Aus der

Radiologischen Universitätsklinik Tübingen

Abteilung Diagnostische und Interventionelle Neuroradiologie

Investigation to Improve the Display of Language Areas using Functional Magnetic Resonance Imaging

Inaugural-Dissertation
zur Erlangung des Doktorgrades
der Medizin

der Medizinischen Fakultät
der Eberhard Karls Universität
zu Tübingen

vorgelegt von

Fetscher, Leonard Johannes

2023

Dekan:	Professor Dr. B. Pichler
1. Berichterstatter:	Professor Dr. U. Klose
2. Berichterstatter:	Professor Dr. T. Ethofer
Tag der Disputation:	14.08.2023

Table of Contents

Abbreviations

ASFN	American Society of Functional Neuroradiology
BOLD	Blood Oxygenation Level Dependent
C	Carbon
CSF	Cerebrospinal fluid
ECS	Electro Cortical Stimulation
EEG	Electroencephalogram
e.g.	Exempli gratia (for example)
EPI	Echo planar imaging
ESM	Electro Stimulation Mapping
F	Female
FDG	Fluorodeoxyglucose
fMRI	Functional Magnetic Resonance Imaging
FoV	Field of view
GLM	General-Linear-Model
HRF	Hemodynamic response function
L	Left
LA	Language area
M	Male
mm	Millimeter
MRI	Magnetic Resonance Imaging
N	Natural number
NT	Naming-task

PET	Positron Emission Tomography
PET-CT	Positron Emission Tomography Computer Tomography
Pre-SMA	supplementary speech area
qEEG	Quantitative Electroencephalography
ROI	Region of Interest
s	Second
SC	Sentence-completion
SWG	Silent-word-generation
TC	Time course
TE	Echo time
TR´	Repetition time
VG	Verb-generation

1 Introduction

1.1 Mapping the Brain

The first and probably most famous attempt to map the cerebrum was published as early as 1909 in “Vergleichende Lokalisationslehre der Großhirnrinde. In ihren Principien dargestellt auf Grund des Zellenbaues.“ By the German neuroanatomist Korbinian Brodmann. The classification of the 52 Brodmann areas is based on differences in cyto- and myelo-architectonics and is still being used today, with only minor changes.

Nowadays, this original classification has been expanded with the help of new techniques and devices, functions have been assigned to several brain areas, and the courses of fiber strands have been identified. The methods used today are divided into two groups. One group of techniques is based on cellular and molecular mechanisms in neurons and glial cells, the other group of techniques investigates the function and functional regions of the brain (Eliassen et al. 2008; Garrett et al. 2012).

The focus of this study lay on the improvement of the preoperative mapping of the language areas (LAs) relevant in clinical practice (Broca’s and Wernicke’s area) using functional Magnetic Resonance Imaging (fMRI). In the following, the advantages of fMRI were shown through an overview and comparison of different mapping techniques, emphasizing the importance of fMRI as a mapping technique in everyday clinical practice.

1.1.1 Invasive brain mapping techniques

The surgical ablative procedures should be mentioned as they represent the basic principle of these techniques. The functions of a region are investigated by destroying the brain tissue in the region of interest and observing the changes in behavior or (cognitive/motoric) abilities of the subject, using the negative proof. This procedure is irreversible and therefore actively only used in animal studies. To gather information about the human brain, changes in the cognitive abilities and the behavior of patients with brain damage in a restricted and defined area caused e.g., by injuries or by medical interventions, such as brain resection (due to e.g., brain tumors or epilepsy), can be investigated and the lost functions can be assigned to the damaged brain regions.

The other group of invasive mapping techniques are the stimulation techniques. There are two different approaches to stimulation methods. In one method, electric current is applied to a brain region, in the other method, anesthetics are applied to a brain region. Both methods temporarily activate or inhibit certain brain areas.

A commonly known method to determine the lateralization of the LAs is the WADA testing (Wada und Rasmussen 2007). A catheter is inserted into the dexter or sinister internal carotid artery of the awake patient. Sodium amytal is injected into the artery to temporarily disable the respective brain hemisphere. Language tasks are performed by the patient after the contralateral hemisphere is disabled, to determine the language abilities of the active hemisphere. If the language function is unaffected, the lateralization of the language center is assumed to be in the active hemisphere.

Another more invasive technique is the Electro Stimulation Mapping (ESM), or the Electro Cortical Stimulation (ECS). ESM is one of the first reversible mapping techniques. Its basic principle goes as far back as 1807, when Luigi Rolando stimulated the cortices of living animals with galvanic current (Caputi et al). Nowadays, ESM is performed by inserting uni- or bipolar electrodes into the brain tissue in regions of interest to either record action potentials or to stimulate the region with electrical current. The recording of action potentials is mainly used extra operatively on epileptic patients to record the action potentials of the affected regions during seizures. In this way, affected regions can be precisely localized and separated from unaffected regions. For this study, the intraoperative use of this method is of greater importance. Stimuli in form of electric current are applied to regions of the brain to receive either positive or negative responses. The stimulation of functional brain areas (e.g., motor cortex), leads to a positive response. The stimulation evokes an action potential innervating the targeted area (e.g., a muscle). The response (e.g., movement of muscle) is then recorded and a function (the evoked response) can be assigned to the corresponding brain region. This examination can be performed under general anesthesia.

To identify regions with higher cognitive functions, such as the LAs, negative responses come into play. In this process, brain areas are stimulated intraoperatively on the awake patient to temporarily disable them. The patient fulfils a language task (e.g., speaking) while different locations of the brain are disabled by being over-stimulated with electric current. A certain function can be assigned to a region, based on the location of the electrode at the time of the loss of function. This method is invasive but very precise and therefore remains the gold standard for functional mapping of the brain (Ritaccio et al. 2018).

1.1.2 Non-invasive brain mapping techniques

Non-invasive mapping techniques are widely used in clinical practice, even tough, these methods have a lower accuracy and significance compared to the invasive methods. The important prerequisites for their widespread use in clinical practice are the easy access and the strongly reduced risk for the patient. An electrophysiological approach is the qEEG, which can assign certain brain/thought processes to certain regions by a dynamic measurement of the change in the EEG signal (Swatzyna et al. 2015). Another option originating from nuclear medicine is the Positron Emission Tomography (PET). Here, the patient is injected with positron-emitting radiopharmaceuticals (Tracer). To measure brain activity, fluorodeoxyglucose (^{18}FDG) and 15-oxygen labeled water is widely used, because glucose and oxygen are the main substances that are required for the function of the brain (Raichle et al. 1983; Fowler et al. 2007). However, FDG is not perfectly suited as a radiotracer, as it achieves only limited specificity due to unspecific high cortical uptake of the tracer. Other radiotracers consisting of amino acids (e.g., ^{11}C methionine) show a higher specificity (D'Souza et al. 2011). However, due to its lack of temporal and spatial resolution, PET-CT is rarely used for imaging brain functions.

Finally, MRI-based imaging methods are to be mentioned, especially the fMRI used in this study. This widely accessible and easy to perform examination provides a satisfactory spatial and temporal resolution, a very low complication rate and very few contra indications, making it a good technique for every day clinical practice.

1.1.2.1 Mapping with MRI

The most commonly used technique for the mapping of brain functions with MRI is fMRI, a contrast technique using the Blood Oxygenation Level Dependent (BOLD)-effect (Ogawa und Lee 1990; Ogawa et al. 1990; Ogawa 2012). Fulfilling cognitive tasks leads to dilatation of arterioles and arteries in the activated regions (known as Hemodynamic response function (HRF)) causing an increased blood flow in the activated regions. Thereby, the amount of diamagnetic oxyhemoglobin increases, affecting the T2- and T2*-

weighted imaging. Through the detection of changes in the regional blood flow with fMRI, brain activity can be displayed and localized.

A typical fMRI examination in block design consists of two blocks, an activation, and a rest blocks. During the activation block, the subject/patient is required to perform tasks (e.g., language or motoric). During the rest block, the subject either lays still or performs an alternate task, not stimulating the region of interest. To see, which brain regions are activated by the tasks, the regional blood flow of activation block and rest block is compared, using statistical tools (e.g., a GLM (General-Linear-Model)). The design of the tasks depends on the region of interest. The more complex the task, the more precise is the location of the region (Perthen et al. 2008; Chen und Glover 2015).

1.2 Language fMRI

Language fMRI is widely used in clinical practice for preoperative diagnostics and planning, often to localize the LAs before neurosurgical interventions (Benjamin et al. 2017; Kundu et al. 2013; Petrella et al. 2006). It offers a low-risk and radiation-free imaging method for the creation of activation maps with a good spatial resolution of activated brain areas. Areas in the brain that are relevant for language are activated using language tasks that are presented to the patient/subject during the fMRI examination.

To avoid damaging the LAs, often resulting in aphasia, during a neurosurgical intervention, a precise preoperative localization of the LA is of great importance. Brain matter and especially functional areas such as the LAs are very sensitive, so that even minor injuries to these areas can cause serious deficits (Kundu et al. 2013; Sanai et al. 2008; Lemée et al. 2019; Benjamin et al. 2017).

However, it is also important to precisely limit the extent of the LAs. In many interventions, such as tumor resection, it is important to remove as much (tumor) tissue as possible. Inaccurate spatial limitation of the LAs can result in too little tissue being removed, which can lead to recurrence of the tumor and can also interfere with radiochemical therapy resulting in a reduced outcome (Hong et al. 2013; Duffau 2016; Lacroix et al. 2001).

Good preoperative planning has a direct impact on the success of the therapy/surgery and language fMRI is a good and frequently used method for this purpose. Nevertheless, the localizations of the LAs on the activation maps generated with language fMRI are in some cases too imprecise to make a valid statement for the preoperative planning, especially in the case of lesions in the immediate surroundings of the LAs. The precision of the fMRI maps is, beside the parameters of the MRI device itself, affected by the setup of the examination and the preprocessing of the data, therefore the right choice of language tasks, measurement settings and analysis of the data can lead to an improved localization of the LAs (Black et al. 2017; Fetscher et al. 2022).

