independently using standardized techniques (i.e., seeding/targeting ROIs in standard locations, see (Wakana et al., 2007), or to use ROIs from the fMRI as seeds and targets. The former has the advantage of relative standardization of methodology in a way that accommodates structural, but not functional, variance between individuals. The latter may better accommodate functional variance, but is somewhat circular and may be subject to propagation and magnification of errors (i.e., a false negative or false positive in fMRI may cause a magnified error / bias in diffusion tractography). Therefore, **this Working Group recommends separate, independent analysis of diffusion tractography and fMRI data.**

In the clinical setting, diffusion MRI for DTI and tractography can be performed at approximately 2mm isotropic resolution and 30-60 or more unique diffusion encoding directions within acceptable imaging times. At that spatial and directional resolution, many large white matter pathways can be quite reliably identified and reconstructed. Accuracy of fiber tract reconstruction benefits from using algorithms that model at least two fibers per voxel and, as with fMRI, diffusion tractography results may be unreliable in the setting of head motion. Language pathways of surgical relevance most agreed upon include the: (1) arcuate fasciculus (AF), (2) superior longitudinal fasciculus (especially its 'horizontal' segment known as SLF-III and its 'posterior' segment, also known as the ascending or temporoparietal segment (SLF-tp) and sometimes referred to as SLF-V), (3) inferior occipitofrontal fasciculus (IOFF, also

known as IFOF), (4) (posterior) inferior longitudinal fasciculus (ILF) and possibly (5) the uncinate fasciculus (UF) and (6) frontal aslant tract (FAT) (for reviews, see (Abhinav et al., 2015; Chang et al., 2015; Martino et al., 2011; Voets et al., 2017; Yagmurlu et al., 2016)).

The extensive work in this field cannot be adequately reviewed in this Appendix. However, a summary overview of some key studies / findings are listed below. The AF consists of a long, deep segment, which runs parallel and medial to the anterior (aka horizontal) segment (SLF-III) and the posterior segment. The latter is also known as the ascending or vertical branch of the AF, corresponds to the temporoparietal part of the superior longitudinal fasciculus (SLF-tp or SLF-V) and overlaps with the anterior vertical occipital fasciculus (VOF). Within the dual-stream theoretical model of language (Fig 3a, main manuscript), the anterior segment of the AF and adjacent SLF-III can be conceptualized as the primary structural connection of the strongly lateralized, dorsal phonological (arcuate) and articulatory (SLF-III) stream while the ILF, IOFF, and UF would represent structural connections of less interhemispheric asymmetry within a ventral lexical-semantic stream assumed to integrate cross-modal content and conceptual meaning (see (Duffau, 2015) for review). According to the classical Wernicke-Lichtheim-Geschwind model, disruption of the dominant AF leads to conduction (i.e., repetition) aphasia, while its electrical stimulation induces phonological and repetition errors (Duffau, 2015; Sierpowska et al., 2017). Specific effects of damage to the ventral stream tracts remain ambiguous, however, electrical stimulation applied to the IOFF/IFOF reliably induces semantic paraphasias during simple picture naming, while more cognitively demanding naming processes may be disrupted by stimulation of the UF (Duffau, 2015; Papagno et al., 2011). Disconnection of the (posterior) ILF and anterior VOF may be associated with pure “preangular” alexia without agraphia and hemianopsia (Bartsch et al., 2013; Epelbaum et al., 2008; Greenblatt, 1990). The FAT is a structural connection between SMA/pre-SMA, the precentral area, and IFG (Broca’s area, in particular (Dick et al., 2019; Vergani et al., 2014)). Its disconnection may be associated with usually transient mutism in SMA syndromes [354].

Prior to resective lesion surgery, it is recommended to integrate fMRI and diffusion tractography results for interdisciplinary discussion and transfer into the

neuronavigation system. Additionally, asymmetries of language-related fiber pathways may provide useful supporting information when lateralization of language is assessed, although a cautious interpretation is warranted given the paucity of data and initial diverging findings surrounding asymmetry of some tracts (e.g., the arcuate) in individuals with right hemisphere language lateralization (Vernooij et al., 2007).

Appendix E. Pitfalls in the pre-processing and analysis of clinical fMRI

The recent Neuroimaging Analysis Replication and Prediction Study identified substantial variations in analysis workflows and accordingly in the outcomes produced by 70 groups analyzing an identical dataset (Botvinik-Nezer et al., 2020). These findings highlight the importance of and need for standardization of fMRI analysis practices in the field of neuroscience that apply as much (and with potentially even greater consequences) to the analysis of clinical, presurgical single subject fMRI. **Standards and agreed methods for evaluating clinical fMRI workflows do not yet exist and are an important area for development**. The general practices and the consensus in relation to preprocessing and analyzing clinical fMRI data, which translate beyond language mapping, are summarized in Box S3 and Table 2 of the main manuscript, and further expanded on below.

Commercially automated fMRI analysis

A few neuronavigation-compatible analysis tools are available that offer automated fMRI analysis and fusing to surgical planning scans. These generally have the advantage of automated registration to structural images and neuronavigational suites, as well as appropriate (e.g., FDA/CE/equivalent) conformity approvals for use in clinical settings (and which reportedly facilitate both purchasing and reimbursement). Important limitations, however, are that they typically only use basic block-design techniques (some requiring block lengths to be a multiple of the scan TR, so that adapting tasks to patient performance is constrained), and do not (at present) offer flexible options for complex task designs (i.e. task contrasts to analyze) and data processing (e.g. distortion correction, easily examining and comparing motion effects, correction for multiple comparisons or statistical inference protecting from type II rather than type I errors) that may have a profound impact on activation maps and their clinical interpretation. **In-depth evaluation by a trained and experienced user, therefore, remains necessary to confirm the utility and accuracy of fMRI language activation maps** produced by these tools.

Distortions and distortion correction

By optimizing sensitivity to the magnetic field effects of T2*, the basis of the BOLD contrast, Gradient Echo (GE)-EPI is unfortunately also negatively affected by other local magnetic-susceptibility-induced inhomogeneities. These lead to signal loss and spatial image distortions at air-tissue boundaries and, for example, around drilling abrasions, trepanation fixes, inserted catheters or ventricular drainages. Because the distortions in the EPI images compared to T1 images differ, they can easily alter the apparent 'projected' location of fMRI activity with respect to a subject's anatomy. Thus, **spatial image distortions are of obvious clinical concern** (main manuscript, Fig 7) **and need to be corrected.** While signal loss generally cannot be recovered, geometric EPI distortions can be reduced and, to a substantial extent, corrected using a conventionally acquired dual-echo GE fieldmap or approaches based on the acquisition of reversed phase-encoded spin echo images (Andersson et al., 2003). These methods can provide effective distortion correction and can be acquired in approximately 1 minute or less. Further improvements to software for automatic fieldmap processing and correction are desirable.

Head motion and motion correction

Some amount of head motion is inevitable during fMRI and clinical populations generally move more during scanning than healthy research participants. In neuroscientific studies, it is common practice to exclude, from final group analysis, those participants who moved an 'excessive' amount during fMRI, commonly 5mm or more. However, **the direction and frequency of head motion can be more important for the usability of fMRI data than the absolute amount of any single event of movement**. Consequently, outright rejection of data by absolute motion criteria may not be necessary or appropriate, unless head motion is visible during real-time data acquisition and the task can be repeated before the patient leaves the scanner.

The most effective approach for dealing with head motion is prevention. Approaches that help to reduce movement during the scan include: explaining to the patient why head motion is problematic and why they must remain still when the scanner is making noise; using comfortable head and scanner table padding; verifying the patient can see/hear task stimuli in the head coil without moving before starting;

and reassuring patients that small adjustments are permissible when the scanner is not making noise. Some startle movements at the beginning of a run and residual head motion during it are, of course, inevitable and various methods for head motion ‘correction’ exist.

Real-time head motion correction is provided by some scanner platforms and can be combined with *retrospective* approaches (with motion determined from the reconstructed images after the imaging session). Options for dealing with *residual* head motion, after one or both of these motion correction methods are applied, include discarding or excluding specific timepoints, and to model variations in the BOLD signal associated with the estimated motion parameters at the analysis stage. The former can be achieved through a combination of calculating and down-weighting outlier timepoints, and this can also include some or all of the estimated motion parameter time-courses (e.g., rotations and translations) as confound regressors in the statistical model, or removing the variance associated with these motion parameters using data de-noising or “scrubbing” approaches.

In all cases, it is important to inspect individual motion plots and **the effect of correction choices** on the results. In general, more frequent or pronounced head motion in patient (especially pediatric) populations is one reason why **multiple short fMRI tasks are advisable, with real-time visual inspection of the EPI images during acquisition** to spot major problems with head motion and re-instruct the patient prior to repeating any affected scans**.** Motion correction practices varied among this Working Group. **The consensus of this group is that neither form of head motion correction should be assumed to fix all issues**. **We strongly advocate assessing head motion in each fMRI task through a variety of means.** Approaches (ideally combined) include inspecting the raw images and retrospective motion / realignment plots (for both the magnitude of head movement and the pattern, e.g., whether it is correlated with the task design periods), as well as performing multiple analyses to visually compare the effects of different motion correction choices on the activation map. Visual assessments firstly aim to detect artefactual ‘activation’ of the CSF or characteristics ‘rings’ at the edge of the brain. A second aim is to appreciate the spatial extent, appearance and disappearance of activations with/without retrospective motion correction, noting that motion correction

can both remove ‘true’ activations (Glasser et al., 2018) or introduce ‘false’ ones (Yakupov et al., 2017).

Physiological noise / motion and its removal

Respiration and cardiac pulsations present additional sources of “motion” or unwanted signals that can mimic or hide the brain activation signals of interest (Murphy et al., 2013). These signals can be largely removed from the statistical activation maps either through retrospective data ‘cleaning’ (Griffanti et al., 2014), or by prospectively recording these biological traces and removing variance associated with these signals when analyzing the fMRI data (Birn et al., 2006; Brooks et al., 2008; Glover et al., 2000). Physiological noise removal is not generally employed by the members of this working group for clinical fMRI, however, as it is considered more crucial for infratentorial and spinal fMRI.

Brain extraction and statistical mask generation

Some data pre-processing pipelines remove non-brain tissue including the eyes, orbits, skull and dura from the images. For example, low-signal lesions (e.g., Zabramski type IV cavernomas or flow-voids in AVMs) and their perilesional areas (due to “blooming” of signal loss in T2*-weighted EPI) can be erroneously removed in this step, with the consequence of excluding areas of interest from later statistical analysis (Jo et al., 2008). Therefore, if brain extraction / mask generation is part of the initial processing pathway, it is crucial to **verify that brain extraction hasn’t also ‘removed’ the pathological lesion and the area around it from the analysis** and, if it has, then edit or replace the brain mask.

Spatial Smoothing

It is common practice in fMRI analysis to spatially smooth the images in order to increase the SNR (signal-to-noise ratio) and to satisfy minimum smoothness requirements for certain types of statistical analysis. Smoothing leads to noticeable benefits in statistical power, or the ability to distinguish real activations from noise. However, this is only true if the extent of the smoothing is less than the size of the

activations. In clinical fMRI settings, small amounts of smoothing are therefore often favored and large amounts of smoothing avoided, since the latter can spatially displace the focus of detected activations (Andrade et al., 2001; Jo et al., 2008).

Smoothing is also often performed in a research context so that Gaussian Random Field theory (GRF) can be applied in the statistical analysis, which is one way of calculating corrected statistical thresholds. However, in a clinical context the rigid control of false positives at the individual patient level is not necessarily the primary concern, as discussed below (see *Statistical Inference* section).

While smoothing is typically implemented in 3D (i.e., volumetric smoothing), an emerging development is to perform smoothing on 2D cortical surfaces (surface-based smoothing), because smoothing performed on the typical 3D volumes biases the spatial locations of activations and makes discrimination of where activations are coming from more difficult (Coalson et al., 2018). There are advantages in smoothing on the cortical surface, as it avoids smoothing between gray and white matter as well as across sulci (Andrade et al., 2001) and thus improves spatial accuracy (Jo et al., 2007). However, at present, 2D surface-based smoothing requires segmentation of the grey matter surface through a separate (time consuming) pre-processing step using dedicated software to accurately extract the cortical surface and gray/white matter interface from a high-resolution volumetric anatomical scan providing high contrast to noise (CNR) between cortical gray matter and underlying white matter. This step generally ‘fails’ in the presence of brain lesions (such as a tumor or focal cortical dysplasia), such that the cortical surface cannot be accurately defined. Registration of the fMRI to the cortical surface also needs to be very precise so that activations are accurately mapped onto the cortical surface. Because surface reconstructions commonly ‘fail’ at and around intra-axial brain lesions, such as gliomas, **volumetric (3D) smoothing is still the most common form of spatial smoothing for clinical fMRI.** Further development and optimization of 2D segmentation approaches is therefore still needed prior to application in clinical populations.

Smoothing (most often by) a Gaussian kernel with a fixed full-width at half maximum (FWHM) size comes at the cost of smearing out larger activations and wiping out

smaller activations. Fig 8 (main manuscript) illustrates that even a modest amount of smoothing not only expands and blurs the boundaries of larger activations but also obliterates smaller activations. Smoothing away smaller activations may lead to ‘false negative’ interpretations and run the obvious risk of post-operative deficits that clinical fMRI aims to avoid, for example if these areas are indispensable for language processing but surgically damaged because no fMRI activation was reported. Excessive blurring, on the other hand, can also overestimate the extent of activations and could lead to incomplete or suboptimal neurosurgical resections due to preservation of tissue that appears to be activated but does not contribute to function.

Standard approaches to spatial smoothing of fMRI data are not informed by the geometry of the cortex and tissue types. Partial voluming and volumetric smoothing, in particular, mixes the signal of CSF, white matter and cortex of, for example, opposite banks of gyri across a sulcus. Therefore, statistical analysis methods have been proposed, especially for presurgical language fMRI, which avoid smoothing during pre-processing altogether, or conditionally adapt it to different spatial locations in the later analysis stage, to help improve spatial accuracy (Andrade et al., 2001; Johnson et al., 2013; Z. Liu et al., 2016). These approaches do not smooth uniformly across space, but instead adapt the amount of smoothing effectively applied to the data conditionally, so that less smoothing is performed in areas where it is not warranted and desired, i.e., at the interface between activated and null regions. However, the clinical utility and benefits of such methods still need to be evaluated.

Therefore, this Working Group recommends **avoiding or minimizing spatial smoothing**, and if performed, using no more than 1 to 2 times the voxel dimensions (P. Liu et al., 2017), **ideally 4mm (and no more than 5mm)** isotropic FWHM (for a 2 to 3mm voxel dimension). This recommendation minimizes the chances of blurring together non-contiguous cortical speech processing sites (e.g., for word production versus word hearing) identified 4mm apart using high density subdural electrode grids (Flinker et al., 2011) and is in keeping with the 4mm optimal smoothing filter identified by Pajula & Tohka (Pajula & Tohka, 2014) (for data acquired with 2mm isotropic voxels).

Registration of clinical fMRI data and high-resolution structural images

Image registration and resampling are used to align and overlay brain fMRI activations onto high-resolution structural images. The challenge of how to most accurately represent fMRI activation results on an individual patient's anatomy is amplified in the pre-surgical context. Unlike in neuroscientific studies at the group level, it is neither necessary nor advisable to register individual level clinical fMRI data to a standard brain.

Whenever fMRI results are transformed to an anatomical scan, it is important to be aware of the limitations in the underlying source data. If, for example, the mastoid air cells and middle ear pneumatization cause extensive EPI signal loss and distortions extending into and obscuring the inferolateral temporal lobes, this can lead to false interpretations when only overlays of fMRI results onto anatomic images are considered. In some instances, such misinterpretations can be avoided by viewing the raw data in some form.

Typically, fMRI analysis is first conducted in the low-resolution (compared to anatomical scans) fMRI "space". When the clinical question posed of fMRI is to determine which hemisphere is dominant for language, it is typically sufficient to analyze the data in EPI (fMRI) space without resampling the fMRI statistical maps to a high-resolution anatomical scan. However, when the question is to more precisely visualize the location of fMRI activations with respect to a lesion, it is necessary to register (or align) the fMRI results to an appropriate pre- or post-contrast structural image (such as a T1-weighted or FLAIR image). This registration step is needed both because the anatomical detail in the fMRI images is poor, due to limited gray-white contrast and lower spatial resolution, and because pathological features such as hemosiderin staining can obscure the boundaries of the lesion (see Appendix B Fig S1a).

Within many specialist analysis packages, a variety of registration methods are available, including those that apply registration with linear or nonlinear transformations, use tissue segmentations to optimize alignments (e.g., a boundary-based registration (BBR) cost function), and using surface or volumetric approaches. **When it is available / accessible, we recommend BBR as it tends to avoid the**

main errors associated with shifting GE-EPI (and the corresponding fMRI results) downwards to the skull base with respect to the underlying anatomy (so-called “z-shift”; Fig S3). The outcome of any registration approach should always be carefully scrutinized, keeping in mind the fact that signal loss areas (inferior medial frontal and anterior as well as infero-lateral temporal areas) present artefactual boundaries, as the “edge” of the brain in the affected areas is the border of the signal loss rather than a true anatomical boundary.

Once registration has been performed, the fMRI statistical result maps can be rendered onto high-resolution anatomical scans and fed into neuronavigation systems. The latter can be tricky because these systems usually possess limited data import functionalities. Most neuronavigation systems are able to perform some form of automatic registration for the user internally, but are not typically accurate enough when provided only with the (non-aligned) fMRI or tractography results. A practical workaround is to render the relevant fMRI maps and other data such as diffusion tractography results onto the structural scans using overlay tools provided by most fMRI analysis software packages, convert these overlay images (similar to a hybrid exam or ‘burnt-in’ images) to DICOM image format and import these into the neuronavigation for display, disabling any additional (now superfluous) automatic registration features within the neuronavigation system.

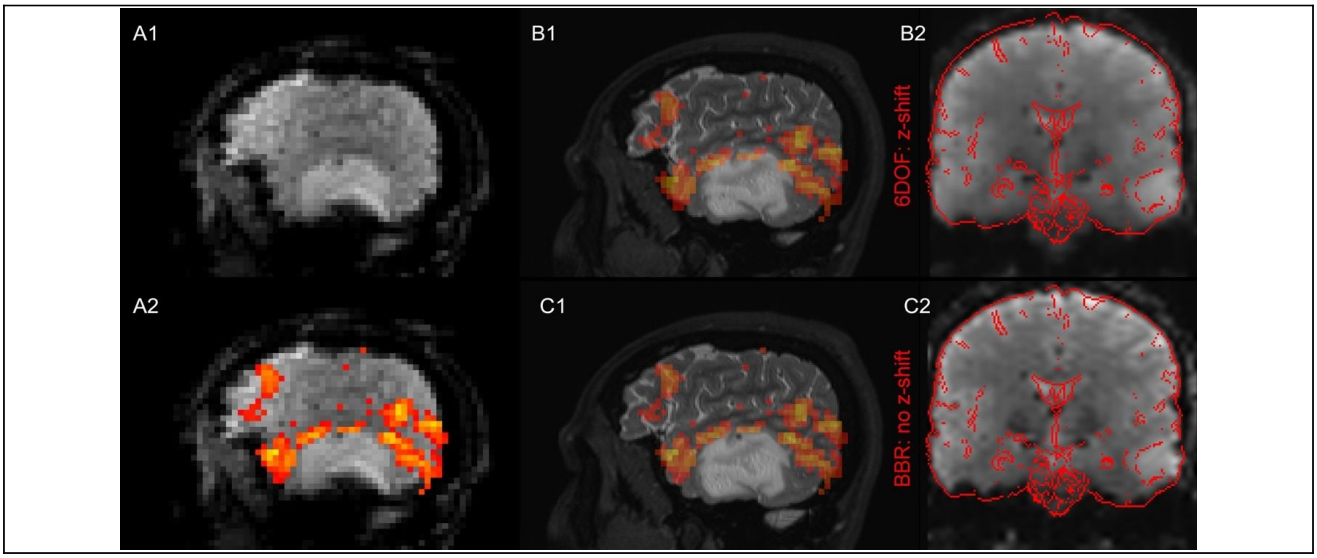


Fig S3. Clinical impact of spatial accuracy / inaccuracies when rendering fMRI onto anatomical scans. *Legend: A1: Distortion-corrected fMRI-EPI. A2: Task-based activation map overlayed onto the underlying fMRI-EPI background (i.e., no registration to an anatomical scan involved) Note that in functional space activations are located just above but not within the tumor. B1: Same fMRI task results transformed and overlaid on the high resolution, T2-weighted structural scan, with activation inaccurately displaced into the tumor (see sagittal image, top of the lesion) due to a - not immediately apparent - misalignment. Note that here activations are, in anatomical space, misplaced into the upper tumor part. B2: Carefully inspecting the registration between the fMRI-EPI (background image) and the anatomical T2-weighted scan (red outline) reveal a small but crucial misalignment due to "z-shift", introduced by EPI signal loss at the skull base and consecutive downward misplacement of the EPI with respect to the contours of the anatomical (see e.g., top of the head and at the level of the ventricles when comparing images B2 and C2) when a rigid-body (6 DOF) registration by a commonly used correlation ratio cost function is applied. Instead, transforming the same fMRI results to anatomical space using boundary-based registration (but still 6 DOF) optimizes the alignment, visible when comparing corresponding activation overlays (A2 and C1) and inspecting the alignment between the EPI image and anatomical scan (C2). Obviously, the clinical consequences are profound: With (C2), the neurosurgeon may aim for cautious but macroscopically total resection while (C1) would falsely suggest "function in tumor" and that only partial resection may be feasible.*

Data analysis and statistical inference

Finally, accurate interpretation of clinical fMRI results requires knowledge of what can vary with different processing options/choices and how to use these methods well.

Hypothesis- vs. data-driven fMRI analysis and the issue of "de-"activations

In principle, fMRI data analysis may be performed using a hypothesis- or data-driven approach. The former case asks the data a specific question (which voxels show signal increase in a way that is correlated with this task design), while data-driven methods ask the data to reveal the structured signals that it can detect. It is recommended to use both, if available, because these are highly complementary and offer their own

advantages and limitations. **Hypothesis-driven general linear model (GLM) fitting** remains the most prevalent of all fMRI analysis approaches, and **is most often applied to analyze clinical fMRI data, also among this Working Group** (for a detailed review of the GLM approach, see (Monti, 2011)). The utility of this approach comes from the fact that, aside from a simple baseline, different active and control conditions can subsequently be contrasted against each other to unveil specific language-related activations.

However, the GLM assumes that the temporal responses / signals consistently evoked by task blocks or events are comprehensively modelled by the design, which determines the testable hypotheses. The GLM attempts to capture brain activations by convolving stimuli (events or task blocks) with a single, uniform, “canonical” hemodynamic response function (HRF) and by using the derived model time series as an explanatory variable (EV) of interest in the GLM. We know, however, that even in healthy individuals, the hemodynamic response varies both between people and across brain regions (Handwerker et al., 2004; Huettel et al., 2001; Miezin et al., 2000; Taylor et al., 2018). This variability is greater still in patients with different brain lesions such as primary brain tumors, with cerebrovascular abnormalities, shunting malformations (e.g., AVM) or instances of potential “neurovascular uncoupling” (NVU). For tasks investigating activations during speech, language and reading, which involve complex cognitive processes performed at different speed, latencies and strategies, the timing and shape of the temporal response may easily deviate from canonical HRF. To better capture such ‘delayed’ responses, temporal derivatives of the EV of interest are often routinely used in GLM analyses. Note that inclusion of temporal derivatives can reduce residual error but does not generally alter the estimated primary signal estimate. The fMRI data is to more likely to diverge from the predefined GLM when more complex processes are probed and/or less challenging tasks are constructed (such as passive story listening vs listening to pseudowords / -sentences), e.g., to facilitate the mapping of children or disabled individuals. Also, unmodeled structured noise (e.g., associated with the ‘resting’ brain networks) is not accurately captured and this can reduce the sensitivity and specificity of hypothesis-driven fMRI analyses using the standard GLM.

Independent component analyses (ICA) have gained favor as a way to identify and statistically separate effects in fMRI data. The data-driven approach of ICA is, per se, not susceptible to deviations from and incompleteness of a prespecified model, and can be used to analyze language task fMRI data. ICA generates its statistical maps in a model-free style. One downside is that different conditions cannot be directly contrasted against each other. However, after ICA has been performed it is possible to temporally correlate the time-courses of the identified components with the time-course of the experimental design, *post hoc,* to determine whether a detected component fits the expected neural activation profile after HRF convolution (Durnez et al., 2013). It is up to the user to decide which degree of correlation (r-value) and which significance levels (p-value) are considered, but generally relaxed criteria (i.e., non-significant p-values) can be recommended to allow for deviations from a simplified model in the time-domain. It is also essential to carefully inspect the spatial maps, because multiple 'language-like' maps can emerge from ICA on a single task, each correlating to some extent with the task design; this is more often the case when the task engages multiple language (sub)processes. Several maps may therefore reflect surgically-meaningful results. In the case of multiple task-related ICA components, it is not always clear how to interpret results or how to decide which map(s) are relevant to show. For the above reasons, and given the relative paucity of published use of ICA for presurgical language fMRI, **this Working Group explicitly encourages clinicians to explore the utility of currently research-based data-driven analysis techniques**.

In both data-driven ICA and model-driven GLM analyses, **'activations' as well as potentially meaningful "de-"activations should be considered (tested for) and inspected. Additional consideration should be given to temporally delayed activation, for example due to delayed task performance or blood flow**. The resulting maps from ICA often easily identify possible delays (or even inversions) in the patient's fMRI signal compared to the expected task time-course (obtained using a standard HRF). Consequently, visual inspection not only of the spatial maps but also of the signal time-course is strongly advocated for clinical fMRI. In the GLM, a temporal shift of the measured response with respect to the expected time-course can mimic either no activations or deactivations (i.e., paradoxical negative BOLD responses), which can be potentially misinterpreted as "neurovascular uncoupling" (NVU; see below) so that the possibility of a temporally delayed activation

needs to be explicitly tested for by introducing variable delays in the permitted/expected time-course in the GLM.

Effect size maps, statistical thresholding and inference

Currently, there is no real consensus on how to best threshold activation maps for clinical purposes. To avoid the occurrence of postsurgical deficits, for example, FN classifications should be minimized, while for maximizing the extent of resection FP detections should be minimal. Additional considerations are that the amplitude of elicited fMRI responses in an individual patient vary substantially based on incompletely understood factors that likely include variable effects of performance and specific medications. The 'optimal' threshold likely differs between patients (W. D. Gaillard, 2000). Therefore, expertise is needed to identify task-relevant activations in each individual. Some practitioners, also among this working group, routinely inspect (and sometimes report) results without thresholding, so as to preserve the continuous nature of the fMRI maps. The decision of thresholds chosen to present results (including unthresholded) should be decided within the multidisciplinary group depending on how the results are to be used (for example, visually inspected for dominance or reported / exported as thresholded maps for neuronavigation).

In neuroscientific research, fMRI results are generated in a way that aims to guard against FP detections. For this purpose, the null hypothesis is formulated as such that *no activation* is assumed. Activations / deactivations are only confirmed by probabilities below a given threshold (usually less than 5% type I error probability) against the null. Given that clinical fMRI primarily strives for FN protection, classical statistical hypothesis testing under the assumption of null activation across the entire brain in a patient who can actually perform the (adjusted, if necessary) task is clearly the wrong question to ask.

Statistical methods of correcting for multiple comparisons (e.g., with Gaussian Random Field theory, maximum height or different cluster statistics to control family-wise error rates FWER) or controlling the false-discovery rate (FDR) are also aiming to control FP rates and hence are often not necessary or advisable in clinically-applied analyses.

Instead, approaches that asymmetrically treat the various types of losses that can occur seem to be generally advantageous in the clinical context (Johnson et al., 2013). In principle, there are three types of losses: misclassification of activated, deactivated or null voxels / clusters. To avoid the occurrence of aphasic / alexic deficits upon surgical resection of an intra-axial brain lesion, it is more important to correctly classify voxels and regions of the brain that are activated during a speech related task than to correctly classify those regions that are not language-critical as null.

Several approaches aiming to control the balance between FNs and FPs are available. These may or may not be suitable for clinical applications and require further evaluations. Alternate hypothesis testing by mixture modelling, both unenhanced (Beckmann & Smith, 2004; Everitt & Bullmore, 1999) or spatially regularized (Hartvig & Jensen, 2000), can be used to explicitly model the tails (which are thought to represent local activations and de-activations) of the distribution of Z-statistics across the brain. Fitting a mixture of, for example, a central Gaussian (for the null) and two Gamma functions (for activations and deactivations in the tails) allows voxels to be labelled as null, activated or "de"-activated at flexible true-positive probabilities / rates (TPRs). Other asymmetric loss functions have been developed specifically for presurgical fMRI applications (Gorgolewski et al., 2012; Johnson et al., 2013; Z. Liu et al., 2016). Alternative-hypothesis-based thresholding reflects evidence against activations and combines different measures of significance for each voxel, producing layered statistical maps, with one layer marking evidence against the traditional null of no activation (classical *p*-value), a second where activation cannot be rejected and a third where activations can be confidently rejected (Durnez et al., 2013).

Preprocessing & analysis considerations	**• BOX S3 •**
o **Perform quality assurance at each step.** Inspect raw images being acquired for excess motion, unexpected signal loss and other image artefacts.	
Image distortions	
o **Image distortions** occur at air-tissue boundaries and around drilling abrasions, trepanation fixes, inserted catheters or ventricular drainages, etc. These are important sources of spatial inaccuracy in clinical fMRI that need to	

be corrected using a separately acquired or generated **fieldmap**.

- **Signal loss** in general **cannot be recovered** but **geometric EPI distortions can be substantially corrected** using a conventionally acquired dual-echo GE fieldmap or approaches using fieldmaps from SE-EPI with different phase encoding directions (Andersson et al., 2003).

Sources of motion

- **Head motion**: Clinical populations tend to move more than research participants. In the worst cases, head motion can make the scan uninterpretable. The **most effective approach** for dealing with head motion is **prevention**.
- For inevitable head motion, **real-time (prospective) head motion correction** is provided by some scanner platforms and, alongside **post-hoc (retrospective) approaches,** may be helpful to generate fMRI data less contaminated by motion. It is important to visualize motion plots and in the case of retrospective correction, visualize the effects of a number of different correction choices on the results.
- The higher incidence of head motion in patient (especially pediatric) populations is one reason why **multiple short fMRI tasks are advisable**, with visual inspection of the EPI images during acquisition to spot problems with head motion, re-instruct the patient and repeat affected scans while they are still in the MRI scanner.
- **Physiological noise**: Respiration and cardiac pulsations present additional unwanted signals that can mimic or hide brain activation signals of interest, especially in inferior regions (Murphy et al., 2013). These **can be largely removed** through retrospective data 'cleaning' (Griffanti et al., 2014)- although this comes with some risks of removing activations of interest - or by prospectively recording these biological traces and removing variance associated with these signals during analysis (Birn et al., 2006; Brooks et al., 2008; Glover et al., 2000). Physiological noise removal is not generally employed by members of this Working Group for clinical fMRI.

Automated brain extraction

- **Inspect and, if necessary, edit brain masks** generated from automated steps. **Low-signal lesions** (e.g., Zabramski type IV cavernomas or flow-voids in AVMs) and perilesional areas (due to “blooming” of signal loss in T2*-weighted EPI) **can be erroneously removed** during automated brain-extraction steps (Haller & Bartsch, 2009). Since the intra- and perilesional area is of interest prior to resective lesion surgery, it is important to ensure that the lesion/peri-lesional area has not been excluded from statistical inference (and therefore inevitably shows ‘no activation’ simply because the area was not included in the analysis).

Appendix F. Cerebrovascular reactivity mapping

Cerebrovascular reactivity mapping (CVRM) has been promoted as an emerging standard of clinical care for presurgical assessment of peri- and intralesional (such as -tumoral) BOLD responsiveness prior or in addition to conventional task- (or resting state-) fMRI (Pillai & Mikulis, 2015).

Originally, CVRM was suggested for measuring the cerebrovascular reserve capacity in steno-occlusive arterial vasculopathies (arteriosclerotic macroangiopathy, Moya-Moya disease and similar disorders). Here, insufficient perfusion and collateralization is presumed to lead to maximal vasodilation irresponsive to hypercapnic stimulation requiring interventional or surgical revascularization. A detailed critical evaluation exceeds the scope of this paper but in case of territorial perfusion deficits the "vasomotor reserve" can often more easily and unambiguously be probed both temporally and quantitatively by transcranial duplex ultrasound, 5 % CO2 / 95 % O2- (such as by carbogen) inhalation and hyperventilation.

CVRM is itself an EPI-fMRI technique based on hypercapnia-induced BOLD signal changes (Bandettini & Wong, 1997). Notably, these exceed the neurogenic BOLD effect an order of magnitude. Thus, CVRM is expected to measure the near maximum cerebrovascular response which can presumably induced. Hypercapnia is readily evoked by breath-hold maneuvers which require less additional equipment than inhalation of hypercapnic gas mixtures. Respiration can be monitored by respiratory belts, and simple end-expiratory (or, if the patient does not tolerate or cannot adequately perform this, mid or end-inspiratory) breath-holding for about 10 to 20 secs can be alternated with blocks of normal breathing for about 40 to 60 secs upon command (Fig S4).

However, a popular misconception is that CVRM is the method of choice to detect "neurovascular uncoupling" (NVU): CVRM cannot detect NVU because it does not use a primary neural stimulus, but rather a respiratory blood-gas stimulus (hypercapnia). In patients with gross delayed perfusion, CVRM may mimic NVU while the cerebrovascular reserve and neurovascular coupling is actually preserved. Therefore,

steno-occlusive extra- and intra-cranial vasculopathies of brain-supplying arteries and consecutive macrovascular territorial perfusion asymmetries should be excluded, and there are good reasons to at least obtain perfusion data to complement each clinical fMRI study prior to CVRM.

At the microvascular level, high-grade brain tumors and AVMs are known to induce neoangiogenesis. Newly formed, pathological vessels are likely to lack normal autoregulation and BOLD responsivity, but so may the pre-existing vascular bed within the microenvironment of a focal brain pathology. CVRM aims to demonstrate such a lack of normal BOLD responsiveness (“NVU”) to avoid FN interpretations of conventional fMRI. However, incomplete and incorrect hypothesis-driven modeling in the temporal domain can cause FNs of CVRM (“GLM pitfall”; while ICA may struggle to converge sensibly in the presence of global hypercapnia-induced BOLD signal changes over the entire cortex). Generally, confirming NVU would require demonstrating a lack of activation, i.e., failing to reject the null hypothesis, which is quite challenging (“statistical pitfall”).

However, CVRM can identify regions with reduced BOLD responsiveness in and around brain lesions (Fig S4). This has potential to localize regions at risk for FNs, prior and in addition to mapping speech, language and reading functions.

This working group considers, at least currently, conventional task fMRI to map language or to establish language lateralization not necessarily superfluous when cerebrovascular reactivity to hypercapnic stimulation is reduced in or around a lesion / vascular territory.

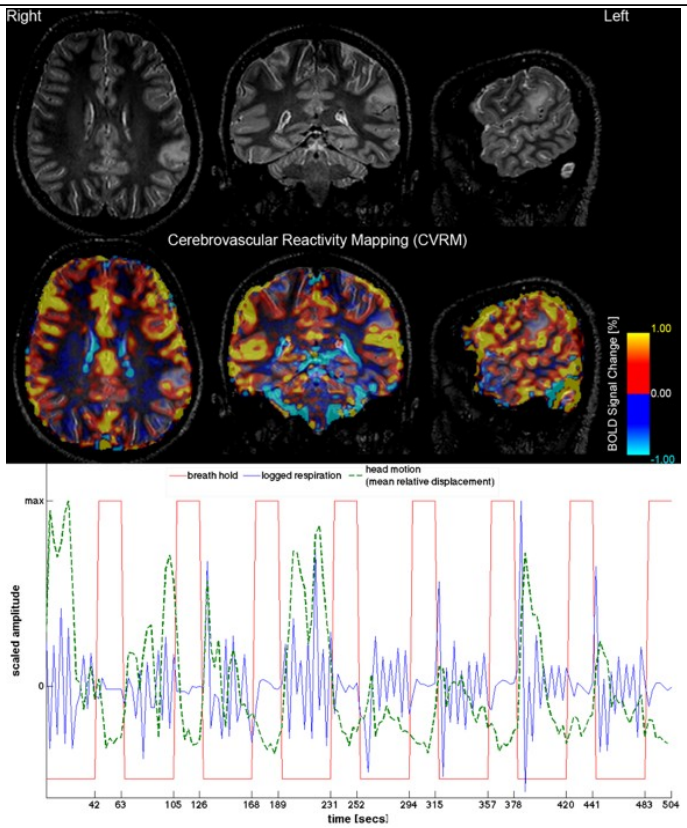


Fig S4: Cerebrovascular reactivity mapping (CVRM) under hypercapnia using breath-holding episodes. *Legend. Top: CVRM identified reduced cerebrovascular BOLD reactivity around a left supramarginal, intra-axial brain lesion. Speech and language fMRI confirmed a relative lack of BOLD responses in and around the lesion (not shown). ESM was performed during awake surgery to guide a safe resection. Bottom: Logged respiration (blue) confirmed patient compliance with breath-hold commands (red), motion correction estimates reveal increased head motion (green) during free breathing.*

Appendix G. Good practices in clinical fMRI reporting

Minimum fMRI reporting checklists and practices have been proposed for the neuroscience community and complement existing guidelines for BOLD-fMRI dictation issued by the ASFNR [†]. Our Working Group recommends specific additions relating to patient performance, description of tasks and analysis methods used, including the following concise points:

Recommended elements of a clinical fMRI report	**• BOX S4 •**

1. Indication

Highlight the clinical purpose for fMRI, i.e., whether performed as part of the presurgical workup for language **lateralization,** for **localization** of specific language functions (e.g., to guide access to an intra-axial lesion), **or a combination of both** (e.g., right hemispheric intra-axial lesion in left-handed/ambidextrous person).

2. Patient history & condition

- Briefly describe the patient's **handedness and current language difficulties**, as well as other known **impairments relevant to language fMRI** (e.g., in memory or ability to focus), preferably with reference to the results of prior comprehensive neuropsychological evaluation.
- Where possible, note recent and current **medications**, especially those that may affect the fMRI BOLD signal and / or performance (e.g., common antiepileptic drugs such as topiramate).

3. fMRI tasks

- **Indicate which paradigm(s) were chosen** and type of presentation (auditory or visual). ‡
- **State if and what adjustments** were made to each task (such as in speed, task duration, number of repetitions) based on known / observed patient performance.

4. Data assessment

- Mention if **prospective motion correction** was used**,** and what data **pre-processing steps** were done (e.g., distortion and retrospective motion correction).
- **Describe quality control measures** (e.g., amount of motion detected, magnetic susceptibility / signal loss artefacts affecting the raw fMRI images especially due to lesion characteristics (e.g., blood products), quality of registration of functional data to the anatomical scan).
- **General analysis approach** used (± any supplemental tests performed,

e.g., ICA, alternative hypothesis testing, if analysis was constrained to prespecified ROIs)

- If final results are shown on the anatomical scan, **highlight if fMRI results involve structures close to or directly affected by signal drop-out /distortion**.
- Note other structures which impact the fMRI results (e.g., large 'activated' draining veins).

5. Findings & Conclusion

- **Mention** any **performance issues** that may impact the fMRI findings, e.g., based on task responses recorded during fMRI or known from prior cognitive testing/practice.
- If relevant, assess **lateralization** of findings and report how these were calculated (e.g., whole-brain or ROI-based, approach to thresholding).
- In the case of resective surgery of a focal lesion, note the **distribution of activation patterns** relative to the lesion. Consider the proximity of relevant language-related activations irrespective of overall laterality (which may be overlook small but essential language sites). Special reference should be made to "function in tumor", displacement of perilesional activations, and surrounding intra- and extra-axial structures of interest.
- **Results of complementary tractography, perfusion or CVRM studies** should ideally be integrated with the fMRI report.

6. Disclaimer

- Indicate that functional mapping by fMRI is based on multiple biophysiological and mathematical assumptions, and results should be interpreted with caution and in conjunction with other tests including neuropsychological data, and if indicated, Wada testing and/or intraoperative electrocortical stimulation.

† https://www.asfnr.org/wp-content/uploads/BOLD-fMRI-Dictation-Guidelines.pdf

‡ It is recommended to indicate in the legend of any generated figures what specific task contrast is shown in a given statistical map (i.e., baseline and control

conditions used for the given language map) and the level of thresholding applied to the t or z-map.

***Presentation and in-person discussion** of the fMRI findings **in a multidisciplinary team meeting is strongly advocated** to consider their impact on: surgical candidacy, likely risks, nature and prognosis of post-surgical deficits, guidance of potential intraoperative stimulation, and implications for presurgical counseling based on multidisciplinary input and consensus. This allows communication around nuances in the interpretation of results. For example, in the case of pediatric patients where a caretaker was present in the MRI room, the timing / type of interaction may be relevant to task performance.

References

1. Pillai JJ. The evolution of clinical functional imaging during the past 2 decades and its current impact on neurosurgical planning. *American Journal of Neuroradiology*. 2010;31(2):219-225. doi:10.3174/ajnr.A1845

2. Junck L, Hervey-Jumper SL, Sagher O. Resection of gliomas around language areas : Can fMRI contribute? *Neurology*. 2015;84(6):550-551. doi:10.1212/WNL.0000000000001241

3. Duffau H. Awake mapping is not an additional surgical technique but an alternative philosophy in the management of low-grade glioma patients. *Neurosurg Rev*. 2018;41(2):689-691. doi:10.1007/s10143-017-0937-6

4. Benjamin CFA, Dhingra I, Li AX, et al. Presurgical language fMRI: Technical practices in epilepsy surgical planning. *Hum Brain Mapp*. 2018;39(10):4032-4042. doi:10.1002/hbm.24229

5. Szaflarski JP, Gloss D, Binder JR, et al. Practice guideline summary: Use of fMRI in the presurgical evaluation of patients with epilepsy. *Neurology*. 2017;88(4):395-402. doi:10.1212/WNL.0000000000003532

6. Black XDF, Vachha XB, Mian XA, et al. American Society of Functional Neuroradiology – Recommended fMRI Paradigm Algorithms for Presurgical Language Assessment. Published online 2017:1-9.