1.2.1 Language Tasks

Depending on the setting and objective, various language tasks are being used for fMRI examinations. Therefore, a variety of recommendations regarding the creation and execution of language tasks exist in the literature. For instance, neuroscience studies that focus on the preoperative localization of the LAs of epilepsy patients with epilepsy that often have little to no cognitive impairment aim to obtain the most precise results possible by using complex to perform and cognitively demanding language tasks (Benjamin et al. 2017; Gaillard et al. 2002; Gaillard et al. 2004). The workload and the complexity of these language tasks is not necessarily useful in clinical practice and/or when working with elderly or cognitively impaired (e.g., due to a brain tumor) patients.

The "American Society of Functional Neuroradiology" (ASFN) has published criteria for creating language tasks, based on interviews with various physicians and neuroscientists that are working in the field of language fMRI. The relevant criteria were an appropriate task difficulty, the balance between sensitivity and specificity, and the representation of inter-hemispheric lateralization and intra-hemispheric localization. The recommended language task for activations in the frontal lobe, the location of Broca's area, was a silent-word-generation task (SWG), for the temporal lobe, where Wernicke's area is located, the language task of choice was a sentence-completion-task (SC) (Black et al. 2017). The correct selection of tasks, or task groups, is fundamental for a full display of LAs (Black et al. 2017; Gaillard et al. 2004; Połczyńska et al. 2017). The display of language activation can further be improved by using two different tasks activating the same brain region. Auditory tasks can be used in addition to visual, as the second or third task group

(Gaillard et al. 2004). Furthermore, tasks requiring higher grammatical skills can lead to increased activation signals in the LAs compared to pure word generation tasks (Połczyńska et al. 2017).

Until now, no gold standard approach to clinical language fMRI exists; neither in terms of language tasks, the examination itself, or the evaluation of it has been described (Benjamin et al. 2017). Nevertheless, efforts have been made on several occasions and the clinical approaches tend to assimilate. For instance, the successful replication of an fMRI examination across different imaging centers has been demonstrated (Binder et al. 2011) and in some cases multicenter imaging consortia using same fMRI protocol have been established (Evans 2006; Friedman und Glover 2006; Mueller et al. 2005). Statistically, the three most used tasks in a clinical setting are noun-prompted verb generation tasks (e.g., SWG), verbal fluency tasks, and semantic decision tasks (e.g., SC) (Black et al. 2017). Also 95% of clinical trials were conducted with two or more language tasks (Benjamin et al. 2018b).

The grammatical and vocabulary content of the tasks in language fMRI examinations of adults should not exceed the level of the 7^{th} school grade (Black et al. 2017). To form sentences and select words, the vocabulary used in children's books may be used to estimate the vocabular and grammatical level of the language tasks (Benjamin et al. 2017; Black et al. 2017).

In this study, the language tasks noun-prompted word generation, responsive naming, and sentence completion were used. These language tasks were compared with each other to determine their suitability in clinical practice, as these were often described in connection with the clinical use-case. Each task was performed in two ways, using visual stimulus presentation, and using auditory stimulus presentation. With this experiment setup, each region is activated by different tasks, improving the quality of the results (Gaillard et al. 2002; Gaillard et al. 2004). The goal was, to identify the language task that combines the accuracy of neuroscientific examinations with the simplicity and robustness required in clinical examinations.

Another focus lay on making this examination accessible to as many patients as possible. The ongoing demographic transition is leading to an aging population in modernized countries (Coale 1984). As this is leading to an increasing prevalence of age-related symptoms the modern medicine and clinical diagnostics must adapt. The sensory system is often affected, manifesting in visual impairment or hearing loss (Tang et al. 2015; Maberley et al. 2006; Pascolini und Mariotti 2012; Goman und Lin 2016; Herbst und Humphrey 1981). A well-functioning sensory system is important for the reception and processing of the language tasks; therefore, task dependent fMRI examinations are affected strongly by these impairments. To solve this problem, the acoustic stimulus presentation was compared to the commonly used visual stimulus presentation, to evaluate if the acoustic stimulus is a valid alternative for the presentation of language tasks. Being able to adjust the presentation of the stimulus to respond to the individual condition of the patient gives more people the opportunity to undergo a language fMRI examination.

1.2.2 Analysis of language fMRI

The display of activations in a task-based block-design language fMRI is based on differences in the HRF between the activation block and the rest block (Figure 1).

The fMRI data can be analyzed in different ways. In the clinical practice, evaluation programs (e.g., syngo.MR Neuro fMRI (Siemens Healthcare GmbH, Germany)) included in the evaluation software included in the MRI-scanner are used frequently. These programs are usually based on a GLM and additionally include various data processing steps, such as smoothing, linear correction, realignment, normalization, and other preprocessing tools. This way solid and evaluable activation maps are created, that give the examiner an idea of the location of the LAs. When processing the data, however, it must be borne in mind that the changes made to the raw data lead to an altered representation of the actual measured location of activations, which can lead to incorrect localization of the LAs. Especially the smoothing has a severe impact on the localization and limitation of the spatial extend of the LAs, as it alters the true location and extend of measured activations and thus makes it difficult for the examiner to make valid predictions.

The examiner's task is to decide whether the displayed activations are artifacts or true activations. If they are true activations, the examiner must determine which functional area these activations belong to. Once the region of interest has been identified, in this case the LAs, the spatial extent of this region must be estimated to make a relevant statement for pre-operative planning. However, the activations shown on the activation map only represent probable locations of the functional areas (Benjamin et al. 2018b). The displayed spatial extent of the activations is variable, because changes in the threshold and the use of smoothing can drastically change the appearance of activations on the map.

The approach used in this study attempts to achieve a more reliable and constant representation of activations to facilitate and improve the spatial estimation of activations. Simple t-maps based on unsmoothed and unprocessed data were created. When creating a t-map, the HRF-signal of a voxel in the activation block is compared with the HRF-signal of the pause block. If a significant difference, which is defined by the t-limit, is detected, the voxel is marked as an activation and displayed on the t-map. During this process, however, a lot of relevant information that is included in the HRF is lost. To compensate for this, the time courses (TC) of individual voxels, were analyzed using parameters to include temporal information and to improve the separation of true activations from artifacts and experiment independent activations.

The task-based language fMRI method used in this study allowed an estimation of the specific shape of the HRF of in voxels that are located in the brain areas activated by the language tasks (Figure 1). By knowing about the likely behavior of a HRF and the BOLD-Effect (Chen et al. 2015), it is possible to estimate parameters such as time to peak, height and width of the activation (Lindquist und Wager 2007; Lindquist et al. 2009) in advance. This knowledge enables the identification and exclusion of activations that occur independently of the experiment. Also, there are also region-specific HRF differences (Lindquist et al. 2009; Lindquist und Wager 2007; Aguirre et al. 1998; Bielczyk et al. 2017) that can be used to highlight LAs among all regions stimulated by the experiment (e.g., visual or auditory cortex).

Figure 1 shows the TC of an activation typical for our experiment, low signal during the rest block and increased signal during the activation block. The TC shows the signal of one rest block (second 0-24) and one activation block (second 24-48). The signal is shown as deviation of the mean in percent. The black bar at 24s indicates the start of the activation block. It is visible that at the beginning of the rest block the signal is still elevated (due to the activation of the previous activation block), but it quickly decreases and forms a plateau with low signal strength. From 24s, the beginning of the activation block, the signal starts to increase until it forms a plateau at an elevated signal level at about 30s (6s after the beginning of the activation block, as it is to be expected due to the BOLD effect). During this study, the specificizes included in TCs of the LAs were analyzed precisely leading to the definition of parameters and their limits. The knowledge about LA specific TCs was then used to differentiate between the different activations and to only display the activations of the LAs on the activation maps.

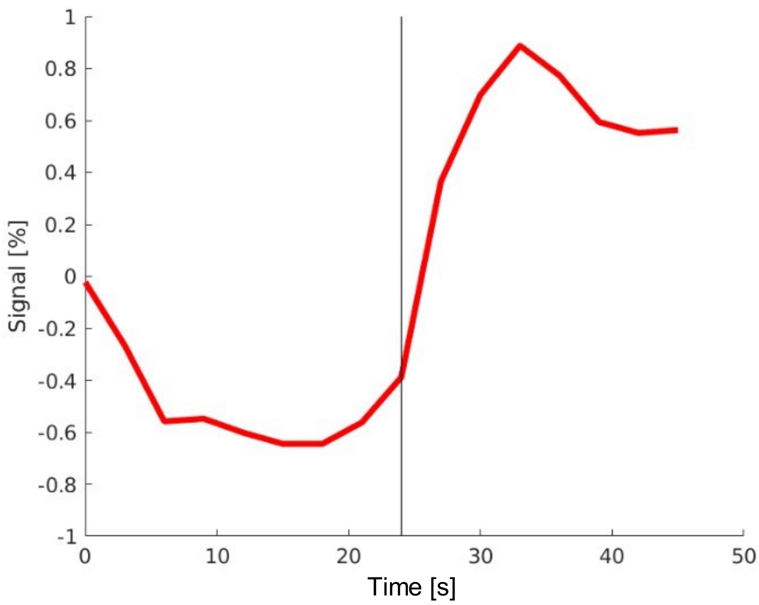


Figure 1: Typical TC of the HRF of a language activation stimulated by a language task.
Red line: Signal of the HRF (0-24s rest block, 24-48s activation block). It is shown as deviation of the mean in percent.
The black bar at 24s indicates the start of the activation block.

1.3 Language Areas

In clinical routine of planning neurosurgical procedures, only Broca's and Wernicke's areas are usually considered as language critical areas (Figure 2). However, these two regions do not represent all functional regions involved in language production. Benjamin et al. 2017 mentioned the following six language-critical regions:

Broca's area, also known as the motoric center of language, was discovered in the 19th century by the French neurosurgeon Paul Broca. He described two patients with impaired language motor function but intact language comprehension. During the autopsy of the patients, Broca found lesions in the inferior frontal gyrus (Brodmann areas 44, 45), today better known as Broca's area (Dronkers et al. 2007). From Broca's area, efferent tracks project to the basal ganglia and then on to the premotor cortex for activation of language associated muscles.