7. Bargalló N, Cano-López I, Rosazza C, et al. Clinical practice of language fMRI in epilepsy centers: a European survey and conclusions by the ESNR Epilepsy Working Group. *Neuroradiology*. 2020;62(5). doi:10.1007/s00234-020-02397-w

8. Benjamin CFA. Cognitive Biomarkers in the Clinic: Lessons From Presurgical fMRI. *J Clin Neurophysiol*. 2022;39(2). doi:10.1097/WNP.0000000000000834

9. Official position of the Division of Clinical Neuropsychology (APA Division 40) on the role of neuropsychologists in Clinical Use of fMRI: Approved by the Division 40 Executive Committee July 28, 2004. *Clinical Neuropsychologist*. 2004;18(3). doi:10.1080/1385404049088718

10. Black DF, Vachha B, Mian A, et al. American Society of Functional Neuroradiology-Recommended fMRI Paradigm Algorithms for Presurgical Language Assessment. *AJNR Am J Neuroradiol*. 2017;38(10):E65-E73. doi:10.3174/ajnr.A5345

11. Carp J. The secret lives of experiments: Methods reporting in the fMRI literature. *Neuroimage*. 2012;63(1):289-300. doi:10.1016/j.neuroimage.2012.07.004

12. Mikl M, Mareček R, Hluštík P, et al. Effects of spatial smoothing on fMRI group inferences. *Magn Reson Imaging*. 2008;26(4). doi:10.1016/j.mri.2007.08.006

13. Coalson TS, Van Essen DC, Glasser MF. The impact of traditional neuroimaging methods on the spatial localization of cortical areas. *Proc Natl Acad Sci U S A*. 2018;115(27). doi:10.1073/pnas.1801582115

14. Nichols TE, Das S, Eickhoff SB, et al. Best practices in data analysis and sharing in neuroimaging using MRI. *Nat Neurosci*. 2017;20(3). doi:10.1038/nn.4500

15. Pernet C, Garrido MI, Gramfort A, et al. Issues and recommendations from the OHBM COBIDAS MEEG committee for reproducible EEG and MEG research. *Nat Neurosci*. 2020;23(12). doi:10.1038/s41593-020-00709-0

16. Ding D, Starke RM, Liu KC, Crowley RW. Cortical plasticity in patients with cerebral arteriovenous malformations. *Journal of Clinical Neuroscience*. 2015;22(12):1857-1861. doi:10.1016/j.jocn.2015.06.014

17. Knecht S, Dräger B, Deppe M, et al. Handedness and hemispheric language dominance in healthy humans. *Brain*. 2000;123(12). doi:10.1093/brain/123.12.2512

18. Somers M, Aukes MF, Ophoff RA, et al. On the relationship between degree of hand-preference and degree of language lateralization. *Brain Lang*. 2015;144:10-15. doi:10.1016/j.bandl.2015.03.006

19. Stewart CC, Swanson SJ, Sabsevitz DS, Rozman ME, Janecek JK, Binder JR. Predictors of language lateralization in temporal lobe epilepsy. *Neuropsychologia*. 2014;60(1). doi:10.1016/j.neuropsychologia.2014.05.021

20. Mazoyer B, Zago L, Jobard G, et al. Gaussian mixture modeling of hemispheric lateralization for language in a large sample of healthy individuals balanced for handedness. *PLoS One*. 2014;9(6):9-14. doi:10.1371/journal.pone.0101165

21. Springer JA, Binder JR, Hammeke TA, et al. Language dominance in neurologically normal and epilepsy subjects. A functional MRI study. *Brain*. 1999;122(11):2033-2045. doi:10.1093/brain/122.11.2033

22. Janecek JK, Swanson SJ, Sabsevitz DS, et al. Language lateralization by fMRI and Wada testing in 229 patients with epilepsy: Rates and predictors of discordance. *Epilepsia*. 2013;54(2):314-322. doi:10.1111/epi.12068

23. Cargnelutti E, Ius T, Skrap M, Tomasino B. What do we know about pre- and postoperative plasticity in patients with glioma? A review of neuroimaging and intraoperative mapping studies. *Neuroimage Clin*. 2020;28(September):102435. doi:10.1016/j.nicl.2020.102435

24. Marcelle M, You X, Fanto EJ, Sepeta LN, Gaillard WD, Berl MM. Impact of development and recent-onset epilepsy on language dominance. *Epilepsia*. 2022;63(10):2637-2649. doi:10.1111/epi.17383

25. Rasmussen T, Milner B. THE ROLE OF EARLY LEFT-BRAIN INJURY IN DETERMINING LATERALIZATION OF CEREBRAL SPEECH FUNCTIONS. *Ann N Y Acad Sci*. 1977;299(1). doi:10.1111/j.1749-6632.1977.tb41921.x

26. Mandonnet E, Mellerio C, Barberis M, Poisson I, Jansma JM, Rutten GJ. When Right Is on the Left (and Vice Versa): A Case Series of Glioma Patients with Reversed Lateralization of Cognitive Functions. *J Neurol Surg A Cent Eur Neurosurg*. 2020;81(2). doi:10.1055/s-0040-1701625

27. DeSalvo MN, Tanaka N, Douw L, et al. Resting-state functional MR imaging for determining language laterality in intractable epilepsy. *Radiology*. 2016;281(1). doi:10.1148/radiol.2016141010

28. Tie Y, Rigolo L, Norton IH, et al. Defining language networks from resting-state fMRI for surgical planning- A feasibility study. *Hum Brain Mapp*. 2014;35(3):1018-1030. doi:10.1002/hbm.22231

29. Cochereau J, Deverdun J, Herbet G, et al. Comparison between resting state fMRI networks and responsive cortical stimulations in glioma patients. *Hum Brain Mapp*. 2016;37(11):3721-3732. doi:10.1002/hbm.23270

30. Parker Jones O, Voets NL, Adcock JE, Stacey R, Jbabdi S. Resting connectivity predicts task activation in pre-surgical populations. *Neuroimage Clin*. 2017;13:378-385. doi:10.1016/j.nicl.2016.12.028

31. Ghinda DC, Wu JS, Duncan NW, Northoff G. How much is enough—Can resting state fMRI provide a demarcation for neurosurgical resection in glioma? *Neurosci Biobehav Rev*. 2018;84. doi:10.1016/j.neubiorev.2017.11.019

32. Park KY, Lee JJ, Dierker D, et al. Mapping language function with task-based vs. resting-state functional MRI. *PLoS One*. 2020;15(7 July). doi:10.1371/journal.pone.0236423

33. Busch RM, Floden DP, Prayson B, et al. Estimating risk of word-finding problems in adults undergoing epilepsy surgery. *Neurology*. 2016;87(22). doi:10.1212/WNL.0000000000003378

34. Thiel A, Habedank B, Winhuisen L, et al. Essential language function of the right hemisphere in brain tumor patients. *Ann Neurol*. 2005;57(1):128-131. doi:10.1002/ana.20342

35. Geranmayeh F, Brownsett SLE, Wise RJS. Task-induced brain activity in aphasic stroke patients: What is driving recovery? *Brain*. 2014;137(10):2632-2648. doi:10.1093/brain/awu163

36. Fetzer S, Dibué M, Nagel AM, Trollmann R. A systematic review of magnetic resonance imaging in patients with an implanted vagus nerve stimulation system. *Neuroradiology*. 2021;63(9). doi:10.1007/s00234-021-02705-y

37. Bartsch AJ, Homola G, Biller A, Solymosi L, Bendszus M. Diagnostic functional MRI: Illustrated clinical applications and decision-making. *Journal of Magnetic Resonance Imaging*. 2006;23(6). doi:10.1002/jmri.20579

38. Peck KK, Bradbury M, Petrovich N, et al. Presurgical evaluation of language using functional magnetic resonance imaging in brain tumor patients with previous surgery. *Neurosurgery*. 2009;64(4):644-652. doi:10.1227/01.NEU.0000339122.01957.0A

39. Hua J, Miao X, Agarwal S, et al. Language Mapping Using T2-Prepared BOLD Functional MRI in the Presence of Large Susceptibility Artifacts—Initial Results in Patients With Brain Tumor and Epilepsy. *Tomography*. 2017;3(2):105-113. doi:10.18383/j.tom.2017.00006

40. Dewey M, Schink T, Dewey CF. Claustrophobia during magnetic resonance imaging: Cohort study in over 55,000 patients. *Journal of Magnetic Resonance Imaging*. 2007;26(5). doi:10.1002/jmri.21147

41. Ragnehed M, Hakansson I, Nilsson M, Lundberg P, Soderfeldt B, Engstrom M. Influence of Diazepam on Clinically Designed fMRI. *J Neuropsychiatr*. 2007;19(2):164-172. doi:10.1176/appi.neuropsych.19.2.164

42. Chwang WB, Iv M, Smith J, et al. Reducing functional MR imaging acquisition times by optimizing workflow. *Radiographics*. 2017;37(1). doi:10.1148/rg.2017160035

43. Bauer PR, Reitsma JB, Houweling BM, Ferrier CH, Ramsey NF. Can fMRI safely replace the Wada test for preoperative assessment of language lateralisation? A meta-analysis and systematic review. *J Neurol Neurosurg Psychiatry*. 2014;85(5):581-588. doi:10.1136/jnnp-2013-305659

44. Binder JR. Functional MRI is a valid noninvasive alternative to Wada testing. *Epilepsy and Behavior*. 2011;20(2):214-222. doi:10.1016/j.yebeh.2010.08.004

45. Deng X, Zhang Y, Xu L, et al. Comparison of language cortex reorganization patterns between cerebral arteriovenous malformations and gliomas: a functional MRI study. *J Neurosurg*. Published online 2015:996-1003. doi:10.3171/2014.12.jns14629

46. Binder JR. The Wernicke area: Modern evidence and a reinterpretation. *Neurology*. 2015;85(24):2170-2175. doi:10.1212/WNL.0000000000002219

47. Price CJ. A review and synthesis of the first 20years of PET and fMRI studies of heard speech, spoken language and reading. *Neuroimage*. 2012;62(2):816-847. doi:10.1016/j.neuroimage.2012.04.062

48. Tremblay P, Dick AS. Broca and Wernicke are dead, or moving past the classic model of language neurobiology. *Brain Lang*. 2016;162:60-71. doi:10.1016/j.bandl.2016.08.004

49. Jansen A, Müller S, Bedenbender J, et al. Determination of crossed language dominance: dissociation of language lateralization within the temporoparietal cortex. *Neurocase*. 2013;19(4):348-350. doi:10.1080/13554794.2012.667129

50. Gaillard WD, Balsamo L, Xu B, et al. fMRI language task panel improves determination of language dominance. *Neurology*. 2004;63(8):1403-1408. https://www.ncbi.nlm.nih.gov/pubmed/15505156

51. Połczyńska M, Japardi K, Curtiss S, et al. Improving language mapping in clinical fMRI through assessment of grammar. *Neuroimage Clin*. 2017;15(October 2016):415-427. doi:10.1016/j.nicl.2017.05.021

52. Seghier ML, Kherif F, Josse G, Price CJ. Regional and hemispheric determinants of language laterality: Implications for preoperative fMRI. *Hum Brain Mapp*. 2011;32(10):1602-1614. doi:10.1002/hbm.21130

53. Vingerhoets G, Verhelst H, Gerrits R, et al. Laterality indices consensus initiative (LICI): A Delphi expert survey report on recommendations to record, assess, and report asymmetry in human behavioural and brain research. *Laterality*. Published online 2023. doi:10.1080/1357650X.2023.2199963

54. Stopa BM, Senders JT, Broekman MLD, Vangel M, Golby AJ. Preoperative functional MRI use in neurooncology patients: A clinician survey. *Neurosurg Focus*. 2020;48(2):E11. doi:10.3171/2019.11.FOCUS19779

55. Thust SC, Heiland S, Falini A, et al. Glioma imaging in Europe: A survey of 220 centres and recommendations for best clinical practice. *Eur Radiol*. 2018;28(8):3306-3317. doi:10.1007/s00330-018-5314-5

56. Binder JR, Swanson SJ, Hammeke TA, Sabsevitz DS. A comparison of five fMRI protocols for mapping speech comprehension systems. *Epilepsia*. 2008;49(12):1980-1997. doi:10.1111/j.1528-1167.2008.01683.x

57. Bradshaw AR, Thompson PA, Wilson AC, Bishop DVM, Woodhead ZVJ. Measuring language lateralisation with different language tasks: a systematic review. *PeerJ*. 2017;5:e3929. doi:10.7717/peerj.3929

58. Carpentier A, Pugh KR, Westerveld M, et al. Functional MRI of language processing: Dependence on input modality and temporal lobe epilepsy. *Epilepsia*. 2001;42(10):1241-1254. doi:10.1046/j.1528-1157.2001.35500.x

59. Gonzálvez GG, Trimmel K, Haag A, et al. Activations in temporal areas using visual and auditory naming stimuli: A language fMRI study in temporal lobe epilepsy. *Epilepsy Res*. 2016;128:102-112. doi:10.1016/j.eplepsyres.2016.10.009

60. Woermann FG, Jokeit H, Luerding R, et al. Language lateralization by Wada test and fMRI in 100 patients with epilepsy. *Neurology*. 2003;61(5):699-701. http://www.ncbi.nlm.nih.gov/pubmed/12963768

61. Benjamin CFA, Li AX, Blumenfeld H, et al. Presurgical language fMRI: Clinical practices and patient outcomes in epilepsy surgical planning. *Hum Brain Mapp*. 2018;39(7):2777-2785. doi:10.1002/hbm.24039

62. Dym RJ, Burns J, Freeman K, Lipton ML. Is functional MR imaging assessment of hemispheric language dominance as good as the wada test?: A meta-analysis. *Radiology*. 2011;261(2):446-455. doi:10.1148/radiol.11101344

63. Massot-Tarrús A, Mousavi SR, Mirsattari SM. Comparing the Intracarotid Amobarbital Test and Functional MRI for the Presurgical Evaluation of Language in Epilepsy. *Curr Neurol Neurosci Rep*. 2017;17(7). doi:10.1007/s11910-017-0763-9

64. Rutten GJM, Ramsey NF, Van Rijen PC, Alpherts WC, Van Veelen CWM. fMRI-determined language lateralization in patients with unilateral or mixed language dominance according to the Wada test. *Neuroimage*. 2002;17(1):447-460. doi:10.1006/nimg.2002.1196

65. Gourovitch ML, Kirkby BS, Goldberg TE, et al. A comparison of rCBF patterns during letter and semantic fluency. *Neuropsychology*. 2000;14(3). doi:10.1037/0894-4105.14.3.353

66. Baldo J V, SChwartz S, Wilkins D, Dronkers NF. Role of frontal versus temporal cortex in verbal fluency as revealed by voxel-based lesion symptom mapping. *Journal of the International Neuropsychological Society*. 2006;12(6):896-900. doi:10.1017/S1355617706061078

67. Barnett A, Marty-Dugas J, McAndrews MP. Advantages of sentence-level fMRI language tasks in presurgical language mapping for temporal lobe epilepsy. *Epilepsy and Behavior*. 2014;32:114-120. doi:10.1016/j.yebeh.2014.01.010

68. Binder JR, Desai RH, Graves WW, Conant LL. Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex*. 2009;19(12):2767-2796. doi:10.1093/cercor/bhp055

69. Lambon Ralph MA, Jefferies E, Patterson K, Rogers TT. The neural and computational bases of semantic cognition. *Nat Rev Neurosci*. 2017;18(1):42-55. doi:10.1038/nrn.2016.150

70. Jackson RL. The neural correlates of semantic control revisited. *Neuroimage*. 2021;224. doi:10.1016/j.neuroimage.2020.117444

71. Arora J, Pugh K, Westerveld M, Spencer S, Spencer DD, Todd Constable R. Language lateralization in epilepsy patients: FMRI validated with the Wada procedure. *Epilepsia*. 2009;50(10):2225-2241. doi:10.1111/j.1528-1167.2009.02136.x

72. Zaca D, Nickerson JP, Deib G, Pillai JJ. Effectiveness of four different clinical fMRI paradigms for preoperative regional determination of language lateralization in

patients with brain tumors. *Neuroradiology*. 2012;54(9):1015-1025. doi:10.1007/s00234-012-1056-2

73. Bookheimer SY, Zeffiro TA, Blaxton TA, Gaillard WD, Malow B, Theodore WH. Regional cerebral blood flow during auditory responsive naming: Evidence for cross-modality neural activation. *Neuroreport*. 1998;9(10). doi:10.1097/00001756-199807130-00047

74. Weyer-Jamora C, Brie MS, Luks TL, et al. Cognitive impact of lower-grade gliomas and strategies for rehabilitation. *Neurooncol Pract*. 2021;8(2):117-128. doi:10.1093/nop/npaa072

75. Stippich C, Mohammed J, Kress B, et al. Robust localization and lateralization of human language function: An optimized clinical functional magnetic resonance imaging protocol. *Neurosci Lett*. 2003;346(1-2):109-113. doi:10.1016/S0304-3940(03)00561-5

76. Partovi S, Jacobi B, Rapps N, et al. Clinical standardized fMRI reveals altered language lateralization in patients with brain tumor. *American Journal of Neuroradiology*. 2012;33(11):2151-2157. doi:10.3174/ajnr.A3137

77. Ekert JO, Kirkman MA, Seghier ML, Green DW, Price CJ. A Data-Based Approach for Selecting Pre- and Intra-Operative Language Mapping Tasks. *Front Neurosci*. 2021;15. doi:10.3389/fnins.2021.743402

78. Dronkers NF, Wilkins DP, Van Valin RD, Redfern BB, Jaeger JJ. Lesion analysis of the brain areas involved in language comprehension. *Cognition*. 2004;92(1-2):145-177. doi:10.1016/j.cognition.2003.11.002

79. Matchin W, Hickok G. The Cortical Organization of Syntax. *Cerebral Cortex*. 2020;30(3). doi:10.1093/cercor/bhz180

80. Freyschlag CF, Krieg SM, Kerschbaumer J, et al. Imaging practice in low-grade gliomas among European specialized centers and proposal for a minimum core of imaging. *J Neurooncol*. 2018;139(3):699-711. doi:10.1007/s11060-018-2916-3

81. Brennan NP, Peck KK, Holodny A. Language Mapping Using fMRI and Direct Cortical Stimulation for Brain Tumor Surgery. *Topics in Magnetic Resonance Imaging*. 2016;25(1):1-10. doi:10.1097/RMR.0000000000000074

82. Binder JR. fMRI of Language Systems: Methods and Applications. In: Faro SH, Mohamed FB, eds. *Functional Neuroradiology: Principles and Clinical Applications*. Springer International Publishing; 2023:565-592. doi:10.1007/978-3-031-10909-6_26

83. Yarkoni T, Poldrack RA, Nichols TE, Van Essen DC, Wager TD. Large-scale automated synthesis of human functional neuroimaging data. *Nat Methods*. 2011;8(8):665-670. doi:10.1038/nmeth.1635

84. Dockes J, Poldrack RA, Primet R, et al. NeuroQuery, comprehensive meta-analysis of human brain mapping. *Elife*. 2020;9. doi:10.7554/eLife.53385

85. Benjamin CF, Walshaw PD, Hale K, et al. Presurgical language fMRI: Mapping of six critical regions. *Hum Brain Mapp*. 2017;38(8):4239-4255. doi:10.1002/hbm.23661

86. Krainik A, Lehericy S, Duffau H, et al. Postoperative speech disorder after medial frontal surgery: role of the supplementary motor area. *Neurology*. 2003;60(4):587-594. http://www.ncbi.nlm.nih.gov/pubmed/12601097

87. Damasio H, Tranel D, Grabowski T, Adolphs R, Damasio A. Neural systems behind word and concept retrieval. *Cognition*. 2004;92(1-2):179-229. doi:10.1016/j.cognition.2002.07.001

88. Drane DL, Pedersen NP. Knowledge of language function and underlying neural networks gained from focal seizures and epilepsy surgery. *Brain Lang*. 2019;189(September 2018):20-33. doi:10.1016/j.bandl.2018.12.007

89. Roux FE, Draper L, Köpke B, Démonet JF. Who actually read Exner? Returning to the source of the frontal " writing centre" hypothesis. *Cortex*. 2010;46(9):1204-1210. doi:10.1016/j.cortex.2010.03.001

90. Chang EF, Kurteff G, Andrews JP, et al. Pure Apraxia of Speech After Resection Based in the Posterior Middle Frontal Gyrus. *Neurosurgery*. 2020;87(3):E383-E389. doi:10.1093/neuros/nyaa002

91. Hazem SR, Awan M, Lavrador JP, et al. Middle Frontal Gyrus and Area 55b: Perioperative Mapping and Language Outcomes. *Front Neurol*. 2021;12. doi:10.3389/fneur.2021.646075

92. Anderson SW, Damasio AR, Damasio H. Troubled letters but not numbers: Domain specific cognitive impairments following focal damage in frontal cortex. *Brain*. 1990;113(3). doi:10.1093/brain/113.3.749

93. Duncan J, Owen AM. Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends Neurosci*. 2000;23(10):475-483. doi:10.1016/s0166-2236(00)01633-7

94. Damasio AR. Time-locked multiregional retroactivation: a systems-level proposal for the neural substrates of recall and recognition. *Cognition*. 1989;33(1-2):25-62. doi:10.1016/0010-0277(89)90005-x

95. Binder JR, Gross WL, Allendorfer JB, et al. Mapping anterior temporal lobe language areas with fMRI: A multicenter normative study. *Neuroimage*. 2011;54(2):1465-1475. doi:10.1016/j.neuroimage.2010.09.048

96. Rao SC, Rainer G, Miller EK. Integration of what and where in the primate prefrontal cortex. *Science (1979)*. 1997;276(5313):821-824. doi:DOI 10.1126/science.276.5313.821

97. Tseng SY, Chettih SN, Arlt C, Barroso-Luque R, Harvey CD. Shared and specialized coding across posterior cortical areas for dynamic navigation decisions. *Neuron*. 2022;110(15):2484-2502 e16. doi:10.1016/j.neuron.2022.05.012

98. Haak K V, Marquand AF, Beckmann CF. Connectopic mapping with resting-state fMRI. *Neuroimage*. 2018;170:83-94. doi:10.1016/j.neuroimage.2017.06.075

99. Affan RO, Scott BB. Everything, everywhere, all at once: Functional specialization and distributed coding in the cerebral cortex. *Neuron*. 2022;110(15):2361-2362. doi:10.1016/j.neuron.2022.07.006

100. Price C, Thierry G, Griffiths T. Speech-specific auditory processing: where is it? *Trends Cogn Sci*. 2005;9(6):271-276. doi:10.1016/j.tics.2005.03.009

101. DeLeon J, Gottesman RF, Kleinman JT, et al. Neural regions essential for distinct cognitive processes underlying picture naming. *Brain*. 2007;130(Pt 5):1408-1422. doi:10.1093/brain/awm011

102. Hillis AE. Aphasia: progress in the last quarter of a century. *Neurology*. 2007;69(2):200-213. doi:10.1212/01.wnl.0000265600.69385.6f

103. Wise RJ. Language systems in normal and aphasic human subjects: functional imaging studies and inferences from animal studies. *Br Med Bull*. 2003;65:95-119. doi:10.1093/bmb/65.1.95

104. Bernal B, Ardila A. The role of the arcuate fasciculus in conduction aphasia. *Brain*. 2009;132(Pt 9):2309-2316. doi:10.1093/brain/awp206

105. Pylkkänen L. The neural basis of combinatory syntax and semantics. *Science (1979)*. 2019;366(6461):62-66. doi:10.1126/science.aax0050

106. Leonard MK, Chang EF. Dynamic speech representations in the human temporal lobe. *Trends Cogn Sci*. 2014;18(9):472-479. doi:10.1016/j.tics.2014.05.001

107. Nelson MJ, El Karoui I, Giber K, et al. Neurophysiological dynamics of phrase-structure building during sentence processing. *Proc Natl Acad Sci U S A*. 2017;114(18):E3669-E3678. doi:10.1073/pnas.1701590114

108. Guidali G, Pisoni A, Bolognini N, Papagno C. Keeping order in the brain: The supramarginal gyrus and serial order in short-term memory. *Cortex*. 2019;119:89-99. doi:10.1016/j.cortex.2019.04.009

109. Fedorenko E, Thompson-Schill SL. Reworking the language network. *Trends Cogn Sci*. 2014;18(3):120-126. doi:10.1016/j.tics.2013.12.006

110. Hagoort P, Indefrey P. The neurobiology of language beyond single words. *Annu Rev Neurosci*. 2014;37:347-362. doi:10.1146/annurev-neuro-071013-013847

111. Amunts K, Zilles K. Architecture and organizational principles of Broca's region. *Trends Cogn Sci*. 2012;16(8):418-426. doi:10.1016/j.tics.2012.06.005

112. Hickok G, Poeppel D. The cortical organization of speech processing. *Nat Rev Neurosci*. 2007;8(5):393-402. doi:10.1038/nrn2113

113. Rauschecker JP, Scott SK. Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nat Neurosci*. 2009;12(6):718-724. doi:10.1038/nn.2331

114. DeWitt I, Rauschecker JP. Wernicke's area revisited: Parallel streams and word processing. *Brain Lang*. 2013;127(2):181-191. doi:10.1016/j.bandl.2013.09.014

115. Chang EF, Raygor KP, Berger MS. Contemporary model of language organization: An overview for neurosurgeons. *J Neurosurg*. 2015;122(2):250-261. doi:10.3171/2014.10.JNS132647

116. Hagoort P. MUC (memory, unification, control) and beyond. *Front Psychol*. 2013;4(JUL). doi:10.3389/fpsyg.2013.00416

117. Middlebrooks EH, Yagmurlu K, Szaflarski JP, Rahman M, Bozkurt B. A contemporary framework of language processing in the human brain in the context of preoperative and intraoperative language mapping. *Neuroradiology*. 2017;59(1):69-87. doi:10.1007/s00234-016-1772-0

118. Purcell JJ, Shea J, Rapp B. Beyond the visual word form area: The orthography-semantics interface in spelling and reading. *Cogn Neuropsychol*. 2014;31(5-6):482-510. doi:10.1080/02643294.2014.909399

119. Shaywitz SE, Shaywitz BA. Paying attention to reading: the neurobiology of reading and dyslexia. *Dev Psychopathol*. 2008;20(4):1329-1349. doi:10.1017/S0954579408000631

120. Roux FE, Durand JB, Réhault E, Planton S, Draper L, Démonet JF. The neural basis for writing from dictation in the temporoparietal cortex. *Cortex*. 2014;50:64-75. doi:10.1016/j.cortex.2013.09.012

121. Rapp B, Purcell J, Hillis AE, Capasso R, Miceli G. Neural bases of orthographic long-term memory and working memory in dysgraphia. *Brain*. 2016;139(2):588-604. doi:10.1093/brain/awv348

122. van Ierschot F, Bastiaanse R, Miceli G. Evaluating Spelling in Glioma Patients Undergoing Awake Surgery: a Systematic Review. *Neuropsychol Rev*. 2018;28(4):470-495. doi:10.1007/s11065-018-9391-7

123. Saur D, Kreher BW, Schnell S, et al. Ventral and dorsal pathways for language. *Proc Natl Acad Sci U S A*. 2008;105(46):18035-18040. doi:10.1073/pnas.0805234105

124. Parker GJM, Luzzi S, Alexander DC, Wheeler-Kingshott CAM, Ciccarelli O, Lambon Ralph MA. Lateralization of ventral and dorsal auditory-language pathways in the human brain. *Neuroimage*. 2005;24(3):656-666. doi:10.1016/j.neuroimage.2004.08.047

125. Mesulam MM, Thompson CK, Weintraub S, Rogalski EJ. The Wernicke conundrum and the anatomy of language comprehension in primary progressive aphasia. *Brain*. 2015;138(8):2423-2437. doi:10.1093/brain/awv154

126. Seghier ML, Neufeld NH, Zeidman P, et al. Reading without the left ventral occipito-temporal cortex. *Neuropsychologia*. 2012;50(14):3621-3635. doi:10.1016/j.neuropsychologia.2012.09.030

127. Fridriksson J, den Ouden DB, Hillis AE, et al. Anatomy of aphasia revisited. *Brain*. 2018;141(3):848-862. doi:10.1093/brain/awx363

128. Head H. *Aphasia and Kindred Disorders of Speech*. Cambridge University Press; 1926.

129. Rutten GJ. Broca-Wernicke theories: A historical perspective. In: *Handbook of Clinical Neurology*. Vol 185. ; 2022. doi:10.1016/B978-0-12-823384-9.00001-3

130. Ojemann GA, Whitaker HA. Language localization and variability. *Brain Lang*. 1978;6(2):239-260. doi:10.1016/0093-934X(78)90061-5

131. Price CJ, Seghier ML, Leff AP. Predicting language outcome and recovery after stroke: The PLORAS system. *Nat Rev Neurol*. 2010;6(4):202-210. doi:10.1038/nrneurol.2010.15

132. Southwell DG, Riva M, Jordan K, et al. Language outcomes after resection of dominant inferior parietal lobule gliomas. *J Neurosurg*. 2017;127(4):781-789. doi:10.3171/2016.8.JNS16443

133. Binder JR, Tong JQ, Pillay SB, et al. Temporal lobe regions essential for preserved picture naming after left temporal epilepsy surgery. *Epilepsia*. 2020;61(9):1939-1948. doi:10.1111/epi.16643

134. Rhoton AL. THE CEREBRUM. *Neurosurgery*. 2007;61(1). doi:10.1227/01.NEU.0000255490.88321.CE

135. Friston KJ, Zarahn E, Josephs O, Henson RNA, Dale AM. Stochastic designs in event-related fMRI. *Neuroimage*. 1999;10(5):607-619. doi:10.1006/nimg.1999.0498

136. Petersen SE, Dubis JW. The mixed block/event-related design. *Neuroimage*. 2012;62(2). doi:10.1016/j.neuroimage.2011.09.084

137. Unadkat P, Fumagalli L, Rigolo L, et al. Functional MRI Task Comparison for Language Mapping in Neurosurgical Patients. *Journal of Neuroimaging*. 2019:1-27.

138. Ekert JO, Lorca-Puls DL, Gajardo-Vidal A, et al. A functional dissociation of the left frontal regions that contribute to single word production tasks. *Neuroimage*. 2021;245. doi:10.1016/j.neuroimage.2021.118734

139. Rutten GJM, Ramsey NF, Van Rijen PC, Noordmans HJ, Van Veelen CWM. Development of a functional magnetic resonance imaging protocol for intraoperative localization of critical temporoparietal language areas. *Ann Neurol*. 2002;51(3):350-360. doi:10.1002/ana.10117

140. Niskanen E, Könönen M, Villberg V, et al. The effect of fMRI task combinations on determining the hemispheric dominance of language functions. *Neuroradiology*. 2012;54(4):393-405. doi:10.1007/s00234-011-0959-7

141. Stippich C, Rapps N, Dreyhaupt J, et al. Localizing and Lateralizing Language in Patients with Brain Tumors: Feasibility of Routine Preoperative Functional MR Imaging

in 81 Consecutive Patients 1. *Radiology*. 2008;243(3):828-836. doi:10.1148/radiol.2433060068

142. Rofes A, Miceli G. Language mapping with verbs and sentences in awake surgery: a review. *Neuropsychol Rev*. 2014;24(2):185-199. doi:10.1007/s11065-014-9258-5

143. De Witte E, Satoer D, Robert E, et al. The Dutch Linguistic Intraoperative Protocol: a valid linguistic approach to awake brain surgery. *Brain Lang*. 2015;140:35-48. doi:10.1016/j.bandl.2014.10.011

144. Bookheimer S. Pre-Surgical Language Mapping with Functional Magnetic Resonance Imaging. *Neuropsychol Rev*. 2007;17(2):145-155. doi:10.1007/s11065-007-9026-x

145. Junck L, Hervey-Jumper SL, Sagher O. Resection of gliomas around language areas : Can fMRI contribute? *Neurology*. 2015;84(6):550-551. doi:10.1212/WNL.0000000000001241

146. Friston KJ, Price CJ. Modules and brain mapping. *Cogn Neuropsychol*. 2011;28(3-4):241-250. doi:10.1080/02643294.2011.558835

147. Morcom AM, Fletcher PC. Does the brain have a baseline? Why we should be resisting a rest. *Neuroimage*. 2007;37(4):1073-1082. doi:10.1016/j.neuroimage.2007.06.019

148. Stark CEL, Squire LR. When zero is not zero: The problem of ambiguous baseline conditions in fMRI. *Proc Natl Acad Sci U S A*. 2001;98(22):12760-12765. doi:10.1073/pnas.221462998

149. Binder JR, Frost JA, Hammeke TA, Bellgowan PSF, Rao SM, Cox RW. Conceptual processing during the conscious resting state: A functional MRI study. *J Cogn Neurosci*. 1999;11(1):80-93. doi:10.1162/089892999563265

150. Hund-Georgiadis M, Lex U, Von Yves Cramon D. Language dominance assessment by means of fMRI: Contributions from task design, performance, and stimulus modality. *Journal of Magnetic Resonance Imaging*. 2001;13(5):668-675. doi:10.1002/jmri.1094

151. Newman SD, Twieg DB, Carpenter PA. Baseline conditions and subtractive logic in neuroimaging. *Hum Brain Mapp*. 2001;14(4):228-235. doi:10.1002/hbm.1055

152. Seghier ML, Price CJ. Dissociating functional brain networks by decoding the between-subject variability. *Neuroimage*. 2009;45(2):349-359. doi:10.1016/j.neuroimage.2008.12.017

153. Binder JR, Frost JA, Hammeke TA, Cox RW, Rao SM, Prieto T. Human brain language areas identified by functional magnetic resonance imaging. *J Neurosci*. 1997;17(1):353-362. https://www.ncbi.nlm.nih.gov/pubmed/8987760

154. Price CJ, Wise RJS, Warburton EA, et al. Hearing and saying. The functional neuro-anatomy of auditory word processing. *Brain*. 1996;119(3):919-931. doi:10.1093/brain/119.3.919

155. Brown EC, Muzik O, Rothermel R, et al. Evaluating reverse speech as a control task with language-related gamma activity on electrocorticography. *Neuroimage*. 2012;60(4):2335-2345. doi:10.1016/j.neuroimage.2012.02.040

156. Gene M, Brennan NP, Holodny AI. Patient Preparation and Paradigm Design in fMRI. *Neuroimaging Clin N Am*. 2021;31(1):11-21. doi:10.1016/j.nic.2020.09.007

157. American College of Radiology. *PRACTICE PARAMETER 1 FMRI BRAIN.*; 2022. https://www.acr.org/-/media/ACR/Files/Practice-Parameters/MR-Perfusion.pdf

158. Poldrack RA, Temple E, Protopapas A, et al. Relations between the neural bases of dynamic auditory processing and phonological processing: evidence from fMRI. *J Cogn Neurosci*. 2001;13(5):687-697. doi:10.1162/089892901750363235

159. Noesselt T, Shah NJ, Jäncke L. Top-down and bottom-up modulation of language related areas - An fMRI study. *BMC Neurosci*. 2003;4. doi:10.1186/1471-2202-4-13

160. Vagharchakian L, Dehaene-Lambertz G, Pallier C, Dehaene S. A Temporal Bottleneck in the Language Comprehension Network. *Journal of Neuroscience*. 2012;32(26):9089-9102. doi:10.1523/jneurosci.5685-11.2012

161. Mechelli A, Friston KJ, Price CJ. The effects of presentation rate during word and pseudoword reading: A comparison of PET and fMRI. *J Cogn Neurosci*. 2000;12(SUPPL. 2):145-156. doi:10.1162/089892900564000

162. Christodoulou JA, Del Tufo SN, Lymberis J, et al. Brain bases of reading fluency in typical reading and impaired fluency in dyslexia. *PLoS One*. 2014;9(7). doi:10.1371/journal.pone.0100552

163. Callicott JH, Mattay VS, Bertolino A, et al. Physiological characteristics of capacity constraints in working memory as revealed by functional MRI. *Cerebral Cortex*. 1999;9(1). doi:10.1093/cercor/9.1.20

164. Thomas ML, Duffy JR, Swerdlow N, Light GA, Brown GG. Detecting the Inverted-U in fMRI Studies of Schizophrenia: A Comparison of Three Analysis Methods. *Journal of the International Neuropsychological Society*. 2022;28(3):258-269. doi:10.1017/S1355617721000424

165. Yen M, DeMarco AT, Wilson SM. Adaptive paradigms for mapping phonological regions in individual participants. *Neuroimage*. 2019;189. doi:10.1016/j.neuroimage.2019.01.040

166. Diachek E, Morgan VL, Wilson SM. Adaptive Language Mapping Paradigms for Presurgical Language Mapping. *American Journal of Neuroradiology*. 2022;43(10):1453-1459. doi:10.3174/ajnr.A7629

167. Yerys BE, Jankowski KF, Shook D, et al. The fMRI success rate of children and adolescents: typical development, epilepsy, attention deficit/hyperactivity disorder, and autism spectrum disorders. *Hum Brain Mapp*. 2009;30(10):3426-3435. doi:10.1002/hbm.20767

168. Benson RR, Logan WJ, Cosgrove GR, et al. Functional MRI localization of language in a 9-year-old child. *Canadian Journal of Neurological Sciences*. 1996;23(3):213-219. doi:10.1017/S0317167100038543

169. Souweidane MM, Kim KH, McDowall R, et al. Brain mapping in sedated infants and young children with passive-functional magnetic resonance imaging. *Pediatr Neurosurg*. 1999;30(2):86-92. doi:10.1159/000028768

170. Vannest J, Karunanayaka PR, Schmithorst VJ, Szaflarski JP, Holland SK. Language networks in children: evidence from functional MRI studies. *AJR Am J Roentgenol*. 2009;192(5):1190-1196. doi:10.2214/AJR.08.2246

171. Suarez RO, Taimouri V, Boyer K, et al. Passive fMRI mapping of language function for pediatric epilepsy surgical planning: validation using Wada, ECS, and FMAER. *Epilepsy Res*. 2014;108(10):1874-1888. doi:10.1016/j.eplepsyres.2014.09.016

172. Wilke M, Groeschel S, Lorenzen A, et al. Clinical application of advanced MR methods in children: points to consider. *Ann Clin Transl Neurol*. 2018;5(11):1434-1455. doi:10.1002/acn3.658

173. Raschle NM, Lee M, Buechler R, et al. Making MR imaging child's play - Pediatric neuroimaging protocol, guidelines and procedure. *Journal of Visualized Experiments*. 2009;(29). doi:10.3791/1309

174. Jones JY, Selvaraj B, Ho ML. Pediatric functional neuroimaging: Practical tips and pearls. *American Journal of Roentgenology*. 2020;214(5). doi:10.2214/AJR.19.22178

175. Liegeois F, Elward R. Functional magnetic resonance imaging. *Handb Clin Neurol*. 2020;174:265-275. doi:10.1016/B978-0-444-64148-9.00019-3

176. Rodin D, Bar-Yosef O, Smith M Lou, Kerr E, Morris D, Donner EJ. Language dominance in children with epilepsy: Concordance of fMRI with intracarotid amytal testing and cortical stimulation. *Epilepsy and Behavior*. 2013;29(1):7-12. doi:10.1016/j.yebeh.2013.06.005

177. Meissner TW, Walbrin J, Nordt M, Koldewyn K, Weigelt S. Head motion during fMRI tasks is reduced in children and adults if participants take breaks. *Dev Cogn Neurosci*. 2020;44. doi:10.1016/j.dcn.2020.100803

178. Berl MM, Duke ES, Mayo J, et al. Functional anatomy of listening and reading comprehension during development. *Brain Lang*. 2010;114(2):115-125. doi:10.1016/j.bandl.2010.06.002

179. Morken F, Helland T, Hugdahl K, Specht K. Reading in dyslexia across literacy development: A longitudinal study of effective connectivity. *Neuroimage*. 2017;144(Pt A):92-100. doi:10.1016/j.neuroimage.2016.09.060

180. Vanderwal T, Kelly C, Eilbott J, Mayes LC, Castellanos FX. Inscapes: A movie paradigm to improve compliance in functional magnetic resonance imaging. *Neuroimage*. 2015;122:222-232. doi:10.1016/j.neuroimage.2015.07.069

181. Berl MM, Mayo J, Parks EN, et al. Regional differences in the developmental trajectory of lateralization of the language network. *Hum Brain Mapp*. 2014;35(1):270-284. doi:10.1002/hbm.22179

182. Austermuehle A, Cocjin J, Reynolds R, et al. Language functional MRI and direct cortical stimulation in epilepsy preoperative planning. *Ann Neurol*. 2017;81(4):526-537. doi:10.1002/ana.24899

183. Holland SK, Plante E, Weber Byars A, Strawsburg RH, Schmithorst VJ, Ball Jr. WS. Normal fMRI brain activation patterns in children performing a verb generation task. *Neuroimage*. 2001;14(4):837-843. doi:10.1006/nimg.2001.0875

184. Brown TT, Lugar HM, Coalson RS, Miezin FM, Petersen SE, Schlaggar BL. Developmental changes in human cerebral functional organization for word generation. *Cereb Cortex*. 2005;15(3):275-290. doi:10.1093/cercor/bhh129

185. Hamberger MJ. Object naming in epilepsy and epilepsy surgery. *Epilepsy Behav*. 2015;46:27-33. doi:10.1016/j.yebeh.2014.12.019

186. Maulisova A, Korman B, Rey G, et al. Atypical language representation in children with intractable temporal lobe epilepsy. *Epilepsy and Behavior*. 2016;58:91-96. doi:10.1016/j.yebeh.2016.03.006

187. Altman N, Bernal B. Paediatric applications of fMRI. In: Faro S, Mohamed FB, eds. *Functional MRI. Basic Principles and Clinical Applications*. Springer Science + Business Media; 2006:394-428.