Wernicke's area, also known as the sensory language center, was discovered by the German neurologist Carl Wernicke in 1874. It is located in the posterior segment of the superior gyrus (Brodmann area 22) and encompasses the auditory cortex on the lateral sulcus (Brodmann areas 41/42). Wernicke’s area is mainly responsible for the comprehension of language in spoken and written form (Javed et al. 2021). Efferent fibers of Wernicke's area running through the arcuate fascicle form a connection to Broca's area, a disconnection between these two areas (e.g. due to a lesion of the arcuate fascicle) can lead to a conduction aphasia (Catani und Mesulam 2008).

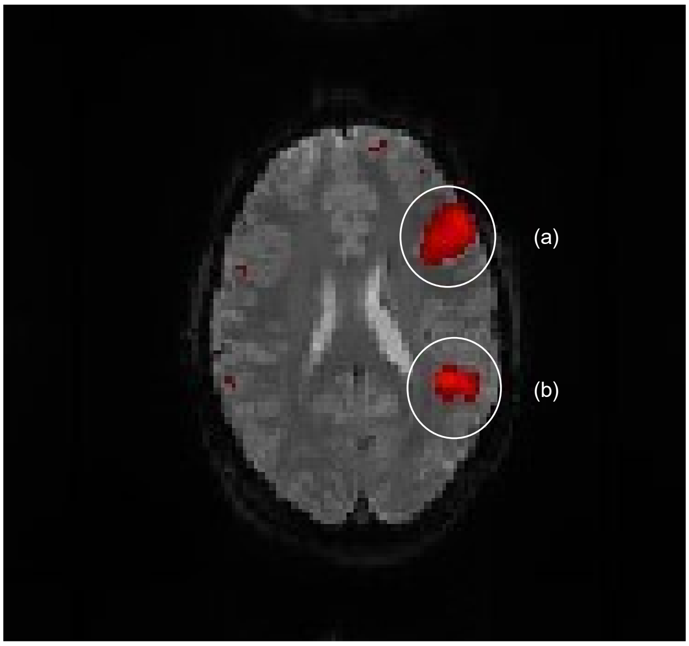


Figure 2: Axial slice of a functional MRI dataset showing the location of the LAs. Red color indicates activations; (a) Broca's area; (b) Wernicke's area.

The remaining four areas are not as fundamental as Broca's or Wernicke's area but are nevertheless important for a perfectly functioning language function. Exner's area is involved in the phonology of words and their translation into writing. It is located in the posterior part of the middle frontal gyrus (Roux et al. 2010). The supplementary speech area (pre-SMA) controls the motor movements needed for speech and is located in the posterior superior and middle frontal cortex (Hiroshima et al. 2014). The angular gyrus forms the connection between Broca's area and Wernicke's area and is involved in reading and in the transfer between read and spoken language (Binder et al. 2009). Lastly, a language-relevant area in the basal temporal region has been described. The loss or stimulation of this region led to various impairments in language production and comprehension (Burnstine et al. 1990; Krauss et al. 1996; Schäffler et al. 1994).

The inclusion of additional regions in basic preoperative planning in clinical practice exceeds the professional and technical resources available today and the preservation of these regions contributes only minor to the patient's outcome. The display and examination of these areas is only performed in special cases, mainly in a research related setting. Since this thesis was conducted in a clinical setting, the focus lay on the display of Broca's and Wernicke's area.

1.4 Injury of language areas

Damage to the LAs is mainly caused by strokes, traumatic brain injury, tumors or brain surgery and can result in aphasia, which is defined as „[…] the loss or impairment of language function caused by brain damage.” (Benson und Ardila 1996).

Motor (Broca's) and sensory (Wernicke's) aphasia must be distinguished. In Broca's aphasia, language comprehension and object naming are preserved, but language production is severely limited. In Wernicke's aphasia, language production is preserved even though the spoken words are incoherent and meaningless in content. In addition, language comprehension and object naming are severely limited. Lesions in both areas result in global aphasia with a failure of both language comprehension and language production (Klinke et al. 2005).

What the preservation of language means for the life and outcome of the patient is impressively shown by studies on stroke patients. Depression (Kauhanen et al. 2000), reduced functional outcome (Paolucci et al. 1998), reduced quality of life and overall increased mortality (Laska et al. 2001; Laures-Gore et al. 2020) are associated with aphasia. Accordingly, it is of great importance to prevent iatrogenic caused aphasia, especially if it is avoidable through preoperative planning.

The incidence of language deficits following surgical intervention near or in the LAs is not consistently reported in the literature, Sawaya 1999 describes of a risk of up to 26%, whereas (Brownsett et al. 2019) summarized various publications with a averaged risk of 5 - 11%. This risk can be reduced by preoperative planning and mapping (Garrett et al. 2012; Ritaccio et al. 2018).

1.5 Aim of the study

In this work, different aspects of the language fMRI examination used in nowadays clinical practice were reviewed to find possibilities to improve the examination in terms of clinical usability, inclusion of elderly patients, realistic spatial representation, and identification of language activations.

The examination setup was adjusted to improve the clinical suitability, including the robustness of the examination, to reduce the number of failed measurements and the accessibility to as many patients as possible, especially elderly patients with age-related impairments (e.g., loss of hearing). This was to be achieved by creating suitable language tasks and adapting the examination settings to the individual conditions of the patients.

The improve the analysis of the data and the identification and precise localization of the LAs themselves an alternative approach was used. By using unprocessed fMRI data and the simple technique of t-map, the real location of the measured activations should be preserved. To make these activation maps more readable and to highlight LAs, the activation maps were filtered by HRF-dependent parameters.

In summary, the aim of this study was to find ways of improving and facilizing preoperative planning of neurosurgical interventions that pose a risk of injury to the LAs and thereby mitigate the number of iatrogenic caused aphasia.

2 Materials and methods

2.1 Data and study population

Two different sets of data were used in this study, one set contained the data of patients, the other set contained the data of healthy subjects. All 30 patients that were enrolled in the dataset of patients, were being treated in the Radiology Clinic, Department of interventional and diagnostic neuroradiology, University Clinic of Tuebingen and had to undergo a language fMRI examination due to clinical indications. All patients gave written consent for the use of their MRI data for scientific purposes. The data was provided in anonymous form, without any information about the patient.

The data of the healthy subjects were obtained by a prospective study. 20 healthy subjects were enrolled in this study. Each underwent three to six language fMRI scans with different experiment settings and language tasks. The duration of the examination was about 60 minutes per subject, depending on the number of scans performed. The total duration of the study was one year. The study was approved by the Ethics Committee of the University of Tuebingen and written consent was given by the subjects about participating in this study and the use of the fMRI data. The inclusion criteria for this study contained the usual criteria of MRI examinations. Also, written consent had been given by the subjects to be informed in the event of an incidental pathological finding. None of the subjects participated in other experimental studies at the time of enrollment. The demographics of the subjects are provided in Table 1.

Table 1: Subject demographics

Subject	Age (years)	Gender	Language Lateralisation
1	25	M	L
2	18	F	L
3	55	M	L
4	50	F	L
5	25	M	L
6	25	M	L
7	22	F	L
8	24	M	L
9	26	M	L
10	25	M	L
11	26	M	L
12	24	M	L
13	25	F	L
14	23	M	L
15	59	F	L
16	25	F	L
17	24	M	L
18	29	M	L
19	23	M	L
20	25	M	L
Mean	28,9	7F/14M	20 L

Abbreviations: F, female; M, male; L, left.

2.2 MRI-Sequence

All scans were performed with Two Siemens 3T whole-body scanner (model: MAGNETOM Skyra and MAGNETOM Prisma, Siemens, Germany) were used to perform the scans. An echo planar imaging sequence (EPI; T2*-weighted, 104 volumes, slice thickness 3.0mm, matrix 96, FoV 245mm x 245mm, gap 0.75mm, TR 3s, TE 36ms, 104 volumes) was used for the functional imaging.

2.3 Experiment design and language tasks

Both patient and subject data in this study were generated using a task-dependent fMRI examination in a block design. The examination consisted of two blocks with a duration of 24s each, a rest block, and an activation block. The blocks were presented alternately starting and ending on a rest block. In total, one examination consisted of 6 activation blocks and 7 rest blocks.

To address the question of how the TCs change over the course of the experiment, some scan performed with more blocks than the standard experiment design. In total 4 scans with 12 periods and 8 scans with 8 periods were implemented. One period is defined as the combination of one activation block and one rest block.

The rest block consisted of an activated pause. The patients/subjects were required to perform an action, namely the opening and closing of the fists. The signal for this action was given either visually or acoustically, matching the stimulus used for the language task, by the word "fist".

The tasks to be performed by the patients/subjects during the activation block were language tasks. Each of the language tasks were designed in a way that they could be presented visually or acoustically. In total, 3 different language tasks were implemented:

Verb-generation (VG): Nouns were presented to which verbs must be formed mentally without speaking. A new word was presented every 3s, there were 8 words per block.

Naming-task (NT): A short word description (3-5 words) was presented. A matching word must be found and formed mentally without speaking. A new word description was presented every 4s, there were 6-word descriptions per block.

Sentence-completion (SC): A short gap text (5-10 words) was presented. The missing word must be found and formed mentally without speaking. A new sentence was presented every 6s, there were 4 sentences per block.

For the subject study, all the language tasks were used, both visually and acoustically, 6 different scans in total; the scans performed with the patients only consisted of the verb-generation task in visual presentation.

The visual tasks were presented to the patients/subjects in written form on a monitor that could be seen by the patient/subject via a mirror attached to the head coil. The acoustic tasks were recorded in spoken form and presented to the patient via headphones (MRI suitable headphones manufactured by the company MSA (MSA - Magdeburger Sonderkonstruktions- und Apparatebau GmbH, Magdeburg, Germany)). All language tasks were created using Presentation® software (version 18.0, Neurobehavioral Systems, Inc., Berkeley, CA, www.neurobs.com).

Patients/subjects were positioned with their heads in a head coil during the examination. The head was held in place with foam cushions to minimize movement artifacts.

2.4 Subject feedback

All subjects participated in a short interview after the scan. The answers were documented by the experiment leader on a prepared questionnaire. The interview included questions on the topics: explanation and practice of the language task before the scan; comprehensibility, quality, and execution of the language tasks; concentration during the scan.

2.5 Creation of activation maps

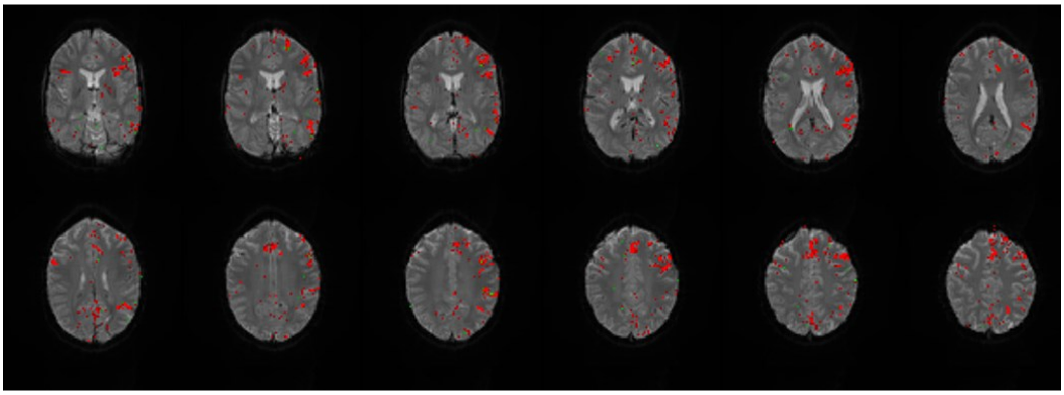
Two different techniques were used to create activation maps (Figure 3). In both cases, only the fMRI data of the EPI-Sequences were used, therefore no co-registration was needed. Therefore, the alteration of the data that could lead to spatial inaccuracies of the activations was avoided, but the anatomical masks of the maps were of limited quality and consisting of only a limited number of slices.

One technique used a GLM. The "Inline BOLD Imaging" included in the syngo.MR software of SIEMENS Healthineers, with "Inline calculation of t-statistics (t-maps) based on a general linear model (GLM) including the hemodynamic response function and correcting for slow drift" (Siemens Medical Solutions USA) was used to calculate the fMRI maps and to define activations. This approach is commonly used in clinical practice and included several preprocessing steps: motion correction, a spatial filter with a filter width of 5.0mm ("smoothing"), a temporal highpass filter and the modeling of transition states. The paradigm (language task) size was 16, but the first time point (TP) of each block (TP 1 and 9) was ignored. These maps (later referred to as "GLM-map") were used to evaluate the performance of the different language tasks and experiment settings implemented during the scans of the subjects. Also, the GLM-maps were compared with the later explained "filter t-maps" to evaluate the quality and performance of this mapping approach.

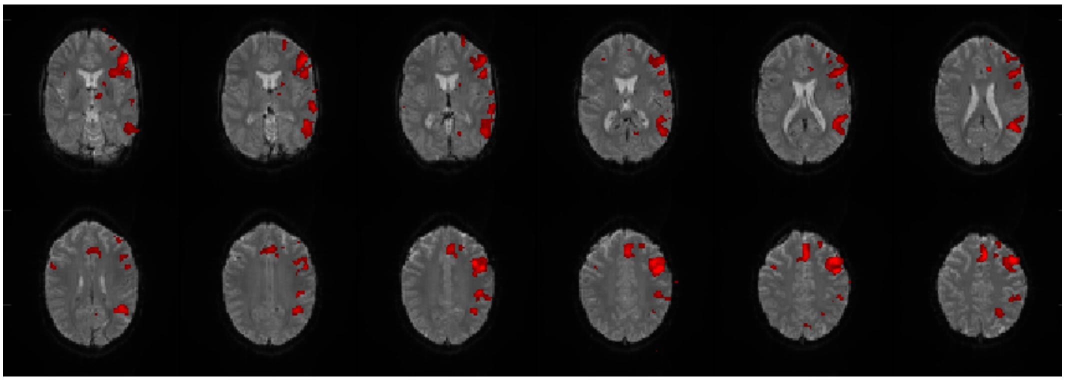
The other mapping approach used in this study was based on a simple t-test and the implementation of a parameter-based filter. The t-maps were created with a Matlab (MATLAB 12.0, The MathWorks, Inc., Natick, Massachusetts, United States) based computer program. Activations on the t-map were defined by comparing the HRF-signal of the activation block to the HRF-signal of the rest block. In both blocks, the 1^{st} TP (1TP = 3s) was excluded. If the HRF differences of the two blocks of a particular voxel exceeded a significant t-value, the voxel was defined as activation and plotted onto the t-map. Usually, this t-value is adjusted individually for each fMRI data set under review. In this study, however, the same t-value ($t = 2.2$) was used for all fMRI data (both patients and subjects).

To define this t-limit, 48 fMRI data sets from the subject study were used compare the number of excluded activations in the LAs with the number of excluded activations of the entire t-map at different t-limits. The number of excluded activations in each region was calculated in dependence of the t-limit and displayed as a histogram. Based on the histogram, the t-limit that offered the best compromise between preserving language activations and deleting artifacts and unwanted activations, was chosen.

The t-maps were created using only the raw fMRI data, without using any preprocessing tools. To improve the appearance of the map without changing the location of activations, a cluster limit was to be integrated, excluding all activations consisting of less than 3 contiguous voxels.



(a)



(b)

Figure 3: Comparison of the two mapping techniques t-map and GLM-map. Measured activations are displayed red on the activation maps.
(a) t-map
(b) GLM-map.

2.6 Regions of interest

Groups of voxels of different regions were marked and saved as Regions of Interests (ROI), to study the region-specific characteristics of the TCs across different scans and patients. A total of seven regions were selected: Broca's area, Wernicke's area (together later referred to as LAs; Figure 1), motor cortex, auditory cortex and visual cortex (all three together referred to as functional regions, Figure 4). Additionally, two apparent activations that were most likely independent of the language tasks and language production were included in the group of ROIs (Figure 4). These activations appeared in most scans, making them well suited as exemplary artifacts for TC analysis. The "ventricular artifact" is located at the transition between brain tissue and the ventricular system and is caused by minimal head movements. Signal changes appear in this region, due to the differences in the magnetic susceptibility of brain tissue and cerebrospinal fluid (CSF). In some cases, these signal changes are detected in the TCs, showing a similar signal pattern as truly activated region, which leads to the display of this region on the activation maps. The other interesting apparent activation is the "frontal artifact". Activations that are independent of the experiment and language production often occur in the frontal regions of the brain. These activations may be caused by subconscious thought processes, large vessels, or other structures.

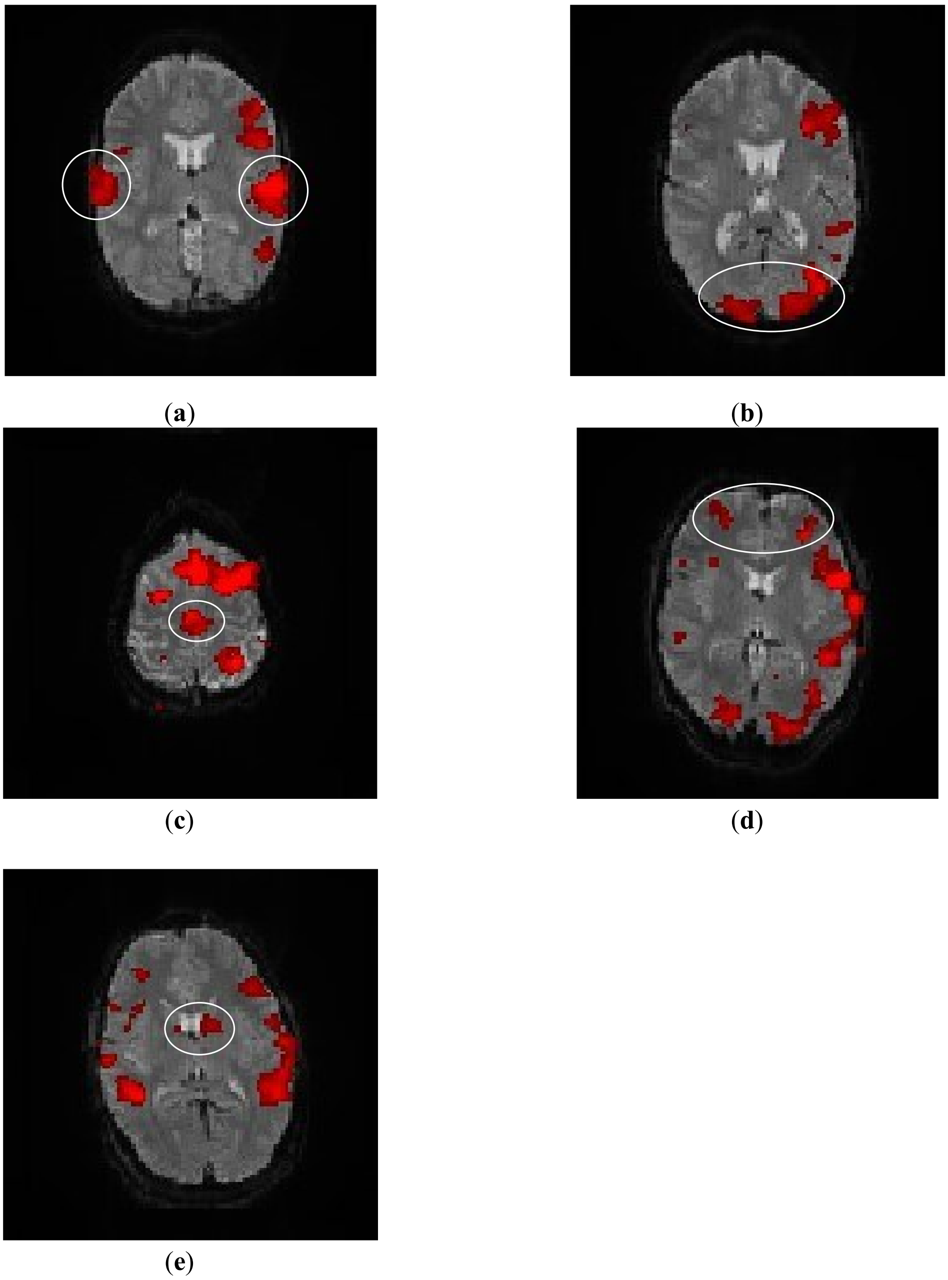


Figure 4: Regions of interest other than the LAs. The white circles show regions that are activated frequently by the tasks during the language fMRI examination. Measured activations are displayed red on the activation maps.
(a) Auditory cortex, coactivated by the acoustic stimulus presentation.
(b) Visual cortex, coactivated by the visual stimulus presentation.
(c) Motor area, coactivated by language production.
(d) "frontal artifact"
(e) "ventricular artifact"