188. Petrovich Brennan NM, Whalen S, de Morales Branco D, O'Shea JP, Norton IH, Golby AJ. Object naming is a more sensitive measure of speech localization than number counting: Converging evidence from direct cortical stimulation and fMRI. *Neuroimage*. 2007;37(SUPPL. 1):S100-S108. doi:10.1016/j.neuroimage.2007.04.052

189. Bookheimer SY, Zeffiro TA, Blaxton TA, Gaillard W, Theodore WH. Activation of language cortex with automatic speech tasks. *Neurology*. 2000;55(8). doi:10.1212/WNL.55.8.1151

190. Gaillard WD, Pugliese M, Grandin CB, et al. Cortical localization of reading in normal children: An fMRI language study. *Neurology*. 2001;57(1). doi:10.1212/WNL.57.1.47

191. Gaillard WD. Structural and functional imaging in children with partial epilepsy. *Ment Retard Dev Disabil Res Rev*. 2000;6(3):220-226. doi:10.1002/1098-2779(2000)6:3<220::AID-MRDD10>3.0.CO;2-C

192. Polczynska MM, Bookheimer SY. Factors Modifying the Amount of Neuroanatomical Overlap between Languages in Bilinguals-A Systematic Review of Neurosurgical Language Mapping Studies. *Brain Sci*. 2020;10(12). doi:10.3390/brainsci10120983

193. Centeno M, Koepp MJ, Vollmar C, et al. Language dominance assessment in a bilingual population: Validity of fMRI in the second language. *Epilepsia*. 2014;55(10):1504-1511. doi:10.1111/epi.12757

194. Abutalebi J, Green DW. Control mechanisms in bilingual language production: Neural evidence from language switching studies. *Lang Cogn Process*. 2008;23(4):557-582. doi:10.1080/01690960801920602

195. Parker Jones O, Green DW, Grogan A, et al. Where, when and why brain activation differs for bilinguals and monolinguals during picture naming and reading aloud. *Cereb Cortex*. 2012;22(4):892-902. doi:10.1093/cercor/bhr161

196. Calabria M, Costa A, Green DW, Abutalebi J. Neural basis of bilingual language control. *Ann N Y Acad Sci*. Published online 2018. doi:10.1111/nyas.13879

197. Pötzl O. Aphasie und Mehrsprachigkeit. *Z f d, g Neur u Psych*. 1930;124:145-162. https://link.springer.com/content/pdf/10.1007/BF02865094.pdf

198. Leischner A. Über die Aphasie der Mehrsprachigen. *Arch F Psychiatr U Z Neur*. 1948;180:731-775. doi:10.1007/BF01066441

199. Paradis M. Bilingualism and Aphasia. In: Whitaker H, Whitaker HA, eds. *Studies in Neurolinguistics*. Vol 3. Academic Press; 1977:65-121. doi:https://doi.org/10.1016/B978-0-12-746303-2.50008-7

200. Fabbro F, Skrap M, Aglioti S. Pathological switching between languages after frontal lesions in a bilingual patient. *J Neurol Neurosurg Psychiatry*. 2000;68(5):650-652. doi:10.1136/jnnp.68.5.650

201. Price CJ, Green DW, von Studnitz R. A functional imaging study of translation and language switching. *Brain*. 1999;122 (Pt 12):2221-2235. doi:10.1093/brain/122.12.2221

202. Hernandez AE, Dapretto M, Mazziotta J, Bookheimer S. Language switching and language representation in Spanish-English bilinguals: an fMRI study. *Neuroimage*. 2001;14(2):510-520. doi:10.1006/nimg.2001.0810

203. Hernandez AE, Martinez A, Kohnert K. In search of the language switch: An fMRI study of picture naming in Spanish-English bilinguals. *Brain Lang*. 2000;73(3):421-431. doi:10.1006/brln.1999.2278

204. Wang Y, Xue G, Chen C, Xue F, Dong Q. Neural bases of asymmetric language switching in second-language learners: an ER-fMRI study. *Neuroimage*. 2007;35(2):862-870. doi:10.1016/j.neuroimage.2006.09.054

205. Kho KH, Duffau H, Gatignol P, et al. Involuntary language switching in two bilingual patients during the Wada test and intraoperative electrocortical stimulation. *Brain Lang*. 2007;101(1):31-37. doi:10.1016/j.bandl.2006.11.009

206. Moritz-Gasser S, Duffau H. Evidence of a large-scale network underlying language switching: A brain stimulation study. Case report. *J Neurosurg*. 2009;111(4):729-732. doi:10.3171/2009.4.JNS081587

207. Borius PY, Giussani C, Draper L, Roux FE. Sentence translation in proficient bilinguals: a direct electrostimulation brain mapping. *Cortex*. 2012;48(5):614-622. doi:10.1016/j.cortex.2011.01.011

208. Lubrano V, Prod'homme K, Demonet JF, Kopke B. Language monitoring in multilingual patients undergoing awake craniotomy: A case study of a German-English-French trilingual patient with a WHO grade II glioma. *J Neurolinguistics*. 2012;25(6):567-578. doi:10.1016/j.jneuroling.2011.08.002

209. Sierpowska J, Gabarrós A, Ripollés P, et al. Intraoperative electrical stimulation of language switching in two bilingual patients. *Neuropsychologia*. 2013;51(13):2882-2892. doi:10.1016/j.neuropsychologia.2013.09.003

210. Sierpowska J, Fernandez-Coello A, Gomez-Andres A, et al. Involvement of the middle frontal gyrus in language switching as revealed by electrical stimulation mapping and functional magnetic resonance imaging in bilingual brain tumor patients. *Cortex*. 2018;99:78-92. doi:10.1016/j.cortex.2017.10.017

211. Wang X, Wang YY, Jiang T, Wang YZ, Wu CX. Direct evidence of the left caudate's role in bilingual control: An intra-operative electrical stimulation study. *Neurocase*. 2013;19(5):462-469. doi:10.1080/13554794.2012.701635

212. Tomasino B, Marin D, Canderan C, et al. Involuntary switching into the native language induced by electrocortical stimulation of the superior temporal gyrus: A multimodal mapping study. *Neuropsychologia*. 2014;62(1):87-100. doi:10.1016/j.neuropsychologia.2014.07.011

213. Herschmann H, Potzl O. Observations on aphasia in polyglots. In: Paradis M, ed. *Readings on Aphasia in Bilinguals and Polyglots*. Marcel Didier; 1983:148-154.

214. Kruger G, Kastrup A, Glover GH. Neuroimaging at 1.5 T and 3.0 T: comparison of oxygenation-sensitive magnetic resonance imaging. *Magn Reson Med*. 2001;45(4):595-604. https://www.ncbi.nlm.nih.gov/pubmed/11283987

215. Constable RT, Spencer DD. Repetition time in echo planar functional MRI. *Magn Reson Med*. 2001;46(4):748-755. https://www.ncbi.nlm.nih.gov/pubmed/11590651

216. Smith SM, Beckmann CF, Andersson J, et al. Resting-state fMRI in the Human Connectome Project. *Neuroimage*. 2013;80:144-168. doi:10.1016/j.neuroimage.2013.05.039

217. Salimi-Khorshidi G, Douaud G, Beckmann CF, Glasser MF, Griffanti L, Smith SM. Automatic denoising of functional MRI data: combining independent component analysis and hierarchical fusion of classifiers. *Neuroimage*. 2014;90:449-468. doi:10.1016/j.neuroimage.2013.11.046

218. Glasser MF, Smith SM, Marcus DS, et al. The Human Connectome Project's neuroimaging approach. *Nat Neurosci*. 2016;19(9). doi:10.1038/nn.4361

219. Uğurbil K, Xu J, Auerbach EJ, et al. Pushing spatial and temporal resolution for functional and diffusion MRI in the Human Connectome Project. *Neuroimage*. 2013;80:80-104. doi:10.1016/j.neuroimage.2013.05.012

220. Triantafyllou C, Polimeni JR, Wald LL. Physiological noise and signal-to-noise ratio in fMRI with multi-channel array coils. *Neuroimage*. 2011;55(2):597-606. doi:10.1016/j.neuroimage.2010.11.084

221. Bodurka J, Ye F, Petridou N, Murphy K, Bandettini PA. Mapping the MRI voxel volume in which thermal noise matches physiological noise--implications for fMRI. *Neuroimage*. 2007;34(2):542-549. doi:10.1016/j.neuroimage.2006.09.039

222. Glasser MF, Sotiropoulos SN, Wilson JA, et al. The minimal preprocessing pipelines for the Human Connectome Project. *Neuroimage*. 2013;80. doi:10.1016/j.neuroimage.2013.04.127

223. Whiteman AS, Bartsch AJ, Kang J, Johnson TD. BAYESIAN INFERENCE FOR BRAIN ACTIVITY FROM FUNCTIONAL MAGNETIC RESONANCE IMAGING COLLECTED AT TWO SPATIAL RESOLUTIONS. *Annals of Applied Statistics*. 2022;16(4). doi:10.1214/22-AOAS1606

224. Lynch CJ, Power JD, Scult MA, Dubin M, Gunning FM, Liston C. Rapid Precision Functional Mapping of Individuals Using Multi-Echo fMRI. *Cell Rep*. 2020;33(12). doi:10.1016/j.celrep.2020.108540

225. Hamberger MJ, Goodman RR, Perrine K, Tamny T. Anatomic dissociation of auditory and visual naming in the lateral temporal cortex. *Neurology*. 2001;56(1):56-61. doi:10.1212/wnl.56.1.56

226. Ellingson BM, Bendszus M, Boxerman J, et al. Consensus recommendations for a standardized Brain Tumor Imaging Protocol in clinical trials. *Neuro Oncol*. 2015;17(9):1188-1198. doi:10.1093/neuonc/nov095

227. Bernasconi A, Cendes F, Theodore WH, et al. Recommendations for the use of structural magnetic resonance imaging in the care of patients with epilepsy: A consensus report from the International League Against Epilepsy Neuroimaging Task Force. *Epilepsia*. 2019;60(6):1054-1068. doi:10.1111/epi.15612

228. Naganawa S, Nihashi T, Fukatsu H, Ishigaki T, Aoki I. Pre-surgical mapping of primary motor cortex by functional MRI at 3 T: Effects of intravenous administration of Gd-DTPA. *Eur Radiol*. 2004;14(1). doi:10.1007/s00330-003-2147-6

229. Bello L, Castellano A, Fava E, et al. Intraoperative use of diffusion tensor imaging fiber tractography and subcortical mapping for resection of gliomas: technical considerations. *Neurosurg Focus*. 2010;28(2):E6. doi:10.3171/2009.12.focus09240

230. Lazar M, Alexander AL, Thottakara PJ, Badie B, Field AS. White matter reorganization after surgical resection of brain tumors and vascular malformations. *AJNR Am J Neuroradiol*. 2006;27(6):1258-1271. http://www.ncbi.nlm.nih.gov/pubmed/16775277

231. Jones DK, Cercignani M. Twenty-five pitfalls in the analysis of diffusion MRI data. *NMR Biomed*. 2010;23(7):803-820. doi:10.1002/nbm.1543

232. Schilling KG, Rheault F, Petit L, et al. Tractography dissection variability: What happens when 42 groups dissect 14 white matter bundles on the same dataset? *Neuroimage*. 2021;243. doi:10.1016/j.neuroimage.2021.118502

233. Martino J, de Witt Hamer PC, Vergani F, et al. Cortex-sparing fiber dissection: An improved method for the study of white matter anatomy in the human brain. *J Anat*. 2011;219(4):531-541. doi:10.1111/j.1469-7580.2011.01414.x

234. Fernandez-Miranda JC, Pathak S, Engh J, et al. High-Definition Fiber Tractography of the Human Brain. *Neurosurgery*. 2012;71(2):430-453. doi:10.1227/NEU.0b013e3182592faa

235. Yagmurlu K, Middlebrooks EH, Tanriover N, Rhoton AL. Fiber tracts of the dorsal language stream in the human brain. *J Neurosurg*. 2016;124(5):1396-1405. doi:10.3171/2015.5.JNS15455

236. Voets NL, Bartsch A, Plaha P. Brain white matter fibre tracts: a review of functional neuro-oncological relevance. *J Neurol Neurosurg Psychiatry*. Published online 2017. doi:10.1136/jnnp-2017-316170

237. Poldrack RA, Fletcher PC, Henson RN, Worsley KJ, Brett M, Nichols TE. Guidelines for reporting an fMRI study. *Neuroimage*. 2008;40(2):409-414. doi:10.1016/j.neuroimage.2007.11.048

238. Jenkinson M, Chappell M. *Introduction to Neuroimaging Analysis*. (Jenkinson M, Chappell M, eds.). Oxford University Press; 2017.

239. Beisteiner R, Pernet C, Stippich C. Can we standardize clinical functional neuroimaging procedures? *Front Neurol*. 2019;10(JAN). doi:10.3389/fneur.2018.01153

240. Zaitsev M, Maclaren J, Herbst M. Motion artifacts in MRI: A complex problem with many partial solutions. *Journal of Magnetic Resonance Imaging*. 2015;42(4). doi:10.1002/jmri.24850

241. Maclaren J, Herbst M, Speck O, Zaitsev M. Prospective motion correction in brain imaging: A review. *Magn Reson Med*. 2013;69(3). doi:10.1002/mrm.24314

242. Thesen S, Heid O, Mueller E, Schad LR. Prospective Acquisition Correction for head motion with image-based tracking for real-time fMRI. *Magn Reson Med*. 2000;44(3). doi:10.1002/1522-2594(200009)44:3<457::AID-MRM17>3.0.CO;2-R

243. Gorgolewski KJ, Storkey AJ, Bastin ME, Whittle I, Pernet C. Single subject fMRI test-retest reliability metrics and confounding factors. *Neuroimage*. 2013;69. doi:10.1016/j.neuroimage.2012.10.085

244. Haller S, Bartsch AJ. Pitfalls in FMRI. *Eur Radiol*. 2009;19(11):2689-2706. doi:10.1007/s00330-009-1456-9

245. Jo HJ, Lee JM, Kim JH, et al. Artificial shifting of fMRI activation localized by volume- and surface-based analyses. *Neuroimage*. 2008;40(3):1077-1089. doi:10.1016/j.neuroimage.2007.12.036

246. Pajula J, Tohka J. Effects of spatial smoothing on inter-subject correlation based analysis of FMRI. *Magn Reson Imaging*. 2014;32(9):1114-1124. doi:10.1016/j.mri.2014.06.001

247. Liu P, Calhoun V, Chen Z. Functional overestimation due to spatial smoothing of fMRI data. *J Neurosci Methods*. 2017;291:1-12. doi:10.1016/j.jneumeth.2017.08.003

248. Flinker A, Chang EF, Barbaro NM, Berger MS, Knight RT. Sub-centimeter language organization in the human temporal lobe. *Brain Lang*. 2011;117(3):103-109. doi:10.1016/j.bandl.2010.09.009

249. Miezin FM, Maccotta L, Ollinger JM, Petersen SE, Buckner RL. Characterizing the hemodynamic response: Effects of presentation rate, sampling procedure, and the possibility of ordering brain activity based on relative timing. *Neuroimage*. 2000;11(6 I):735-759. doi:10.1006/nimg.2000.0568

250. Huettel SA, Singerman JD, McCarthy G. The effects of aging upon the hemodynamic response measured by functional MRI. *Neuroimage*. 2001;13(1):161-175. doi:10.1006/nimg.2000.0675

251. Handwerker DA, Ollinger JM, D'Esposito M. Variation of BOLD hemodynamic responses across subjects and brain regions and their effects on statistical analyses. *Neuroimage*. 2004;21(4):1639-1651. doi:10.1016/j.neuroimage.2003.11.029

252. Taylor AJ, Kim JH, Ress D. Characterization of the hemodynamic response function across the majority of human cerebral cortex. *Neuroimage*. 2018;173:322-331. doi:10.1016/j.neuroimage.2018.02.061

253. Jacobs J, Hawco C, Kobayashi E, et al. Variability of the hemodynamic response as a function of age and frequency of epileptic discharge in children with epilepsy. *Neuroimage*. 2008;40(2):601-614. doi:10.1016/j.neuroimage.2007.11.056

254. Bonakdarpour B, Parrish TB, Thompson CK. Hemodynamic response function in patients with stroke-induced aphasia: implications for fMRI data analysis. *Neuroimage*. 2007;36(2):322-331. doi:10.1016/j.neuroimage.2007.02.035

255. Roc AC, Wang J, Ances BM, Liebeskind DS, Kasner SE, Detre JA. Altered hemodynamics and regional cerebral blood flow in patients with hemodynamically significant stenoses. *Stroke*. 2006;37(2):382-387. doi:10.1161/01.STR.0000198807.31299.43

256. Rutten GJ. What fMRI can contribute to clinical care. In: Rutten GJ, ed. *The Broca-Wernicke Doctrine*. Springer International Publishing; 2017:251-267. doi:10.1007/978-3-319-54633-9

257. Durnez J, Moerkerke B, Bartsch A, Nichols TE. Alternative-based thresholding with application to presurgical fMRI. *Cogn Affect Behav Neurosci*. 2013;13(4):703-713. doi:10.3758/s13415-013-0185-3

258. Johnson TD, Liu Z, Bartsch AJ, Nichols TE. A Bayesian non-parametric Potts model with application to pre-surgical FMRI data. *Stat Methods Med Res*. 2013;22(4):364-381. doi:10.1177/0962280212448970

259. Liu Z, Berrocal VJ, Bartsch AJ, Johnson TD. Pre-Surgical fMRI Data Analysis Using a Spatially Adaptive Conditionally Autoregressive Model. *Bayesian Anal*. 2016;11(2):599-625. doi:10.1214/15-BA972

260. Gorgolewski KJ, Storkey AJ, Bastin ME, Pernet CR. Adaptive thresholding for reliable topological inference in single subject fMRI analysis. *Front Hum Neurosci*. 2012; (AUGUST). doi:10.3389/fnhum.2012.00245

261. Gross WL, Binder JR. Alternative thresholding methods for fMRI data optimized for surgical planning. *Neuroimage*. 2014;84. doi:10.1016/j.neuroimage.2013.08.066

262. Higgins J, Barbieri E, Wang X, et al. Reliability of BOLD signals in chronic stroke-induced aphasia. *Eur J Neurosci*. 2020;52(8):3963-3978. doi:10.1111/ejn.14739

263. Yasuda C, Vollmar C, Centeno M, et al. The effect of topiramate on verbal fluency fMRI: a longitudinal pilot study. *Epilepsia*. 2002;43:1-351. doi:10.1046/j.1528-1157.43.s7.1.x

264. Genovese CR, Lazar NA, Nichols T. Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *Neuroimage*. 2002;15(4):870-878. doi:10.1006/nimg.2001.1037

265. Hou BL, Bradbury M, Peck KK, Petrovich NM, Gutin PH, Holodny AI. Effect of brain tumor neovasculature defined by rCBV on BOLD fMRI activation volume in the primary motor cortex. *Neuroimage*. 2006;32(2):489-497. doi:10.1016/j.neuroimage.2006.04.188

266. Zaca D, Jovicich J, Nadar SR, Voyvodic JT, Pillai JJ. Cerebrovascular Reactivity Mapping in Patients With Low Grade Gliomas Undergoing Presurgical Sensorimotor Mapping With BOLD fMRI. *Journal of Magnetic Resonance Imaging*. 2014;40(2):383-390. doi:10.1002/jmri.24406

267. Ulmer JL, Hacein-Bey L, Mathews VP, et al. Lesion-induced pseudo-dominance at functional magnetic resonance imaging: Implications for preoperative assessments. *Neurosurgery*. 2004;55(3):569-579. doi:10.1227/01.NEU.0000134384.94749.B2

268. Pillai JJ, Mikulis DJ. Cerebrovascular reactivity mapping: An evolving standard for clinical functional imaging. *American Journal of Neuroradiology*. 2015;36(1):7-13. doi:10.3174/ajnr.A3941

269. Ruff IM, Petrovich Brennan NM, Peck KK, et al. Assessment of the language laterality index in patients with brain tumor using functional MR imaging: Effects of thresholding, task selection, and prior surgery. *American Journal of Neuroradiology*. 2008;29(3):528-535. doi:10.3174/ajnr.A0841

270. Seghier ML. Laterality index in functional MRI: methodological issues. *Magn Reson Imaging*. 2008;26(5):594-601. doi:10.1016/j.mri.2007.10.010

271. Sabsevitz DS, Swanson SJ, Hammeke TA, et al. Use of preoperative functional neuroimaging to predict language deficits from epilepsy surgery. *Neurology*. 2003;60(11):1788-1792. http://www.ncbi.nlm.nih.gov/pubmed/12796532

272. Suarez RO, Whalen S, O'Shea JP, Golby AJ. A Surgical Planning Method for Functional MRI Assessment of Language Dominance: Influences from Threshold, Region-of-Interest, and Stimulus Mode. *Brain Imaging Behav*. 2008;2(2):59-73. doi:10.1007/s11682-007-9018-8

273. Pillai JJ, Zaca D. Relative utility for hemispheric lateralization of different clinical fMRI activation tasks within a comprehensive language paradigm battery in brain tumor patients as assessed by both threshold-dependent and threshold-independent analysis methods. *Neuroimage*. 2011;54(SUPPL. 1):S136-S145. doi:10.1016/j.neuroimage.2010.03.082

274. Knecht S, Flöel A, Dräger B, et al. Degree of language lateralization determines susceptibility to unilateral brain lesions. *Nat Neurosci*. 2002;5(7). doi:10.1038/nn868

275. Jones SE, Mahmoud SY, Phillips MD. A practical clinical method to quantify language lateralization in fMRI using whole-brain analysis. *Neuroimage*. 2011;54(4):2937-2949. doi:10.1016/j.neuroimage.2010.10.052

276. Szaflarski JP, Holland SK, Jacola LM, Lindsell C, Privitera MD, Szaflarski M. Comprehensive presurgical functional MRI language evaluation in adult patients with epilepsy. *Epilepsy Behav*. 2008;12(1):74-83. doi:10.1016/j.yebeh.2007.07.015

277. Kemp S, Wilkinson K, Caswell H, Reynders H, Baker G. The base rate of Wada test failure. *Epilepsy Behav*. 2008;13(4):630-633. doi:10.1016/j.yebeh.2008.07.013

278. Papanicolaou AC, Rezaie R, Narayana S, et al. Is it time to replace the Wada test and put awake craniotomy to sleep? *Epilepsia*. 2014;55(5):629-632. doi:10.1111/epi.12569

279. Wilke M, Lidzba K. LI-tool: A new toolbox to assess lateralization in functional MR-data. *J Neurosci Methods*. 2007;163(1). doi:10.1016/j.jneumeth.2007.01.026

280. Uğurbil K, Toth L, Kim DS. How accurate is magnetic resonance imaging of brain function? *Trends Neurosci*. 2003;26(2):108-114. doi:10.1016/S0166-2236(02)00039-5