2.7 Creation of time courses

The TCs were a representation of the activation dependent fluctuations of the BOLD-signal caused by the HRF-dependent changes in the blood-flow of activated brain regions during the language fMRI examination. To make TCs of different signal levels comparable and to give a good representation of positive and negative fluctuations, the mean signal of the TC was set as baseline and changes of signal were defined as the difference to the baseline value.

$$\Delta TC = TC(x) - mean\ (TC)$$

For the standardization of the signal strength (y-value), the TCs were converted to percent:

$$TC\ [\%]\ =\ ((TC -\ mean\ (TC))\ /\ (mean\ (TC))\ \times 100$$

Different versions of TCs were created and used for this study. Figure 5a shows the TC of a selected region (in this case Broca’s area) on one activation map, consisting of 10-100 voxels of one fMRI data set. This provided information about the behavior of activations of individual datasets but was not suitable to study the region dependent characteristics of activations in general.

To receive more general information about the region dependent behavior of TCs, TCs of multiple fMRI datasets from the same region were combined and a TC showing the mean TC of activations in this region was created (Figure 5b). These TCs were designed to find region-specific characteristics of the activations. The graph also indicates the maximum positive and negative deviation of the signal strength (y-axis) at each TP, indicating the variability of the values.

The third plot shows the cumulated information contained in all 6 periods of the plot in Figure 5b, displaying the mean of all periods as one period, representing the mean of the signal (Figure 5c). This provided a good overview of the activation signal and simplified the comparison of the TCs of different regions.

The last plot shows the slope of the TC. The slope of a TP was calculated by the slope to its next TP, representing a timeframe of 3s (Figure 5d, red) or its second next TP, (representing a timeframe of 6s (Figure 5d, blue). In this study, mainly the slope measurement referring to the second next time point (Figure 5d, blue) was used. This step enabled the analysis and definition of slope dependent parameters.

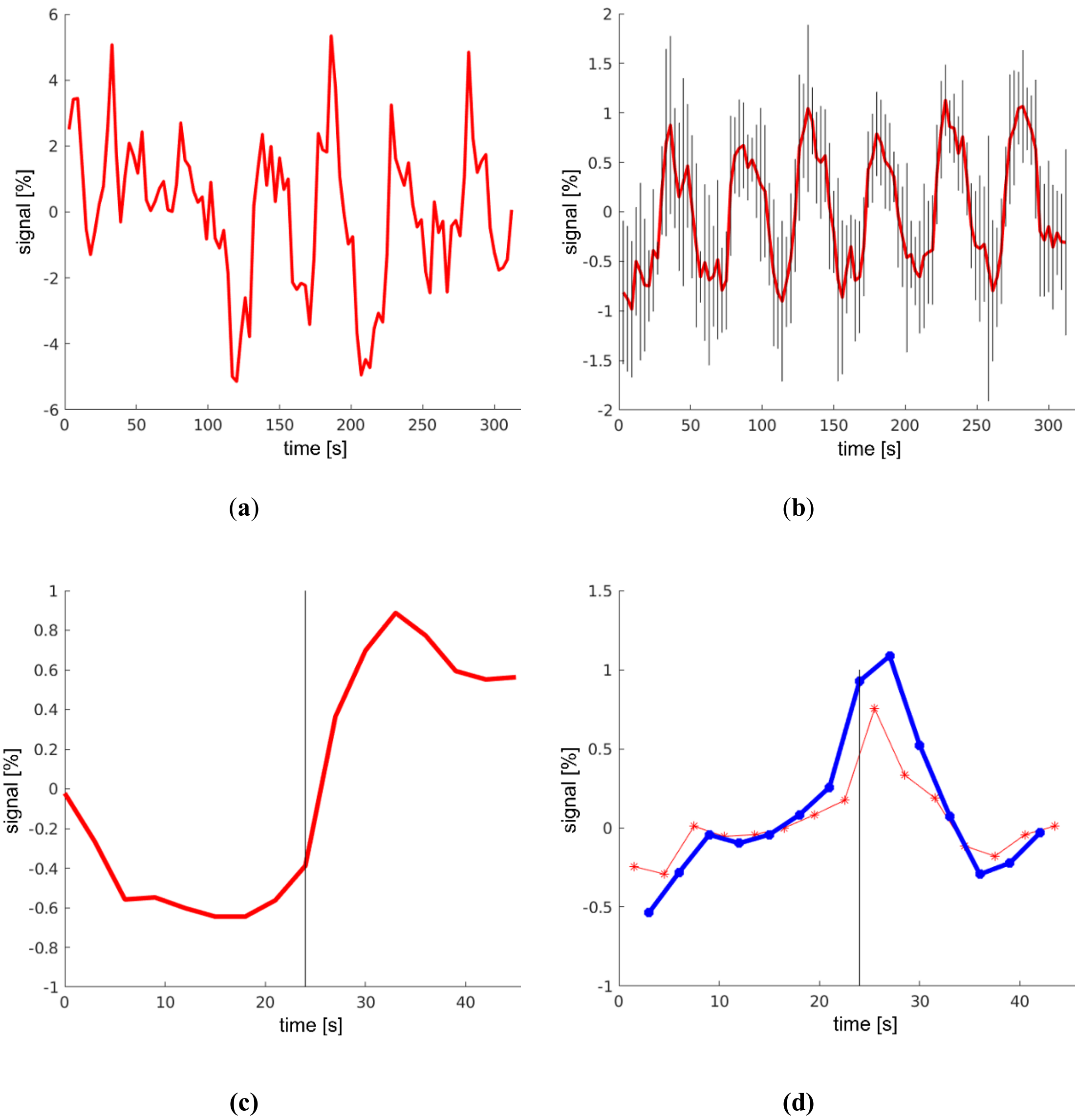


Figure 5: Different ways the TC and the slope of regions in one activation map were displayed.
a) TC of the selected voxels in a region (here: Broca's area) of one scan.
b) TC of a selected region (here: Broca's area) averaged over multiple scans. Black bars show the max. pos./neg. deviation from the mean.
c) TC information of all 6 periods combined and displayed as one period. The black bar at 24s marks the start of the activation-block.
d) The slope of the signal, calculated as the slope of each TP to its next (red) or second next (blue) time point is shown. The black bar at 24s marks the start of the activation-block.

2.8 Defining characteristics of regions and parameters

The TCs that were previously saved as ROIs and grouped depending on the region, were analyzed with the goal of finding characteristics specific for different regions. Initially all TCs included in each group were plotted individually to exclude obvious outliers that complicate the defining of characteristics (Figure 6 a). The individual TCs were then combined, the average TC was displayed (Figure 6 b), and the information of all six periods was summarized and plotted in one period (Figure 6 c), as well as the slope (Figure 6 d), as explained in 2.7.

The TCs displayed in Figure 6 c and d) show the typical TC of language activations evoked by the tasks used in this experiment. By comparing the average TCs of the different regions, six parameters were found to be useful to define the characteristics of language activations: time-to-onset and ending of activation [s] (measured by the time of the maximum and the minimum slope), the maximum and minimum signal strength [%], and the maximum and minimum slope [%/s] (Figure 6 c, d).

To find the appropriate language-specific values for the parameters, the averaged TCs of the different regions were used initially. The values of the parameters within the different regions measured from the averaged TCs were compared. This provided a rough impression of the specificity of the parameters and the range of values. The next, more precise, step consisted of a voxel-wise analysis, in which the values of the parameters measured from the TC of single voxels were monitored. The voxels of the whole map were compared to the voxels located in the LAs. The comparison of these values was then plotted in a histogram to clarify the TC-characteristics of the language activations and to find ranges of values that showed strong differences between activations of the LAs and the entire map, that would define the parameter limits. To verify the limit of each parameter and to minimize the number of false-negative voxels (especially deleted voxels within the LAs), the parameters were individually applied to the t-maps. This showed which activations were excluded through each parameter, making a final verification and adjustment of the limits possible.

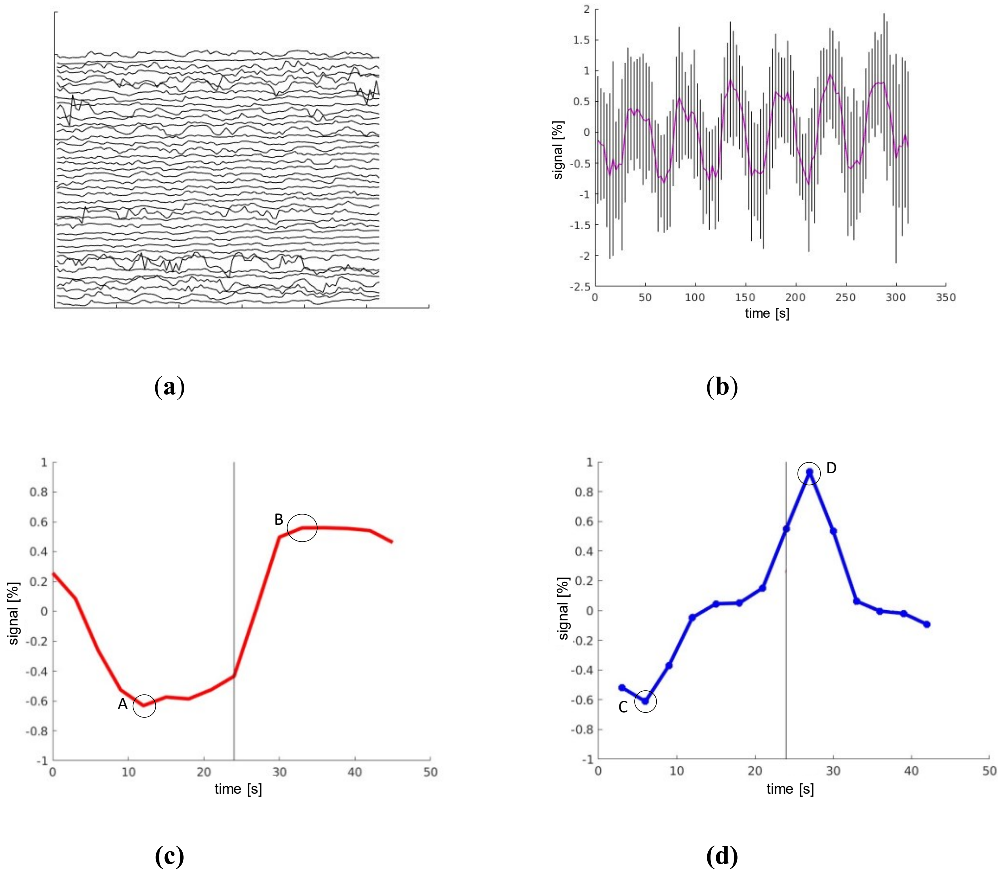


Figure 6: Display of the TC and slope of regions averaged over multiple scans.
a): Display of each TC included in the group of the region (here: Broca's area).
b): Averaged TC of all TCs shown in a). Black bars show the max. positive and negative deviation of the individual TCs.
c): TC consisting of one period (one rest and one activation block), combining the averaged information of all six periods.
A: minimum signal strength.
B: maximum signal strength.
d): Slope of the signal. The slope of one TP was calculated between the TP and its second next TP.
C: minimum slope and timing of the minimum slope.
D: maximum slope and timing of the maximum slope.

2.9 Filter t-maps

The defined parameters were integrated in the creation of the t-maps by using a filter, analyzing the TCs of every voxel marked as activation on the t-map and comparing the values of the TC with the defined limits of each parameter. A voxel was defined as correct activation if the TC-values lay within the limits of the parameters. A voxel was defined as incorrect activation and excluded from the t-map if the TC-values lay outside the limits of the parameters.

Lastly a cluster limit was applied, deleting all groups of activated voxels consisting of less than 3 voxels. This modification was made to reduce small artifacts and noise and thus improve the readability of the filter t-maps.

Matlab (MATLAB 12.0, The MathWorks, Inc., Natick, Massachusetts, United States) was used to create and execute the software for the filter t-maps.

2.10 Rating

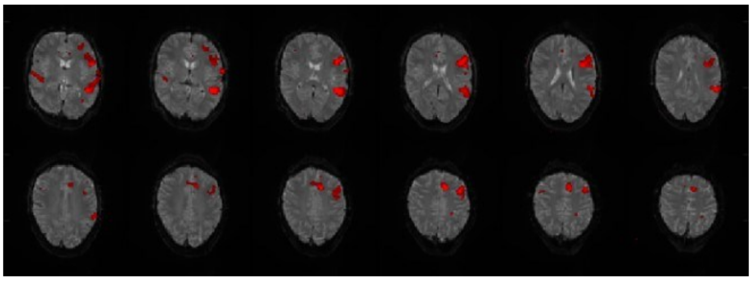
Two different rating scales were implemented, one to compare the used language tasks in terms of the appearance of the activation maps and the display of the LAs. The other rating scale compared the performance of the t-map and the GLM-map.

2.10.1 Rating of language task

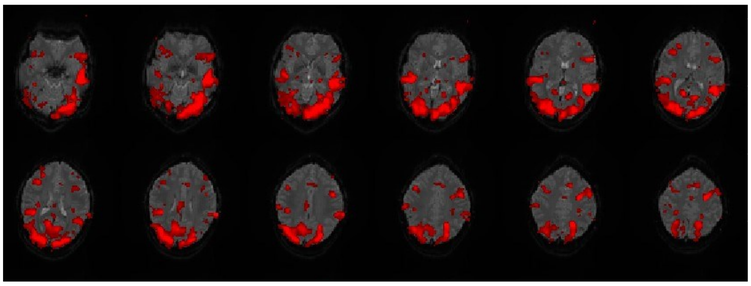
The important aspects for the evaluation of the language tasks were the display of Broca's area, the display of Wernicke's area and the overall quality of the map. In Table 2 the criteria for the assignment of the points were described. Looking at the LAs, the focus lay on the complete display of the region and the separability from neighboring activations. Rating the overall quality, the general appearance, such as the quality of signal and the number of artifacts on the map, was taken into count. Within each aspect, a maximum of 3 points (best) and a minimum of 0 points (worst) were distributed. The maps were rated between 9 and 0 points. An example for the rating of a map is given in Figure 7. Highly rated maps displayed complete and solitary activations of the LAs, making the identification of these and the separation to adjacent activations easy. Poorly rated maps, in comparison, displayed many artifacts, Broca's and Wernicke's area were incompletely displayed and hard to separate from the adjacent artifacts. The final rating is converted into a percentage value by dividing the achieved score by the highest possible score.

Table 2: Rating scale to evaluate the language tasks. Rating descending from 3 (highest score per criterion) to 0 (lowest score per criterion). Overall, the highest achievable score was 9 (3 + 3 + 3), the lowest highest achievable was 0 (0 + 0 + 0)

Points	**Broca**	**Wernicke**	**Signal**
3	very good display of Broca's area without adjacent artifacts	very good display of Wernicke's area without adjacent artifacts	very good display of the LAs, very few artifacts
2	good display of Broca's area with few adjacent artifacts	good display of Wernicke's area with few adjacent artifacts	good Signal, few artifacts
1	poor display of Broca's area with adjacent artifacts	poor display of Wernicke's area with adjacent artifacts	bad signal, many artifacts
0	no/ not usable display of the Broca area	no/ not usable display of the Wernicke area	no signal



a)



b)

Figure 7: Exemplary rating of two activation maps. Measured activations are displayed red on the activation maps.
(a) Map rated 9 (3 + 3+ 3); the display of Broca's and Wernicke's area is complete, no artifacts adjacent to the LAs that could be falsely associated with the LAs. The overall signal is good, and the map shows only few artifacts.
(b) Map rated 3 (1 + 1 + 1). Broca's and Wernicke's area can be identified, but a distinction between the LAs and adjacent artifacts is impossible. The overall signal is bad, as too many unwanted activations and artifacts are displayed all over the map.

2.10.2 Rating of the filter t-map

The performance of the filter t-map was evaluated by a direct comparison of filter t-map to the GLM-map. Two different activation maps of 48 fMRI datasets of the subjects were created, one using the method of the filter t-map, one using the method of the GLM-map. The rating was based on the performance of the filter t-maps. The possible rating of the filter t-map was +1, 0 or -1 point. In case the filtered t-map showed improved display of LAs and general readability, +1 point was given; for equal performance of both maps 0 points were given; for worse performance of the filter t-map, -1 point was given. In Figure 8 an exemplary map evaluation is demonstrated. The rating of Finally the performance of the t-map was rated based on the final score, summarizing the rating of each activation map of all 48 fMRI datasets; a positive score indicated an improvement using the filter t-map.

Since the 48 datasets of the subject study were also used for the creation the parameters of filter and definition of their limits, the evaluation may be biased. Therefore, a second evaluation of the filter t-map was performed with 30 datasets of patients, in the same way the datasets of the subjects were rated. This intended to counteract this possible bias and at the same time verify the clinical suitability of the filter t-map.

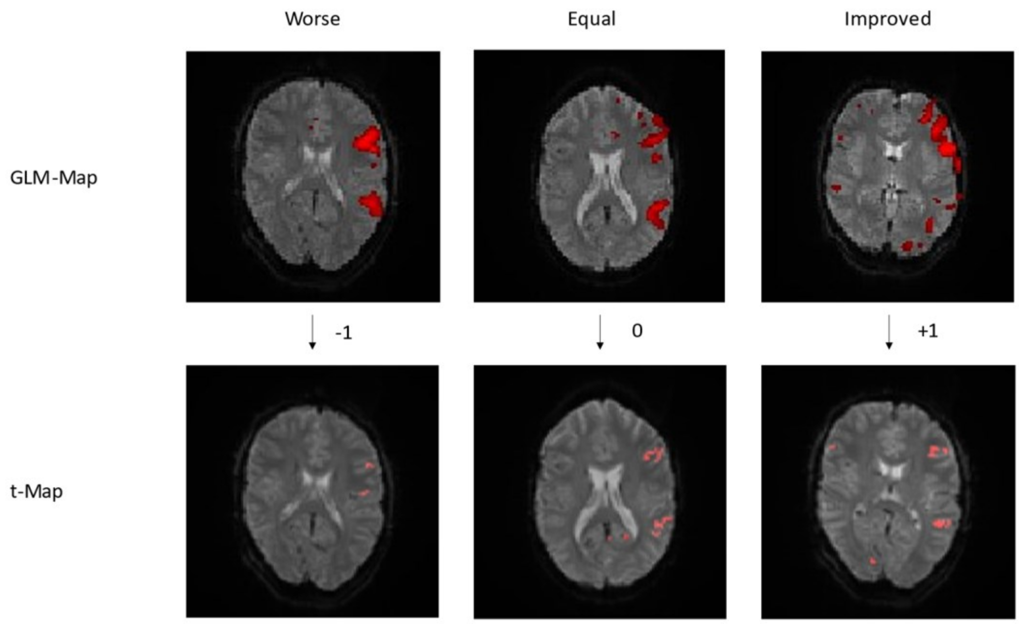


Figure 8: Exemplary comparison between a filter t-map and a GLM-map. Measured activations are displayed red on the activation maps.
Left: Worse performance of the filter t-map compared to the GLM-map: the LAs are displayed incompletely on the filter t-map. -> -1 point
Middle: Both activation maps show similar results. -> 0 points
Right: Improved performance of the filter t-map compared to the GLM-map: fewer artifacts are displayed and the LAs are better distinguishable from adjacent activations. -> +1 point

3 Results

104 scans were conducted during the subject study. 18 scans of the 104 scans had to be excluded from the study due to technical complications, excessive head movement of the subject or bad signal. The remaining 86 scans were successful and of adequate quality. These scans were used for the evaluation of the language tasks and the examination of the experiment settings. For practical reason, only eight fMRI datasets of the subjects, each consisting of six successful scans, were used for the creation of the filter t-map. The eight subjects were selected randomly.