281. Kuchcinski G, C. M, J. P, et al. Three-tesla functional MR language mapping : Comparison with direct cortical stimulation in gliomas. *Neurology*. 2015;84(6):560-568. doi:http://dx.doi.org/10.1212/WNL.0000000000001226

282. Sierpowska J, Rofes A, Dahlslätt K, et al. The Aftercare Survey: Assessment and intervention practices after brain tumor surgery in Europe. *Neurooncol Pract*. 2022;9(4). doi:10.1093/nop/npac029

283. Savoy RL. Evolution and current challenges in the teaching of functional MRI and functional brain imaging. *Neuroimage*. 2012;62(2):1201-1207. doi:10.1016/j.neuroimage.2012.02.029

284. Petrella JR, Shah LM, Harris KM, et al. Preoperative functional MR imaging localization of language and motor areas: effect on therapeutic decision making in patients with potentially resectable brain tumors. *Radiology*. 2006;240(3):793-802. doi:10.1148/radiol.2403051153

285. Incekara F, Olubiyi O, Ozdemir A, Lee T, Rigolo L, Golby A. The Value of Pre- and Intraoperative Adjuncts on the Extent of Resection of Hemispheric Low-Grade Gliomas: A Retrospective Analysis. *J Neurol Surg A Cent Eur Neurosurg*. 2015;77(2). doi:10.1055/s-0035-1551830

286. Ius T, Isola M, Budai R, et al. Low-grade glioma surgery in eloquent areas: Volumetric analysis of extent of resection and its impact on overall survival. A single-institution experience in 190 patients - Clinical article. *J Neurosurg*. 2012;117(6):1039-1052. doi:10.3171/2012.8.JNS12393

287. Liegeois F, Cross JH, Gadian DG, Connelly A. Role of fMRI in the decision-making process: epilepsy surgery for children. *J Magn Reson Imaging*. 2006;23(6):933-940. doi:10.1002/jmri.20586

288. de Ribaupierre S, Fohlen M, Bulteau C, et al. Presurgical language mapping in children with epilepsy: clinical usefulness of functional magnetic resonance imaging for the planning of cortical stimulation. *Epilepsia*. 2012;53(1):67-78. doi:10.1111/j.1528-1167.2011.03329.x

289. Giussani C, Roux FE, Ojemann J, Sganzerla E Pietro, Pirillo D, Papagno C. Is preoperative functional magnetic resonance imaging reliable for language areas mapping in brain tumor surgery? Review of language functional magnetic resonance imaging and direct cortical stimulation correlation studies. *Neurosurgery*. 2010;66(1):113-120. doi:10.1227/01.NEU.0000360392.15450.C9

290. Weng HH, Noll KR, Johnson JM, et al. Accuracy of Presurgical Functional MR Imaging for Language Mapping of Brain Tumors: A Systematic Review and Meta-Analysis. *Radiology*. 2018;286(2):512-523. doi:10.1148/radiol.2017162971

291. Borchers S, Himmelbach M, Logothetis N, Karnath HO. Direct electrical stimulation of human cortex - the gold standard for mapping brain functions? *Nat Rev Neurosci*. 2012;13(1):63-70. doi:10.1038/nrn3140

292. Nossek E, Matot I, Shahar T, et al. Failed awake craniotomy: A retrospective analysis in 424 patients undergoing craniotomy for brain tumor ; Clinical article. *J Neurosurg*. 2013;118(2). doi:10.3171/2012.10.JNS12511

293. Talacchi A, Santini B, Casagrande F, Alessandrini F, Zoccatelli G, Squintani GM. Awake surgery between art and science. Part I: Clinical and operative settings. *Funct Neurol*. 2013;28(3). doi:10.11138/FNeur/2013.28.3.205

294. Mandonnet E, Sarubbo S, Duffau H. Proposal of an optimized strategy for intraoperative testing of speech and language during awake mapping. *Neurosurg Rev*. 2017;40(1):29-35. doi:10.1007/s10143-016-0723-x

295. Roux FE, Djidjeli I, Quéhan R, Réhault E, Giussani C, Durand JB. Intraoperative electrostimulation for awake brain mapping: How many positive interference responses are required for reliability? *J Neurosurg*. 2020;133(4). doi:10.3171/2019.6.JNS19925

296. Ojemann GA. Individual variability in cortical localization of language. *J Neurosurg*. 1979;50(2):164-169. doi:10.3171/jns.1979.50.2.0164

297. Ojemann G, Ojemann J, Lettich E, Berger M. Cortical language localization in left, dominant hemisphere. An electrical stimulation mapping investigation in 117 patients. *J Neurosurg*. 1989;71(3):316-326. doi:10.3171/jns.1989.71.3.0316

298. Mandonnet E, Winkler PA, Duffau H. Direct electrical stimulation as an input gate into brain functional networks: Principles, advantages and limitations. *Acta Neurochir (Wien)*. 2010;152(2):185-193. doi:10.1007/s00701-009-0469-0

299. Perrone-Bertolotti M, Alexandre S, Jobb AS, et al. Probabilistic mapping of language networks from high frequency activity induced by direct electrical stimulation. *Hum Brain Mapp*. 2020;41(14):4113-4126. doi:10.1002/hbm.25112

300. Hamberger MJ, Williams AC, Schevon CA. Extraoperative neurostimulation mapping: Results from an international survey of epilepsy surgery programs. *Epilepsia*. 2014;55(6). doi:10.1111/epi.12644

301. De Witt Hamer PC, Robles SG, Zwinderman AH, Duffau H, Berger MS. Impact of Intraoperative Stimulation Brain Mapping on Glioma Surgery Outcome: A Meta-Analysis. *Journal of Clinical Oncology*. 2012;30(20):2559-2565. doi:10.1200/JCO.2011.38.4818

302. Holloway T, Leach JL, Tenney JR, et al. Functional MRI and electrical stimulation mapping for language localization: A comparative meta-analysis. *Clin Neurol Neurosurg*. 2022;222. doi:10.1016/j.clineuro.2022.107417

303. Trinh VT, Fahim DK, Maldaun MVCC, et al. Impact of preoperative functional magnetic resonance imaging during awake craniotomy procedures for intraoperative guidance and complication avoidance. *Stereotact Funct Neurosurg*. 2014;92(5):315-322. doi:10.1159/000365224

304. Rolinski R, Austermuehle A, Wiggs E, et al. Functional MRI and direct cortical stimulation: Prediction of postoperative language decline. *Epilepsia*. 2019;60(3):560-570. doi:10.1111/epi.14666

305. Crow AJD, Thomas A, Rao Y, et al. Task-based functional magnetic resonance imaging prediction of postsurgical cognitive outcomes in temporal lobe epilepsy: A systematic review, meta-analysis, and new data. *Epilepsia*. 2023;64(2). doi:10.1111/epi.17475

306. Binder JR. Preoperative prediction of verbal episodic memory outcome using FMRI. *Neurosurg Clin N Am*. 2011;22(2):219-232, ix. doi:10.1016/j.nec.2010.12.002

307. Janecek JK, Swanson SJ, Sabsevitz DS, et al. Naming outcome prediction in patients with discordant Wada and fMRI language lateralization. *Epilepsy and Behavior*. 2013;27(2):399-403. doi:10.1016/j.yebeh.2013.02.030

308. Krainik A, Duffau H, Capelle L, et al. Postoperative speech disorder after medial frontal surgery SMA. *Neurology*. 2003;60:391-396.

309. Labudda K, Mertens M, Kalbhenn T, Schulz R, Woermann FG. Partial resection of presurgical fMRI activation is associated with a postsurgical loss of language function after frontal lobe epilepsy surgery. *Neurocase*. 2017;23(3-4):239-248. doi:10.1080/13554794.2017.1383445

310. You X, Zachery AN, Fanto EJ, et al. fMRI prediction of naming change after adult temporal lobe epilepsy surgery: Activation matters. *Epilepsia*. 2019;60(3):527-538. doi:10.1111/epi.14656

311. Lyo JK, Arevalo-Perez J, Petrovich Brennan N, Peck KK, Holodny AI. Pre-operative fMRI localization of the supplementary motor area and its relationship with postoperative speech deficits. *Neuroradiology Journal*. 2015;28(3):281-288. doi:10.1177/1971400915589681

312. Vysotski S, Madura C, Swan B, et al. Preoperative FMRI Associated with Decreased Mortality and Morbidity in Brain Tumor Patients. *Interdiscip Neurosurg*. 2018;13:40-45. doi:10.1016/j.inat.2018.02.001

313. Lin F, Jiao Y, Wu J, et al. Effect of functional MRI-guided navigation on surgical outcomes: a prospective controlled trial in patients with arteriovenous malformations. *J Neurosurg*. 2017;126(6):1863-1872. doi:10.3171/2016.4.JNS1616

314. Tan L, Holland SK, Deshpande AK, Chen Y, Choo DI, Lu LJ. A semi-supervised Support Vector Machine model for predicting the language outcomes following cochlear implantation based on pre-implant brain fMRI imaging. *Brain Behav*. 2015;5(12):1-25. doi:10.1002/brb3.391

315. Yasuda CL, Centeno M, Vollmar C, et al. The effect of topiramate on cognitive fMRI. *Epilepsy Res*. 2013;105(1-2):250-255. doi:10.1016/j.eplepsyres.2012.12.007

316. Jokeit H, Okujava M, Woermann FG. Carbamazepine reduces memory induced activation of mesial temporal lobe structures: a pharmacological fMRI-study. *BMC Neurol*. 2001;1:6. doi:10.1186/1471-2377-1-6

317. Keulers EHH, Stiers P, Nicolson NA, Jolles J. The association between cortisol and the BOLD response in male adolescents undergoing fMRI. *Brain Res*. 2015;1598. doi:10.1016/j.brainres.2014.12.022

318. Poldrack RA, Wagner AD, Prull MW, Desmond JE, Glover GH, Gabrieli JD. Functional specialization for semantic and phonological processing in the left inferior prefrontal cortex. *Neuroimage*. 1999;10(1):15-35. doi:10.1006/nimg.1999.0441

319. Loh KK, Procyk E, Neveu R, et al. Cognitive control of orofacial motor and vocal responses in the ventrolateral and dorsomedial human frontal cortex. *Proc Natl Acad Sci U S A*. 2020;117(9):4994-5005. doi:10.1073/pnas.1916459117

320. Sohoglu E, Peelle JE, Carlyon RP, Davis MH. Predictive top-down integration of prior knowledge during speech perception. *Journal of Neuroscience*. 2012;32(25):8443-8453. doi:10.1523/JNEUROSCI.5069-11.2012

321. Kroczek LOH, Gunter TC, Rysop AU, Friederici AD, Hartwigsen G. Contributions of left frontal and temporal cortex to sentence comprehension: Evidence from simultaneous TMS-EEG. *Cortex*. 2019;115:86-98. doi:10.1016/j.cortex.2019.01.010

322. Friederici AD, Chomsky N, Berwick RC, Moro A, Bolhuis JJ. Language, mind and brain. *Nat Hum Behav*. 2017;1(10):713-722. doi:10.1038/s41562-017-0184-4

323. Rodd JM, Vitello S, Woollams AM, Adank P. Localising semantic and syntactic processing in spoken and written language comprehension: An Activation Likelihood Estimation meta-analysis. *Brain Lang*. 2015;141:89-102. doi:10.1016/j.bandl.2014.11.012

324. Corina DP, Loudermilk BC, Detwiler L, Martin RF, Brinkley JF, Ojemann G. Analysis of naming errors during cortical stimulation mapping: implications for models of language representation. *Brain Lang*. 2010;115(2):101-112. doi:10.1016/j.bandl.2010.04.001

325. Tate MC, Herbet G, Moritz-Gasser S, Tate JE, Duffau H. Probabilistic map of critical functional regions of the human cerebral cortex: Broca's area revisited. *Brain*. 2014;137(10):2773-2782. doi:10.1093/brain/awu168

326. Lubrano V, Filleron T, Demonet JF, Roux FE. Anatomical correlates for category-specific naming of objects and actions: a brain stimulation mapping study. *Hum Brain Mapp*. 2014;35(2):429-443. doi:10.1002/hbm.22189

327. Lu J, Zhao Z, Zhang J, et al. Functional maps of direct electrical stimulation-induced speech arrest and anomia: A multicentre retrospective study. *Brain*. 2021;144(8):2541-2553. doi:10.1093/brain/awab125

328. Schwartz MF, Faseyitan O, Kim J, Coslett HB. The dorsal stream contribution to phonological retrieval in object naming. *Brain*. 2012;135(12):3799-3814. doi:10.1093/brain/aws300

329. Bährend I, Muench MR, Schneider H, et al. Incidence and linguistic quality of speech errors: A comparison of preoperative transcranial magnetic stimulation and intraoperative direct cortex stimulation. *J Neurosurg*. 2021;134(5):1409-1418. doi:10.3171/2020.3.JNS193085

330. Sarubbo S, Tate M, De Benedictis A, et al. Mapping critical cortical hubs and white matter pathways by direct electrical stimulation: an original functional atlas of the human brain. *Neuroimage*. 2020;205(May 2019):116237. doi:10.1016/j.neuroimage.2019.116237

331. Rasmussen T, Milner B. *Clinical and Surgical Studies of the Cerebral Speech Areas in Man*. (Zulch KJ, Creutzfelt O, Galbraith GC, eds.). Springer-Verlag; 1975. doi:https://doi.org/10.1007/978-3-642-66204-1_19

332. Duffau H, Capelle L, Denvil D, et al. The role of dominant premotor cortex in language: A study using intraoperative functional mapping in awake patients. *Neuroimage*. 2003;20(4):1903-1914. doi:10.1016/S1053-8119(03)00203-9

333. Quiñones-Hinojosa A, Ojemann SG, Sanai N, Dillon WP, Berger MS. Preoperative correlation of intraoperative cortical mapping with magnetic resonance imaging landmarks to predict localization of the Broca area. *J Neurosurg*. 2003;99(2):311-318. doi:10.3171/jns.2003.99.2.0311

334. Chang EF, Breshears JD, Raygor KP, Lau D, Molinaro AM, Berger MS. Stereotactic probability and variability of speech arrest and anomia sites during stimulation mapping of the language dominant hemisphere. *J Neurosurg*. 2017;126(1):114-121. doi:10.3171/2015.10.JNS151087

335. Zhou Y, Zhao Z, Zhang J, et al. Electrical stimulation-induced speech-related negative motor responses in the lateral frontal cortex. *J Neurosurg*. 2022;137(2):496-504. doi:10.3171/2021.9.JNS211069

336. Young JS, Lee AT, Chang EF. A Review of Cortical and Subcortical Stimulation Mapping for Language. *Neurosurgery*. 2021;89(3):331-342. doi:10.1093/neuros/nyaa436

337. Almairac F, Herbet G, Moritz-Gasser S, de Champfleur NM, Duffau H. The left inferior fronto-occipital fasciculus subserves language semantics: a multilevel lesion study. *Brain Struct Funct*. 2015;220(4):1983-1995. doi:10.1007/s00429-014-0773-1

338. Herbet G, Moritz-Gasser S, Duffau H. Electrical stimulation of the dorsolateral prefrontal cortex impairs semantic cognition. *Neurology*. 2018;90(12):e1077-e1084. doi:10.1212/WNL.0000000000005174

339. Corina DP, Mcburney SL, Dodrill C, Hinshaw K, Brinkley J, Ojemann G. *Functional Roles of Broca's Area and SMG: Evidence from Cortical Stimulation Mapping in a Deaf Signer*.; 1999. http://www.idealibrary.com

340. Yourganov G, Fridriksson J, Rorden C, Gleichgerrcht E, Bonilha L. Multivariate Connectome-Based Symptom Mapping in Post-Stroke Patients: Networks Supporting Language and Speech. *The Journal of Neuroscience*. 2016;36(25):6668-6679. doi:10.1523/jneurosci.4396-15.2016

341. Wilson SM, Lam D, Babiak MC, et al. Transient aphasias after left hemisphere resective surgery. *J Neurosurg*. 2015;123(3):581-593. doi:10.3171/2015.4.JNS141962

342. Schmidt CSM, Nitschke K, Bormann T, et al. Dissociating frontal and temporal correlates of phonological and semantic fluency in a large sample of left hemisphere stroke patients. *Neuroimage Clin*. 2019;23. doi:10.1016/j.nicl.2019.101840

343. Yourganov G, Smith KG, Fridriksson J, Rorden C. Predicting aphasia type from brain damage measured with structural MRI. *Cortex*. 2015;73:203-215. doi:10.1016/j.cortex.2015.09.005

344. Fridriksson J, Fillmore P, Guo D, Rorden C. Chronic Broca's Aphasia Is Caused by Damage to Broca's and Wernicke's Areas. *Cereb Cortex*. 2015;25(12):4689-4696. doi:10.1093/cercor/bhu152

345. Rolston JD, Englot DJ, Benet A, Li J, Cha S, Berger MS. Frontal operculum gliomas: language outcome following resection. *J Neurosurg*. Published online 2015:725-734. doi:10.3171/2014.11.JNS132172

346. Andrews JP, Cahn N, Speidel BA, et al. Dissociation of Broca's area from Broca's aphasia in patients undergoing neurosurgical resections. *J Neurosurg*. Published online August 5, 2022:1-11. doi:10.3171/2022.6.jns2297

347. Badre D, Wagner AD. Left ventrolateral prefrontal cortex and the cognitive control of memory. *Neuropsychologia*. 2007;45(13):2883-2901. doi:10.1016/j.neuropsychologia.2007.06.015

348. Liuzzi AG, Bruffaerts R, Peeters R, et al. Cross-modal representation of spoken and written word meaning in left pars triangularis. *Neuroimage*. 2017;150(December 2016):292-307. doi:10.1016/j.neuroimage.2017.02.032

349. Vitello S, Rodd JM. Resolving Semantic Ambiguities in Sentences: Cognitive Processes and Brain Mechanisms. *Lang Linguist Compass*. 2015;9(10):391-405. doi:10.1111/lnc3.12160

350. Matchin WG. A neuronal retuning hypothesis of sentence-specificity in Broca's area. *Psychon Bull Rev*. 2018;25(5):1682-1694. doi:10.3758/s13423-017-1377-6

351. Corrivetti F, de Schotten MT, Poisson I, et al. Dissociating motor–speech from lexico-semantic systems in the left frontal lobe: insight from a series of 17 awake intraoperative mappings in glioma patients. *Brain Struct Funct*. 2019;224(3):1151-1165. doi:10.1007/s00429-019-01827-7

352. Spena G, Nava A, Cassini F, et al. Preoperative and intraoperative brain mapping for the resection of eloquent-area tumors. A prospective analysis of methodology, correlation, and usefulness based on clinical outcomes. *Acta Neurochir (Wien)*. 2010;152(11):1835-1845. doi:10.1007/s00701-010-0764-9

353. Vidorreta JG, Garcia R, Moritz-Gasser S, Duffau H. Double dissociation between syntactic gender and picture naming processing: A brain stimulation mapping study. *Hum Brain Mapp*. 2011;32(3):331-340. doi:10.1002/hbm.21026

354. Alyahya RSW, Lambon Ralph MA, Halai A, Hoffman P. The cognitive and neural underpinnings of discourse coherence in post-stroke aphasia. *Brain Commun*. 2022;4(3). doi:10.1093/braincomms/fcac147

355. Davis C, Kleinman JT, Newhart M, Gingis L, Pawlak M, Hillis AE. Speech and language functions that require a functioning Broca's area. *Brain Lang*. 2008;105(1):50-58. doi:10.1016/j.bandl.2008.01.012

356. Adezati E, Thye M, Edmondson-Stait AJ, Szaflarski JP, Mirman D. Lesion correlates of auditory sentence comprehension deficits in post-stroke aphasia. *Neuroimage: Reports*. 2022;2(1):100076. doi:10.1016/j.ynirp.2021.100076

357. Marini A, Urgesi C. Please get to the point! A cortical correlate of linguistic informativeness. *J Cogn Neurosci*. 2012;24(11):2211-2222. doi:10.1162/jocn_a_00283

358. Dickens JV, Fama ME, DeMarco AT, Lacey EH, Friedman RB, Turkeltaub PE. Localization of phonological and semantic contributions to reading. *Journal of Neuroscience*. 2019;39(27):5361-5368. doi:10.1523/JNEUROSCI.2707-18.2019

359. den Ouden DB, Malyutina S, Basilakos A, et al. Cortical and structural-connectivity damage correlated with impaired syntactic processing in aphasia. *Hum Brain Mapp*. 2019;40(7):2153-2173. doi:10.1002/hbm.24514

360. Vigneau M, Beaucousin V, Hervé PY, et al. Meta-analyzing left hemisphere language areas: Phonology, semantics, and sentence processing. *Neuroimage*. 2006;30(4):1414-1432. doi:10.1016/j.neuroimage.2005.11.002

361. Sabb FW, Bilder RM, Chou M, Bookheimer SY. Working memory effects on semantic processing: priming differences in pars orbitalis. *Neuroimage*. 2007;37(1):311-322. doi:10.1016/j.neuroimage.2007.04.050

362. Nakai Y, Jeong JW, Brown EC, et al. Three- and four-dimensional mapping of speech and language in patients with epilepsy. *Brain*. 2017;140(5):1351-1370. doi:10.1093/brain/awx051

363. Binder JR, Desai RH. The neurobiology of semantic memory. *Trends Cogn Sci*. 2011;15(11):527-536. doi:10.1016/j.tics.2011.10.001