Nine of the 30 patient datasets had to be excluded from this study, because they were either incomplete, showed measurement errors, or had other errors that made it impossible to identify the LAs or evaluate the map. For some other aspects of the study, e.g., the training effect, all 41 scans of 30 patients, were used.

In summary, 86 scans from 20 subjects and 41 scans from 30 patients, leading to a total of 127 language fMRI scans that were used in this study.

3.1 Language Tasks

To evaluate the performance of the different language tasks and stimuli (visual or acoustic) used for the scans of the subjects, the rating scale described in 2.10.1 (Table 3) was used. The evaluation of the language tasks was based on 3 aspects: the display of Broca's area, the display of Wernicke's area and the quality of signal/the number of artifacts.

First, the overall effect of the language tasks on the activation maps was evaluated by combining the results of all three rating categories (Table 3, Figure 9). The overall best rated language task, visually and acoustically, was the language task "Naming" with an average score of around 70%. Second best rated were the two acoustic tasks "verb-generation (acoustic)" and "sentence-completion (acoustic)", both rated around 68%. The two visual tasks performed worse in the overall rating: 65% for "verb-generation (visual)" and 62.5% for "sentence-completion (visual)".

Table 3: Language-task-rating including all evaluation aspects (Broca's area, Wernicke's area, and signal/artifacts).

Language Tasks	Mean	N	Std.-Deviation
Verb-Generation (visual)	65.22% (5.87)	15	1.36
Naming (visual)	**69.89% (6.29)**	14	.99
Sentence-Completion (visual)	62.56% (5.63)	16	.96
Verb-Generation (acoustic)	68.33% (6.15)	13	1.73
Naming (acoustic)	**70.67% (6.36)**	14	1.78
Sentence-Completion (acoustic)	68.22% (6.14)	14	1.87
Total	67.33% (6.06)	86	1.46

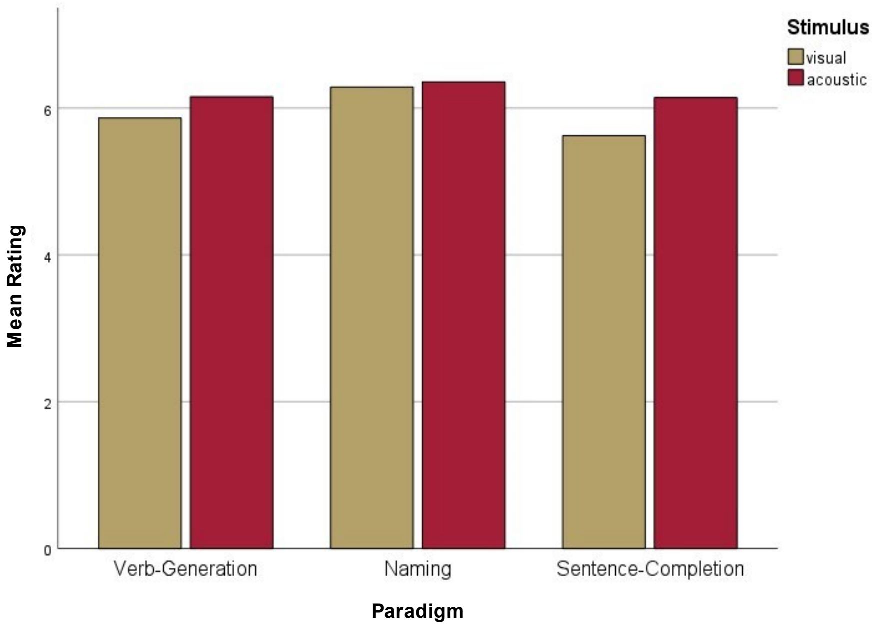


Figure 9: Histogram showing the language-task-rating of all evaluation aspects (Broca's area, Wernicke's area, and signal/artifacts). The mean rating is displayed as the average point-score each language task achieved.

The rating of the display of Broca's area is shown in Table 4 and Figure 10. Maps with the language task "naming (visual)" showed overall the best display of Broca's area, allowing a clear definition and delimitation of this area. Accordingly, "naming (visual)" achieved the highest score (78.67%), followed by the two acoustic tasks "verb-generation (acoustic)" (74.33%) and "naming (acoustic)" (73.67%). The two visual tasks achieved the worst ratings, "sentence completion (visual)" with a rating of 68.67% and "verb-generation (visual)" with 64.33%. These results were consistent with the overall ranking.

Table 4: Language-task-rating of the display of Broca's area.

Language Tasks	**Mean**	**N**	**Std.-Deviation**
Verb-Generation (visual)	64.33% (1,93)	15	.59
Naming (visual)	**78.67% (2,36)**	14	.50
Sentence-Completion (visual)	68.67% (2,06)	16	.44
Verb-Generation (acoustic)	74.33% (2,23)	13	.73
Naming (acoustic)	73.67% (2,21)	14	.70
Sentence-Completion (acoustic)	71.33% (2,14)	14	.66
Total	71.67% (2,15)	86	.60

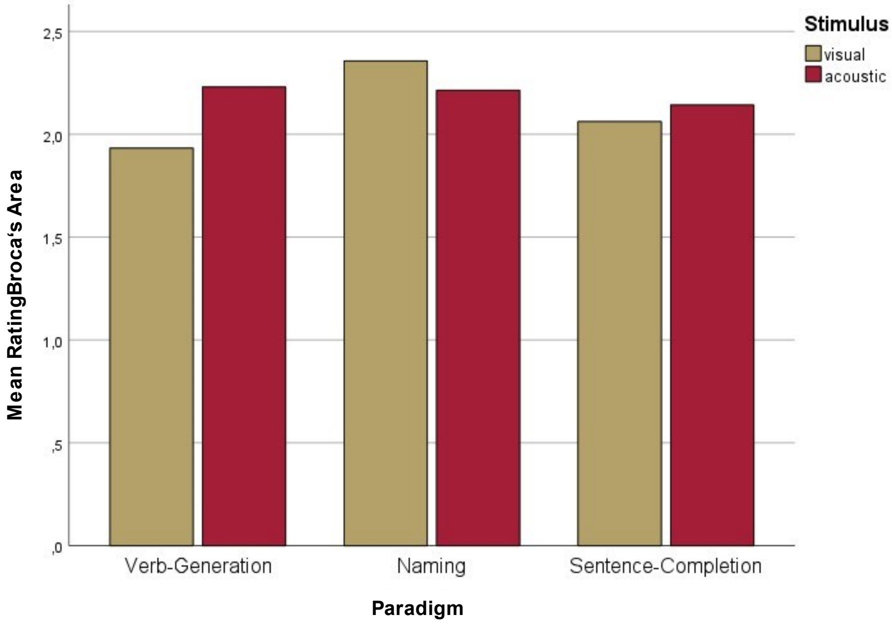


Figure 10: Histogram showing the language-task-rating of the display of Broca's area. The mean rating is displayed as the average point-score each language task achieved.

The rating of the display of Wernicke's area was generally worse than the display of Broca's area, independent of the language task. The reason for this could be that the language tasks were not ideally suited for the activation of Wernicke's area, or, more likely, that the activation signal of Wernicke's area is not as strong as activations in Broca's area, leading to a generally weaker display of Wernicke's area. The rating was around 65% for most language tasks. Only the visual tasks "verb-generation (visual)", and "sentence-completion (visual)" stood out with a rating of around 70%, making those the best rated language task for the display of Wernicke's area (Table 5, Figure 11).

Table 5: Language-task-rating of the display of Wernicke's area.

Language Tasks	Mean	N	Std.-Deviation
Verb-Generation (visual)	**69.00% (2,07)**	15	.88
Naming (visual)	64.33% (1,93)	14	.73
Sentence-Completion (visual)	68.67% (2,06)	16	.44
Verb-Generation (acoustic)	64.00% (1,92)	13	.76
Naming (acoustic)	64.33% (1,93)	14	.83
Sentence-Completion (acoustic)	64.33% (1,93)	14	.73
Total	66.00% (1,98)	86	.72

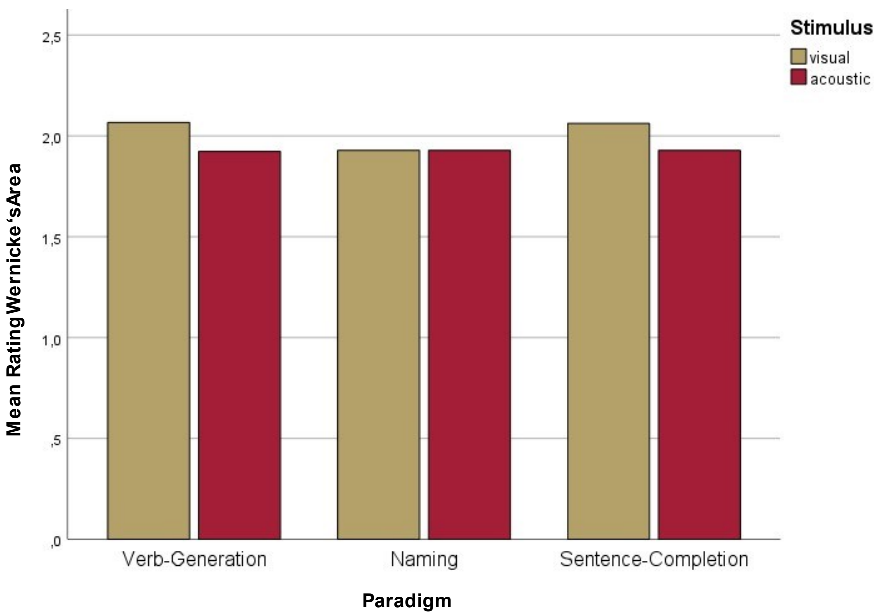


Figure 11: Histogram showing the language-task-rating of the display of Wernicke's area. The mean rating is displayed as the average point-score each language task achieved.