364. Noonan KA, Jefferies E, Visser M, Lambon Ralph MA. Going beyond inferior prefrontal involvement in semantic control: evidence for the additional contribution of dorsal angular gyrus and posterior middle temporal cortex. *J Cogn Neurosci*. 2013;25(11):1824-1850. doi:10.1162/jocn_a_00442

365. Sanai N, Mirzadeh Z, Berger MS. Functional Outcome after Language Mapping for Glioma Resection. *New England Journal of Medicine*. 2008;358(1):18-27. doi:10.1056/nejmoa067819

366. Duffau H, Gatignol P, Mandonnet E, Peruzzi P, Tzourio-Mazoyer N, Capelle L. New insights into the anatomo-functional connectivity of the semantic system: A study using cortico-subcortical electrostimulations. *Brain*. 2005;128(4):797-810. doi:10.1093/brain/awh423

367. Lubrano V, Draper L, Roux FE. What makes surgical tumor resection feasible in Broca's area? Insights into intraoperative brain mapping. *Neurosurgery*. 2010;66(5):868-875. doi:10.1227/01.NEU.0000368442.92290.04

368. Ivanova M V, Dragoy O, Kuptsova S V, et al. Neural mechanisms of two different verbal working memory tasks: A VLSM study. *Neuropsychologia*. 2018;115:25-41. doi:10.1016/j.neuropsychologia.2018.03.003

369. Henseler I, Regenbrecht F, Obrig H. Lesion correlates of patholinguistic profiles in chronic aphasia: Comparisons of syndrome-, modality-and symptom-level assessment. *Brain*. 2014;137(3):918-930. doi:10.1093/brain/awt374

370. Breshears JD, Southwell DG, Chang EF. Inhibition of Manual Movements at Speech Arrest Sites in the Posterior Inferior Frontal Lobe. *Neurosurgery*. 2019;85(3):E496-E501. doi:10.1093/neuros/nyy592

371. Tarapore PE, Findlay AM, Honma SM, et al. Language mapping with navigated repetitive TMS: Proof of technique and validation. *Neuroimage*. 2013;82:260-272. doi:10.1016/j.neuroimage.2013.05.018

372. Vassal F, Boutet C, Lemaire JJ, Nuti C. New insights into the functional significance of the frontal aslant tract: An anatomo-functional study using intraoperative electrical stimulations combined with diffusion tensor imaging-based fiber tracking. *Br J Neurosurg*. 2014;28(5):685-687. doi:10.3109/02688697.2014.889810

373. van Geemen K, Herbet G, Moritz-Gasser S, Duffau H. Limited plastic potential of the left ventral premotor cortex in speech articulation: Evidence From intraoperative awake mapping in glioma patients. *Hum Brain Mapp*. 2014;35(4):1587-1596. doi:10.1002/hbm.22275

374. Makuuchi M, Friederici AD. Hierarchical functional connectivity between the core language system and the working memory system. *Cortex*. 2013;49(9):2416-2423. doi:10.1016/j.cortex.2013.01.007

375. Brass M, Derrfuss J, Forstmann B, von Cramon DY. The role of the inferior frontal junction area in cognitive control. *Trends Cogn Sci*. 2005;9(7):314-316. doi:10.1016/j.tics.2005.05.001

376. Uzawa A, Hiraga A, Kamitsukasa I. Pure dysarthria resulting from a small cortical infarction located at the left middle frontal gyrus. *Internal Medicine*. 2009;48(1):75-76. doi:10.2169/internalmedicine.48.1540

377. Morshed RA, Lee AT, Wang EJ, et al. Functional outcomes after resection of middle frontal gyrus diffuse gliomas. *J Neurosurg*. Published online 2021:1-8. doi:10.3171/2021.8.JNS211624

378. Planton S, Jucla M, Roux FE, Démonet JF. The “handwriting brain”: A meta-analysis of neuroimaging studies of motor versus orthographic processes. *Cortex*. 2013;49(10):2772-2787. doi:10.1016/j.cortex.2013.05.011

379. Lubrano V, Roux FE, Demonet JF. Writing-specific sites in frontal areas: a cortical stimulation study. *J Neurosurg*. 2004;101(5):787-798. doi:10.3171/jns.2004.101.5.0787

380. Roux FE, Dufor O, Giussani C, et al. The graphemic/motor frontal area Exner’s area revisited. *Ann Neurol*. 2009;66(4):537-545. doi:10.1002/ana.21804

381. Abe K, Yokoyama R. A selective agraphia of Kana. *Behav Neurol*. 1994;7(2):79-81. doi:10.3233/BEN-1994-7205

382. Penfield W, Roberts L. *Speech and Brain Mechanisms*. Princeton University Press; 1959.

383. Piras F, Marangolo P. Word and number reading in the brain: Evidence from a Voxel-based Lesion-symptom Mapping study. *Neuropsychologia*. 2009;47(8-9):1944-1953. doi:10.1016/j.neuropsychologia.2009.03.006

384. Crosson B, Sadek JR, Maron L, et al. *Relative Shift in Activity from Medial to Lateral Frontal Cortex During Internally Versus Externally Guided Word Generation*.; 2001.

385. Ford A, McGregor KM, Case K, Crosson B, White KD. Structural connectivity of Broca's area and medial frontal cortex. *Neuroimage*. 2010;52(4):1230-1237. doi:10.1016/j.neuroimage.2010.05.018

386. Alario FX, Chainay H, Lehericy S, Cohen L. The role of the supplementary motor area (SMA) in word production. *Brain Res*. 2006;1076(1):129-143. doi:10.1016/j.brainres.2005.11.104

387. Crosson B, Benefield H, Cato MA, et al. Left and right basal ganglia and frontal activity during language generation: Contributions to lexical, semantic, and phonological processes. *Journal of the International Neuropsychological Society*. 2003;9(7):1061-1077. doi:10.1017/S135561770397010X

388. Penfield W, Welch K. The supplementary motor area of the cerebral cortex; a clinical and experimental study. *AMA Arch Neurol Psychiatry*. 1951;66(3):289-317. doi:10.1001/archneurpsyc.1951.02320090038004

389. Rostomily RC, Berger MS, Ojemann GA, Lettich E. Postoperative deficits and functional recovery following removal of tumors involving the dominant hemisphere supplementary motor area. *J Neurosurg*. 1991;75(1):62-68. doi:10.3171/jns.1991.75.1.0062

390. Rosenberg K, Nossek E, Liebling R, et al. Prediction of neurological deficits and recovery after surgery in the supplementary motor area: A prospective study in 26 patients - Clinical article. *J Neurosurg*. 2010;113(6):1152-1163. doi:10.3171/2010.6.JNS1090

391. Filevich E, Kühn S, Haggard P. Negative motor phenomena in cortical stimulation: Implications for inhibitory control of human action. *Cortex*. 2012;48(10):1251-1261. doi:10.1016/j.cortex.2012.04.014

392. Sierpowska J, Gabarrós A, Fernandez-Coello A, et al. Morphological derivation overflow as a result of disruption of the left frontal aslant white matter tract. *Brain Lang*. 2015;142:54-64. doi:10.1016/j.bandl.2015.01.005

393. Laplane D, Talairach J, Meininger V, Bancaud J, Orgogozo JM. Clinical consequences of corticectomies involving the supplementary motor area in man. *J Neurol Sci*. 1977;34(3):301-314. https://www.ncbi.nlm.nih.gov/pubmed/591992

394. Fontaine D, Capelle L, Duffau H. Somatotopy of the supplementary motor area: evidence from correlation of the extent of surgical resection with the clinical patterns of deficit. *Neurosurgery*. 2002;50(2):297-303; discussion 303-5. doi:10.1097/00006123-200202000-00011

395. Kim YH, Kim CH, Kim JS, et al. Risk factor analysis of the development of new neurological deficits following supplementary motor area resection: Clinical article. *J Neurosurg*. 2013;119(1):7-14. doi:10.3171/2013.3.JNS121492

396. Vassal M, Charroud C, Deverdun J, et al. Recovery of functional connectivity of the sensorimotor network after surgery for diffuse low-grade gliomas involving the supplementary motor area. *J Neurosurg*. 2016;126(4):1181-1190. doi:10.3171/2016.4.jns152484

397. Chivukula S, Pikul BK, Black KL, Pouratian N, Bookheimer SY. Contralateral functional reorganization of the speech supplementary motor area following neurosurgical tumor resection. *Brain Lang*. 2018;183(April):41-46. doi:10.1016/j.bandl.2018.05.006

398. Pinson H, Van Lerbeirghe J, Vanhauwaert D, Van Damme O, Hallaert G, Kalala JP. The supplementary motor area syndrome: a neurosurgical review. *Neurosurg Rev*. 2022;45(1):81-90. doi:10.1007/s10143-021-01566-6

399. Baldo J V., Wilkins DP, Ogar J, Willock S, Dronkers NF. Role of the precentral gyrus of the insula in complex articulation. *Cortex*. 2011;47(7):800-807. doi:10.1016/j.cortex.2010.07.001

400. Oh A, Duerden EG, Pang EW. The role of the insula in speech and language processing. *Brain Lang*. 2014;135:96-103. doi:10.1016/j.bandl.2014.06.003

401. Tang YY, Rothbart MK, Posner MI. Neural correlates of establishing, maintaining, and switching brain states. *Trends Cogn Sci*. 2012;16(6):330-337. doi:10.1016/j.tics.2012.05.001

402. Zaccarella E, Friederici AD. Reflections of word processing in the insular cortex: A sub-regional parcellation based functional assessment. *Brain Lang*. 2015;142:1-7. doi:10.1016/j.bandl.2014.12.006

403. Stefaniak JD, Alyahya RSW, Lambon Ralph MA. Language networks in aphasia and health: A 1000 participant activation likelihood estimation meta-analysis. *Neuroimage*. 2021;233. doi:10.1016/j.neuroimage.2021.117960

404. Eickhoff SB, Heim S, Zilles K, Amunts K. A systems perspective on the effective connectivity of overt speech production. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*. 2009;367(1896):2399-2421. doi:10.1098/rsta.2008.0287

405. Duffau H, Moritz-Gasser S, Gatignol P. Functional outcome after language mapping for insular World Health Organization Grade II gliomas in the dominant hemisphere: experience with 24 patients. *Neurosurg Focus*. 2009;27(2):E7. doi:10.3171/2009.5.focus0938

406. Afif A, Minotti L, Kahane P, Hoffmann D. Anatomofunctional organization of the insular cortex: A study using intracerebral electrical stimulation in epileptic patients. *Epilepsia*. 2010;51(11):2305-2315. doi:10.1111/j.1528-1167.2010.02755.x

407. Ackermann H, Riecker A. The contribution(s) of the insula to speech production: a review of the clinical and functional imaging literature. *Brain Struct Funct*. 2010;214(5-6):419-433. doi:10.1007/s00429-010-0257-x

408. Zarino B, Sirtori MA, Meschini T, et al. Insular lobe surgery and cognitive impairment in gliomas operated with intraoperative neurophysiological monitoring. *Acta Neurochir (Wien)*. 2021;163(5):1279-1289. doi:10.1007/s00701-020-04643-9

409. Kerezoudis P, Singh R, Goyal A, et al. Insular epilepsy surgery: lessons learned from institutional review and patient-level meta-analysis. *J Neurosurg*. 2022;136(2):523-535. doi:10.3171/2021.1.JNS203104

410. Przybylowski CJ, Baranoski JF, So VM, Wilson J, Sanai N. Surgical morbidity of transsylvian versus transcortical approaches to insular gliomas. *J Neurosurg*. 2020;132(6):1731-1738. doi:10.3171/2018.12.JNS183075

411. Hervey-Jumper SL, Li J, Osorio JA, et al. Surgical assessment of the insula. Part 2: validation of the Berger-Sanai zone classification system for predicting extent of glioma resection. *J Neurosurg*. 2016;124(2):482-488. doi:10.3171/2015.4.JNS1521

412. Di Stefano V, De Angelis M V, Montemitro C, et al. Clinical presentation of strokes confined to the insula: a systematic review of literature. *Neurol Sci*. 2021;42(5):1697-1704. doi:10.1007/s10072-021-05109-1

413. Duffau H, Taillandier L, Gatignol P, Capelle L. The insular lobe and brain plasticity: Lessons from tumor surgery. *Clin Neurol Neurosurg*. 2006;108(6):543-548. doi:10.1016/j.clineuro.2005.09.004

414. Sanai N, Polley MY, Berger MS. Insular glioma resection: assessment of patient morbidity, survival, and tumor progression. *J Neurosurg*. 2010;112(1):1-9. doi:10.3171/2009.6.JNS0952

415. Ogar J, Willock S, Baldo J, Wilkins D, Ludy C, Dronkers N. Clinical and anatomical correlates of apraxia of speech. *Brain Lang*. 2006;97(3). doi:10.1016/j.bandl.2006.01.008

416. Bates E, Wilson SM, Saygin AP, et al. Voxel-based lesion-symptom mapping. *Nat Neurosci*. 2003;6(5):448-450. doi:10.1038/nn1050

417. Na Y, Jung JY, Tench CR, Auer DP, Pyun SB. Language systems from lesion-symptom mapping in aphasia: A meta-analysis of voxel-based lesion mapping studies. *Neuroimage Clin*. 2022;35. doi:10.1016/j.nicl.2022.103038

418. Kümmerer D, Hartwigsen G, Kellmeyer P, et al. Damage to ventral and dorsal language pathways in acute aphasia. *Brain*. 2013;136(2):619-629. doi:10.1093/brain/aws354

419. Skrap M, Mondani M, Tomasino B, et al. Surgery of insular nonenhancing gliomas: Volumetric analysis of tumoral resection, clinical outcome, and survival in a consecutive series of 66 cases. *Neurosurgery*. 2012;70(5):1081-1093. doi:10.1227/NEU.0b013e31823f5be5

420. Vinci-Booher S, Cheng H, James KH. An Analysis of the Brain Systems Involved with Producing Letters by Hand. *J Cogn Neurosci*. 2019;31(1):138-154. doi:10.1162/jocn_a_01340

421. Nee DE. Integrative frontal-parietal dynamics supporting cognitive control. *Elife*. 2021;10. doi:10.7554/eLife.57244

422. Mountcastle VB. The parietal system and some higher brain functions. *Cereb Cortex*. 1995;5(5):377-390. doi:10.1093/cercor/5.5.377

423. Coslett HB, Schwartz MF. The parietal lobe and language. *Handb Clin Neurol*. 2018;151:365-375. doi:10.1016/B978-0-444-63622-5.00018-8

424. Bruffaerts R, De Deyne S, Meersmans K, Liuzzi AG, Storms G, Vandenberghe R. Redefining the resolution of semantic knowledge in the brain: Advances made by the introduction of models of semantics in neuroimaging. *Neurosci Biobehav Rev*. 2019;103:3-13. doi:10.1016/j.neubiorev.2019.05.015

425. Lesourd M, Servant M, Baumard J, et al. Semantic and action tool knowledge in the brain: Identifying common and distinct networks. *Neuropsychologia*. 2021;159. doi:10.1016/j.neuropsychologia.2021.107918

426. Magrassi L, Bongetta D, Bianchini S, Berardesca M, Arienta C. Central and peripheral components of writing critically depend on a defined area of the dominant superior parietal gyrus. *Brain Res*. 2010;1346:145-154. doi:10.1016/j.brainres.2010.05.046

427. Hartwigsen G, Baumgaertner A, Price CJ, Koehnke M, Ulmer S, Siebner HR. Phonological decisions require both the left and right supramarginal gyri. *Proceedings of the National Academy of Sciences*. 2010;107(38):16494-16499. doi:10.1073/pnas.1008121107

428. Sliwinska MW, Khadilkar M, Campbell-Ratcliffe J, Quevenco F, Devlin JT. Early and sustained supramarginal gyrus contributions to phonological processing. *Front Psychol*. 2012;3(MAY). doi:10.3389/fpsyg.2012.00161

429. Oberhuber M, Hope TMH, Seghier ML, et al. Four Functionally Distinct Regions in the Left Supramarginal Gyrus Support Word Processing. *Cerebral Cortex*. 2016;26(11):4212-4226. doi:10.1093/cercor/bhw251

430. Menon V, Desmond JE. Left superior parietal cortex involvement in writing: integrating fMRI with lesion evidence. *Brain Res Cogn Brain Res*. 2001;12(2):337-340. doi:10.1016/s0926-6410(01)00063-5

431. Paulesu E, Frith CD, Frackowiak RS. The neural correlates of the verbal component of working memory. *Nature*. 1993;362(6418):342-345. doi:10.1038/362342a0

432. Jobard G, Crivello F, Tzourio-Mazoyer N. Evaluation of the dual route theory of reading: A metanalysis of 35 neuroimaging studies. *Neuroimage*. 2003;20(2):693-712. doi:10.1016/S1053-8119(03)00343-4

433. Yue Q, Martin RC. Maintaining verbal short-term memory representations in non-perceptual parietal regions. *Cortex*. 2021;138:72-89. doi:10.1016/j.cortex.2021.01.020

434. Hashimoto Y, Sakai KL. Brain activations during conscious self-monitoring of speech production with delayed auditory feedback: An fMRI study. *Hum Brain Mapp*. 2003;20(1):22-28. doi:10.1002/hbm.10119

435. Golfinopoulos E, Tourville JA, Bohland JW, Ghosh SS, Nieto-Castanon A, Guenther FH. FMRI investigation of unexpected somatosensory feedback perturbation during speech. *Neuroimage*. 2011;55(3):1324-1338. doi:10.1016/j.neuroimage.2010.12.065

436. Kandylaki KD, Nagels A, Tune S, et al. Predicting “When” in Discourse Engages the Human Dorsal Auditory Stream: An fMRI Study Using Naturalistic Stories. *Journal of Neuroscience*. 2016;36(48):12180-12191. doi:10.1523/jneurosci.4100-15.2016

437. Lukic S, Thompson CK, Barbieri E, et al. Common and distinct neural substrates of sentence production and comprehension. *Neuroimage*. 2021;224. doi:10.1016/j.neuroimage.2020.117374

438. Hamberger MJ, Seidel WT. Auditory and visual naming tests: normative and patient data for accuracy, response time, and tip-of-the-tongue. *J Int Neuropsychol Soc*. 2003;9(3):479-489. https://www.ncbi.nlm.nih.gov/pubmed/12666772

439. Maldonado IL, Moritz-Gasser S, De Champfleur NM, Bertram L, Moulinié G, Duffau H. Surgery for gliomas involving the left inferior parietal lobule: New insights into the functional anatomy provided by stimulation mapping in awake patients: Clinical article. *J Neurosurg*. 2011;115(4):770-779. doi:10.3171/2011.5.JNS112

440. Roux FE, Lubrano V, Lauwers-Cances V, Trémoulet M, Mascott CR, Démonet JF. Intra-operative mapping of cortical areas involved in reading in mono- and bilingual patients. *Brain*. 2004;127(8):1796-1810. doi:10.1093/brain/awh204

441. Quigg M, Geldmacher DS, Elias WJ. Conduction aphasia as a function of the dominant posterior perisylvian cortex. Report of two cases. *J Neurosurg*. 2006;104(5):845-848. doi:10.3171/jns.2006.104.5.845

442. Roux FE, Durand JB, Jucla M, Réhault E, Reddy M, Démonet JF. Segregation of Lexical and Sub-Lexical Reading Processes in the Left Perisylvian Cortex. *PLoS One*. 2012;7(11). doi:10.1371/journal.pone.0050665

443. Papagno C, Comi A, Riva M, et al. Mapping the brain network of the phonological loop. *Hum Brain Mapp*. 2017;38(6):3011-3024. doi:10.1002/hbm.23569

444. Sanai N, Martino J, Berger MS. Morbidity profile following aggressive resection of parietal lobe gliomas: Clinical article. *J Neurosurg*. 2012;116(6):1182-1186. doi:10.3171/2012.2.JNS111228

445. Stark BC, Basilakos A, Hickok G, Rorden C, Bonilha L, Fridriksson J. Neural organization of speech production: A lesion-based study of error patterns in connected speech. *Cortex*. 2019;117:228-246. doi:10.1016/j.cortex.2019.02.029

446. Pisoni A, Mattavelli G, Casarotti A, et al. The neural correlates of auditory-verbal short-term memory: a voxel-based lesion-symptom mapping study on 103 patients after glioma removal. *Brain Struct Funct*. 2019;224(6):2199-2211. doi:10.1007/s00429-019-01902-z

447. Sakurai Y, Asami M, Mannen T. Alexia and agraphia with lesions of the angular and supramarginal gyri: Evidence for the disruption of sequential processing. *J Neurol Sci*. 2010;288(1-2):25-33. doi:10.1016/j.jns.2009.10.015

448. Ghaleh M, Lacey EH, Fama ME, Anbari Z, Demarco AT, Turkeltaub PE. Dissociable Mechanisms of Verbal Working Memory Revealed through Multivariate Lesion Mapping. *Cerebral Cortex*. 2020;30(4):2542-2554. doi:10.1093/cercor/bhz259

449. Scarone P, Gatignol P, Guillaume S, Denvil D, Capelle L, Duffau H. Agraphia after awake surgery for brain tumor: new insights into the anatomo-functional network of writing. *Surg Neurol*. 2009;72(3):223-241; discussion 241. doi:10.1016/j.surneu.2008.10.074

450. Baldo J V., Kacinik N, Ludy C, et al. Voxel-based lesion analysis of brain regions underlying reading and writing. *Neuropsychologia*. 2018;115:51-59. doi:10.1016/j.neuropsychologia.2018.03.021

451. Thye M, Mirman D. Relative contributions of lesion location and lesion size to predictions of varied language deficits in post-stroke aphasia. *Neuroimage Clin*. 2018;20:1129-1138. doi:10.1016/j.nicl.2018.10.017

452. Lau EF, Phillips C, Poeppel D. A cortical network for semantics: (de)constructing the N400. *Nat Rev Neurosci*. 2008;9(12):920-933. doi:10.1038/nrn2532

453. Seghier ML. The angular gyrus: Multiple functions and multiple subdivisions. *Neuroscientist*. 2013;19(1):43-61. doi:10.1177/1073858412440596

454. Humphreys GF, Lambon Ralph MA, Simons JS. A Unifying Account of Angular Gyrus Contributions to Episodic and Semantic Cognition. *Trends Neurosci*. 2021;44(6):452-463. doi:10.1016/j.tins.2021.01.006

455. Humphries C, Binder JR, Medler DA, Liebenthal E. Time course of semantic processes during sentence comprehension: An fMRI study. *Neuroimage*. 2007;36(3):924-932. doi:10.1016/j.neuroimage.2007.03.059

456. Seghier ML, Fagan E, Price CJ. Functional subdivisions in the left angular gyrus where the semantic system meets and diverges from the default network. *Journal of Neuroscience*. 2010;30(50):16809-16817. doi:10.1523/JNEUROSCI.3377-10.2010

457. Bonner MF, Peelle JE, Cook PA, Grossman M. Heteromodal conceptual processing in the angular gyrus. *Neuroimage*. 2013;71:175-186. doi:10.1016/j.neuroimage.2013.01.006

458. Vandenberghe R, Wang Y, Nelissen N, et al. The associative-semantic network for words and pictures: Effective connectivity and graph analysis. *Brain Lang*. 2013;127(2):264-272. doi:10.1016/j.bandl.2012.09.005

459. Graves WW, Binder JR, Desai RH, Humphries C, Stengel BC, Seidenberg MS. Anatomy is strategy: Skilled reading differences associated with structural connectivity differences in the reading network. *Brain Lang*. 2014;133:1-13. doi:10.1016/j.bandl.2014.03.005

460. Palliera C, Devauchellea AD, Dehaenea S. Cortical representation of the constituent structure of sentences. *Proc Natl Acad Sci U S A*. 2011;108(6):2522-2527. doi:10.1073/pnas.1018711108

461. Meyler A, Keller TA, Cherkassky VL, Gabrieli JDE, Just MA. Modifying the brain activation of poor readers during sentence comprehension with extended remedial instruction: A longitudinal study of neuroplasticity. *Neuropsychologia*. 2008;46(10):2580-2592. doi:10.1016/j.neuropsychologia.2008.03.012

462. Boylan C, Trueswell JC, Thompson-Schill SL. Relational vs. attributive interpretation of nominal compounds differentially engages angular gyrus and anterior temporal lobe. *Brain Lang*. 2017;169:8-21. doi:10.1016/j.bandl.2017.01.008

463. Matchin W, Brodbeck C, Hammerly C, Lau E. The temporal dynamics of structure and content in sentence comprehension: Evidence from fMRI-constrained MEG. *Hum Brain Mapp*. 2019;40(2):663-678. doi:10.1002/hbm.24403

464. Thye M, Geller J, Szaflarski JP, Mirman D. Intracranial EEG evidence of functional specialization for taxonomic and thematic relations. *Cortex*. 2021;140:40-50. doi:10.1016/j.cortex.2021.03.018

465. Maldonado IL, De Champfleur NM, Velut S, Destrieux C, Zemmoura I, Duffau H. Evidence of a middle longitudinal fasciculus in the human brain from fiber dissection. *J Anat*. 2013;223(1):38-45. doi:10.1111/joa.12055

466. Pustina D, Avants B, Faseyitan OK, Medaglia JD, Coslett HB. Improved accuracy of lesion to symptom mapping with multivariate sparse canonical correlations. *Neuropsychologia*. 2018;115:154-166. doi:10.1016/j.neuropsychologia.2017.08.027

467. Greenblatt SH. Neurosurgery and the anatomy of reading: a practical review. *Neurosurgery*. 1977;1(1):6-15. doi:10.1227/00006123-197707000-00003

468. Maeda K, Yamamoto Y, Shiraishi T, Fujioka E. Dystypia Associated with Diaschisis of the Middle Frontal Gyri after Left Angular Infarction. *J Stroke Cerebrovasc Dis*. 2021;30(7):105803. doi:10.1016/j.jstrokecerebrovasdis.2021.105803

469. Purcell J, Rapp B, Martin RC. Distinct Neural Substrates Support Phonological and Orthographic Working Memory: Implications for Theories of Working Memory. *Front Neurol*. 2021;12. doi:10.3389/fneur.2021.681141

470. Sul B, Lee KB, Hong BY, et al. Association of lesion location with long-term recovery in post-stroke aphasia and language deficits. *Front Neurol*. 2019;10(JUL). doi:10.3389/fneur.2019.00776

471. Wise RJ, Scott SK, Blank SC, Mummery CJ, Murphy K, Warburton EA. Separate neural subsystems within “Wernicke’s area.” *Brain*. 2001;124(Pt 1):83-95. https://www.ncbi.nlm.nih.gov/pubmed/11133789

472. Chang EF, Edwards E, Nagarajan SS, et al. Cortical spatio-temporal dynamics underlying phonological target detection in humans. *J Cogn Neurosci*. 2011;23(6):1437-1446. doi:10.1162/jocn.2010.21466

473. Turkeltaub PE, Branch Coslett H. Localization of sublexical speech perception components. *Brain Lang*. 2010;114(1):1-15. doi:10.1016/j.bandl.2010.03.008

474. Leff AP, Schofield TM, Crinion JT, et al. The left superior temporal gyrus is a shared substrate for auditory short-term memory and speech comprehension: Evidence from 210 patients with stroke. *Brain*. 2009;132(12):3401-3410. doi:10.1093/brain/awp273

475. Anderson JM, Gilmore R, Roper S, et al. Conduction aphasia and the arcuate fasciculus: A reexamination of the Wernicke-Geschwind model. *Brain Lang*. 1999;70(1):1-12. doi:10.1006/brln.1999.2135

476. Roux FE, Miskin K, Durand JB, et al. Electrostimulation mapping of comprehension of auditory and visual words. *Cortex*. 2015;71:398-408. doi:10.1016/j.cortex.2015.07.001

477. Rogalsky C, LaCroix AN, Chen KH, et al. The neurobiology of agrammatic sentence comprehension: A lesion study. *J Cogn Neurosci*. 2018;30(2):234-255. doi:10.1162/jocn_a_01200

478. Caplan D, Michaud J, Hufford R, Makris N. Deficit-lesion correlations in syntactic comprehension in aphasia. *Brain Lang*. 2016;152:14-27. doi:10.1016/j.bandl.2015.10.005

479. Mesgarani N, Chang EF. Selective cortical representation of attended speaker in multi-talker speech perception. *Nature*. 2012;485(7397):233-236. doi:10.1038/nature11020

480. Scott SK, Blank CC, Rosen S, Wise RJ. Identification of a pathway for intelligible speech in the left temporal lobe. *Brain*. 2000;123 Pt 12:2400-2406. https://www.ncbi.nlm.nih.gov/pubmed/11099443

481. Mesgarani N, Cheung C, Johnson K, Chang EF. Phonetic feature encoding in human superior temporal gyrus. *Science (1979)*. 2014;343(6174):1006-1010. doi:10.1126/science.1245994

482. Haglund MM, Berger MS, Shamseldin M, Lettich E, Ojemann GA. Cortical localization of temporal lobe language sites in patients with gliomas. *Neurosurgery*. 1994;34(4):567-576. doi:10.1227/00006123-199404000-00001

483. Boatman D. Cortical bases of speech perception: evidence from functional lesion studies. *Cognition*. 2004;92(1-2):47-65. doi:10.1016/j.cognition.2003.09.010

484. Bello L, Gallucci M, Fava M, et al. Intraoperative subcortical language tract mapping guides surgical removal of gliomas involving speech areas. *Neurosurgery*. 2007;60(1):67-80. doi:10.1227/01.NEU.0000249206.58601.DE

485. Walenski M, Europa E, Caplan D, Thompson CK. Neural networks for sentence comprehension and production: An ALE-based meta-analysis of neuroimaging studies. *Hum Brain Mapp*. 2019;40(8):2275-2304. doi:10.1002/hbm.24523

486. Whitney C, Kirk M, O'Sullivan J, Lambon Ralph MA, Jefferies E. Executive semantic processing is underpinned by a large-scale neural network: revealing the contribution of left prefrontal, posterior temporal, and parietal cortex to controlled retrieval and selection using TMS. *J Cogn Neurosci*. 2012;24(1):133-147. doi:10.1162/jocn_a_00123

487. Sierpowska J, Gabarrós A, Fernández-Coello A, et al. White-matter pathways and semantic processing: intrasurgical and lesion-symptom mapping evidence. *Neuroimage Clin*. 2019;22. doi:10.1016/j.nicl.2019.101704

488. Sanai N, Berger MS. Glioma extent of resection and its impact on patient outcome. *Neurosurgery*. 2008;62(4):753-764; discussion 264-6. doi:10.1227/01.neu.0000318159.21731.cf

489. Maldonado IL, Moritz-Gasser S, Duffau H. Does the left superior longitudinal fascicle subserve language semantics? A brain electrostimulation study. *Brain Struct Funct*. 2011;216(3):263-274. doi:10.1007/s00429-011-0309-x

490. Kristinsson S, Thors H, Yourganov G, et al. Brain Damage Associated with Impaired Sentence Processing in Acute Aphasia. *J Cogn Neurosci*. 2020;32(2):256-271. doi:10.1162/jocn_a_01478

491. Turken AU, Dronkers NF. The Neural Architecture of the Language Comprehension Network: Converging Evidence from Lesion and Connectivity Analyses. *Frontiers in System Neuroscience*. 2011;5(February):1-20. doi:10.3389/fnsys.2011.00001

492. Baldo J V., Arévalo A, Patterson JP, Dronkers NF. Grey and white matter correlates of picture naming: Evidence from a voxel-based lesion analysis of the Boston Naming Test. *Cortex*. 2013;49(3):658-667. doi:10.1016/j.cortex.2012.03.001

493. Ilmberger J, Eisner W, Schmid U, Reulen HJ. Performance in picture naming and word comprehension: Evidence for common neuronal substrates from intraoperative language mapping. *Brain Lang*. 2001;76(2):111-118. doi:10.1006/brln.2000.2415

494. Pillay SB, Binder JR, Humphries C, Gross WL, Book DS. Lesion localization of speech comprehension deficits in chronic aphasia. *Neurology*. 2017;88:97-975.

495. Dien J, Brian ES, Molfese DL, Gold BT. Combined ERP/fMRI evidence for early word recognition effects in the posterior inferior temporal gyrus. *Cortex*. 2013;49(9):2307-2321. doi:10.1016/j.cortex.2013.03.008

496. Bonilha L, Hillis AE, Hickok G, den Ouden DB, Rorden C, Fridriksson J. Temporal lobe networks supporting the comprehension of spoken words. *Brain*. 2017;140(9):2370-2380. doi:10.1093/brain/awx169

497. Pouratian N, Bookheimer SY, Rubino G, Martin NA, Toga AW. Category-specific naming deficit identified by intraoperative stimulation mapping and postoperative neuropsychological testing. Case report. *J Neurosurg*. 2003;99(1):170-176. doi:10.3171/jns.2003.99.1.0170

498. Bedos Ulvin L, Jonas J, Brissart H, et al. Intracerebral stimulation of left and right ventral temporal cortex during object naming. *Brain Lang*. 2017;175:71-76. doi:10.1016/j.bandl.2017.09.003

499. Abdallah C, Brissart H, Colnat-Coulbois S, et al. Stereoelectroencephalographic language mapping of the basal temporal cortex predicts postoperative naming outcome. *J Neurosurg*. Published online 2021:1-11. doi:10.3171/2020.8.JNS202431

500. Chan-Seng E, Moritz-Gasser S, Duffau H. Awake mapping for low-grade gliomas involving the left sagittal stratum: Anatomofunctional and surgical considerations: Clinical article. *J Neurosurg*. 2014;120(5):1069-1077. doi:10.3171/2014.1.JNS132015

501. Tamai S, Kinoshita M, Nakajima R, Okita H, Nakada M. Two different subcortical language networks supporting distinct Japanese orthographies: morphograms and phonograms. *Brain Struct Funct*. 2022;227(3):1145-1154. doi:10.1007/s00429-022-02454-5

502. Ardila A, Bernal B, Rosselli M. How Localized are Language Brain Areas? A Review of Brodmann Areas Involvement in Oral Language. *Arch Clin Neuropsychol*. 2016;31(1):112-122. doi:10.1093/arclin/acv081

503. Zemmoura I, Herbet G, Moritz-Gasser S, Duffau H. New insights into the neural network mediating reading processes provided by cortico-subcortical electrical mapping. *Hum Brain Mapp*. 2015;36(6):2215-2230. doi:10.1002/hbm.22766

504. Pylkkänen L. *The Neural Basis of Combinatory Syntax and Semantics*. https://www.science.org

505. Wong C, Gallate J. The function of the anterior temporal lobe: A review of the empirical evidence. *Brain Res*. 2012;1449:94-116. doi:10.1016/j.brainres.2012.02.017

506. Hartwigsen G, Henseler I, Stockert A, et al. Integration demands modulate effective connectivity in a fronto-temporal network for contextual sentence integration. *Neuroimage*. 2017;147:812-824. doi:10.1016/j.neuroimage.2016.08.026

507. Cervenka MC, Corines J, Boatman-Reich DF, et al. Electrocorticographic functional mapping identifies human cortex critical for auditory and visual naming. *Neuroimage*. 2013;69:267-276. doi:10.1016/j.neuroimage.2012.12.037

508. Matsumoto R, Imamura H, Inouchi M, et al. Left anterior temporal cortex actively engages in speech perception: A direct cortical stimulation study. *Neuropsychologia*. 2011;49(5):1350-1354. doi:10.1016/j.neuropsychologia.2011.01.023

509. Hermann BP, Perrine K, Chelune GJ, et al. Visual confrontation naming following left anterior temporal lobectomy: a comparison of surgical approaches. *Neuropsychology*. 1999;13(1):3-9. doi:10.1037//0894-4105.13.1.3

510. Ives-Deliperi VL, Butler JT. Naming outcomes of anterior temporal lobectomy in epilepsy patients: a systematic review of the literature. *Epilepsy Behav*. 2012;24(2):194-198. doi:10.1016/j.yebeh.2012.04.115

511. Tsapkini K, Frangakis CE, Hillis AE. The function of the left anterior temporal pole: Evidence from acute stroke and infarct volume. *Brain*. 2011;134(10):3094-3105. doi:10.1093/brain/awr050

512. Tsapkini K, Rapp B. The orthography-specific functions of the left fusiform gyrus: Evidence of modality and category specificity. *Cortex*. 2010;46(2):185-205. doi:10.1016/j.cortex.2009.02.025

513. Turker S, Hartwigsen G. The use of noninvasive brain stimulation techniques to improve reading difficulties in dyslexia: A systematic review. *Hum Brain Mapp*. 2022;43(3):1157-1173. doi:10.1002/hbm.25700

514. Dehaene S, Cohen L. The unique role of the visual word form area in reading. *Trends Cogn Sci*. 2011;15(6):254-262. doi:10.1016/j.tics.2011.04.003

515. Seghier ML, Price CJ. Dissociating frontal regions that co-lateralize with different ventral occipitotemporal regions during word processing. *Brain Lang*. 2013;126(2):133-140. doi:10.1016/j.bandl.2013.04.003

516. Woolnough O, Donos C, Rollo PS, et al. Spatiotemporal dynamics of orthographic and lexical processing in the ventral visual pathway. *Nat Hum Behav*. 2021;5(3):389-398. doi:10.1038/s41562-020-00982-w

517. Yeatman JD, White AL. Reading: The Confluence of Vision and Language. *Annu Rev Vis Sci*. 2021;7:487-517. doi:10.1146/annurev-vision-093019-113509

518. Ludersdorfer P, Wimmer H, Richlan F, Schurz M, Hutzler F, Kronbichler M. Left ventral occipitotemporal activation during orthographic and semantic processing of auditory words. *Neuroimage*. 2016;124(Pt A):834-842. doi:10.1016/j.neuroimage.2015.09.039

519. Price CJ, Devlin JT. The interactive account of ventral occipitotemporal contributions to reading. *Trends Cogn Sci*. 2011;15(6):246-253. doi:10.1016/j.tics.2011.04.001

520. Mattioni S, Rezk M, Battal C, et al. Categorical representation from sound and sight in the ventral occipito-temporal cortex of sighted and blind. *Elife*. 2020;9. doi:10.7554/eLife.50732

521. Gaillard R, Naccache L, Pinel P, et al. Direct intracranial, FMRI, and lesion evidence for the causal role of left inferotemporal cortex in reading. *Neuron*. 2006;50(2):191-204. doi:10.1016/j.neuron.2006.03.031

522. Mani J, Diehl B, Piao Z, et al. Evidence for a basal temporal visual language center: cortical stimulation producing pure alexia. *Neurology*. 2008;71(20):1621-1627. doi:10.1212/01.wnl.0000334755.32850.f0

523. Gil-Robles S, Carvallo A, Jimenez M del M, et al. Double Dissociation Between Visual Recognition and Picture Naming. *Neurosurgery*. 2013;72(4):678-686. doi:10.1227/NEU.0b013e318282a361

524. Hirshorn EA, Li Y, Ward MJ, Richardson RM, Fiez JA, Ghuman AS. Decoding and disrupting left midfusiform gyrus activity during word reading. *Proc Natl Acad Sci U S A*. 2016;113(29):8162-8167. doi:10.1073/pnas.1604126113

525. Woolnough O, Snyder KM, Morse CW, McCarty MJ, Lhatoo SD, Tandon N. Intraoperative localization and preservation of reading in ventral occipitotemporal cortex. *J Neurosurg*. Published online 2022:1-8. doi:10.3171/2022.2.JNS22170

526. Mandonnet E, Gatignol P, Duffau H. Evidence for an occipito-temporal tract underlying visual recognition in picture naming. *Clin Neurol Neurosurg*. 2009;111(7):601-605. doi:10.1016/j.clineuro.2009.03.007

527. Hamberger MJ. Cortical language mapping in epilepsy: A critical review. *Neuropsychol Rev*. 2007;17(4):477-489. doi:10.1007/s11065-007-9046-6

528. Aron O, Jonas J, Colnat-Coulbois S, Maillard L. Language Mapping Using Stereo Electroencephalography: A Review and Expert Opinion. *Front Hum Neurosci*. 2021;15. doi:10.3389/fnhum.2021.619521

529. Luders H, Lesser RP, Hahn J, et al. Basal temporal language area. *Brain*. 1991;114 (Pt 2):743-754. doi:10.1093/brain/114.2.743

530. Schaffler L, Luders HO, Morris 3rd HH, Wyllie E. Anatomic distribution of cortical language sites in the basal temporal language area in patients with left temporal lobe epilepsy. *Epilepsia*. 1994;35(3):525-528. doi:10.1111/j.1528-1157.1994.tb02472.x

531. Usui K, Ikeda A, Takayama M, et al. Processing of Japanese morphogram and syllabogram in the left basal temporal area: electrical cortical stimulation studies. *Brain Res Cogn Brain Res*. 2005;24(2):274-283. doi:10.1016/j.cogbrainres.2005.02.001

532. Mikuni N, Miyamoto S, Ikeda A, et al. Subtemporal hippocampectomy preserving the basal temporal language area for intractable mesial temporal lobe epilepsy: preliminary results. *Epilepsia*. 2006;47(8):1347-1353. doi:10.1111/j.1528-1167.2006.00610.x

533. Shimotake A, Matsumoto R, Ueno T, et al. Direct exploration of the role of the ventral anterior temporal lobe in semantic memory: Cortical stimulation and local field potential evidence from subdural grid electrodes. *Cerebral Cortex*. 2015;25(10):3802-3817. doi:10.1093/cercor/bhu262

534. Forseth KJ, Kadipasaoglu CM, Conner CR, Hickok G, Knight RT, Tandon N. A lexical semantic hub for heteromodal naming in middle fusiform gyrus. *Brain*. 2018;141(7):2112-2126. doi:10.1093/brain/awy120

535. Reindl C, Allgauer AL, Kleiser BA, et al. Resection of dominant fusiform gyrus is associated with decline of naming function when temporal lobe epilepsy manifests

after the age of five: A voxel-based lesion-symptom mapping study. *Neuroimage Clin*. 2022;35:103129. doi:10.1016/j.nicl.2022.103129

536. Sanjuán A, Hope TMH, Parker Jones ’Oiwi, et al. Dissociating the semantic function of two neighbouring subregions in the left lateral anterior temporal lobe. *Neuropsychologia*. 2015;76:153-162. doi:10.1016/j.neuropsychologia.2014.12.004

537. Weber B, Wellmer J, Schür S, et al. Presurgical language fMRI in patients with drug-resistant epilepsy: Effects of task performance. *Epilepsia*. 2006;47(5):880-886. doi:10.1111/j.1528-1167.2006.00515.x

538. Bennett MR, Hacker PMS. *Philosophical Foundations of Neuroscience*. Blackwell Publishing; 2003.

539. Kahn E, Lane M, Sagher O. Eloquent: history of a word’s adoption into the neurosurgical lexicon. *J Neurosurg*. 2017;127(6):1461-1466. doi:10.3171/2017.3.JNS17659