The rating of the signal quality and number of artifacts is shown in (Table 6, Figure 12). The acoustic tasks were generally ranked higher, showing generally a better signal and the display of less artifacts, implying a higher robustness and lower sensibility for subject dependent errors/artifacts. Regardless of the type of the stimulus, the "naming" task again proved to be the language task that resulted in the best signal and the fewest artifacts on the activation map. This result is congruent with overall rating.

Table 6: Language-task-rating of the quality of the overall signal and the number of artifacts.

Language Tasks	**Mean**	**N**	**Std.-Deviation**
Verb-Generation (visual)	62.33% (1.87)	15	.52
Naming (visual)	66.67% (2.00)	14	.39
Sentence-Completion (visual)	50% (1.50)	16	.52
Verb-Generation (acoustic)	66.67% (2.00)	13	.58
Naming (acoustic)	**73.67% (2.21)**	14	.58
Sentence-Completion (acoustic)	69% (2.07)	14	.73
Total	64.33% (1.93)	86	.59

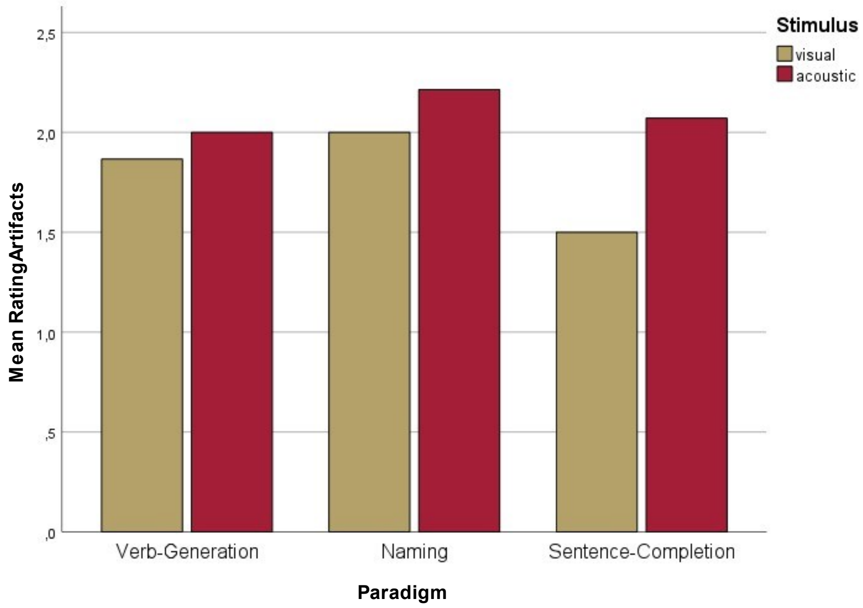


Figure 12: Histogram showing the language task-rating of the overall signal and the display of artifacts. The mean rating is displayed as the average point-score each language task achieved.

In conclusion, it was demonstrated that the choice of the language tasks is important as it led to differences in the quality of the activation maps and the display of the LAs. The language tasks "naming" was the overall an individually best performing language tasks, independent of the stimulus.

To make conclusions about the choice of stimulus presentation, the overall rating scores of all maps using the visual stimulus were compared with all maps using the acoustic stimulation (Table 7), to make conclusions about the choice of stimulus presentation. The acoustic presentation of language tasks (rating 69,11%) achieved better results than the visual presentation (65,67%). It can be assumed that both stimuli can be applied alternately, adjusted to the patient's conditions, without having to expect a significant loss or change of quality in the creation of the activation maps, since the differences between the stimuli were minor.

Table 7: Rating of the different ways of stimulus presentation: visual and acoustical.

Presentation	Mean Score	N	Std.-Deviation
visual	65.67% (5.91)	45	1.13
acoustical	**69.11% (6.22)**	41	1.75
Total	67.33% (6.06)	86	1.46

Two different stimulus-specific artifact-patterns occurred, while comparing the effect of the two different stimuli on the activation maps, (Figure 13). The visual stimuli led to stimulus dependent activations in the visual cortex and tended to show more (apparent) activations and artifacts in total, leading to a slightly lower rating of the map-quality, as it may affect the identification and differentiation of the targeted LAs negatively.

Maps created with acoustic stimuli showed activations in the auditory cortex, but tended to show less artifacts in total, which may explain the overall good rating of the acoustic tasks. The auditory cortex is located closely to Wernicke's area, making the identification of Wernicke's area more difficult. Yet, with basic anatomical knowledge about the location of these regions, the stimulus dependent activation of the auditory cortex can be easily separated from the activations of Wernicke's area (Figure 13).

Activation maps created with visual language tasks showed activations mainly in the visual cortex area. These activations are located far occipital and can hardly be mistaken for language activations. Nevertheless, the visual stimulation resulted in worse ratings compared to the auditory stimulation, as more artifacts also occurred outside the visual cortex, distributed over the whole activation map (Figure 13).

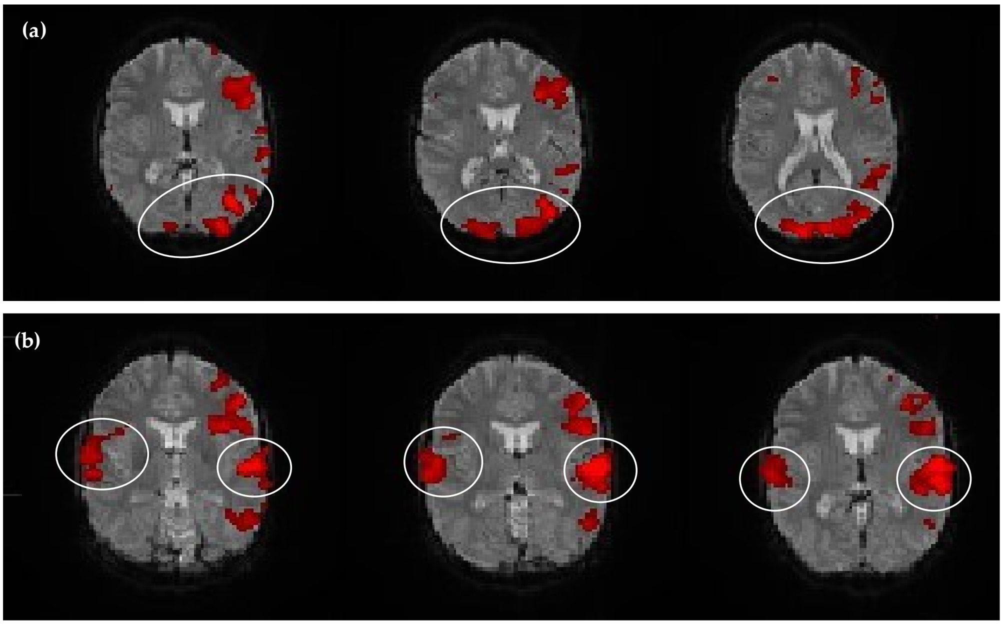


Figure 13: Differences in the stimulus-dependent activation pattern. Measured activations are displayed red on the activation maps.
(a) Visual stimulus with activation of the visual cortex (circled).
(b) Acoustic stimulus with activation of the auditory cortex (circled).

3.2 Training effect

The first period of the TCs, at the beginning of the scan, often performed worse than the following periods (Figure 14). TCs from of all 30 patient's scans (n=41) were analyzed and each period was defined usable or unusable based on the quality of the TC. The results were as expected: the first period was in more than half of the scans (51.2%) unusable, meaning that the TC did not match basic characteristics of the experiment, such as time of activation, etc. (Figure 14,

Table 8). The following periods showed consistently better quality, even though period 2 and 6 showed a couple of outliers. These results indicate that there may be a "training effect", meaning that the patient/subject must learn the required task at first and, after getting used to it, improves in the execution of the task.

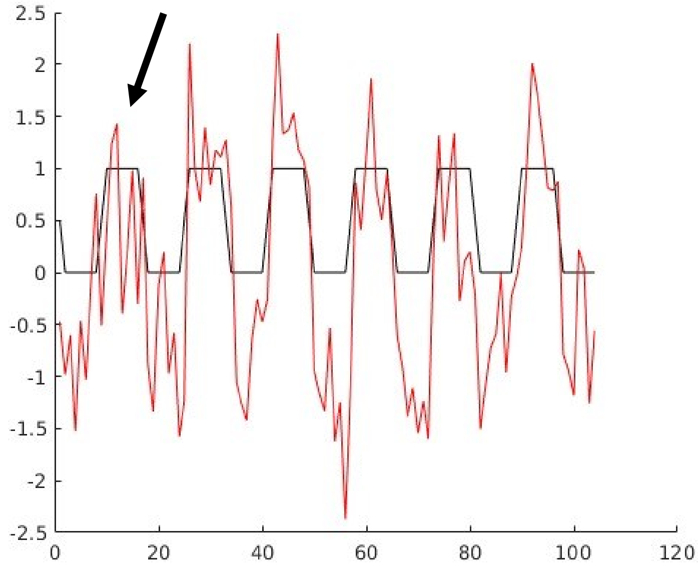


Figure 14: Exemplary TC with poor signal in the first period. TC (red); poor signal in the first period (arrow). The black line demonstrates the prototypical course of signal matching the experiment setup (y = 0 represents the time during the rest block; y = 1 represents the time during the activation block).

Table 8: Evaluation of individual periods in the TCs of patient data (n=41). The periods were defined usable or unusable depending on quality.

Period	**1**	**2**	**3**	**4**	**5**	**6**
Usable quality	48.8% (20)	70.7% (29)	92.7% (38)	90.2% (37)	85.4% (35)	87.8% (31)
unusable quality	51.2% (21)	29.3% (12)	7.3% (3)	9.8% (4)	14.6% (6)	12.2% (10)

To follow up on this theory, activation maps were created without the first TC period and compared to activation maps with all six TC periods. The deletion of the first period barely made a difference. Only a few activation maps showed an improvement caused by the deletion of the first period, while in most cases, the activation maps remained mostly unchanged. In very few cases the deletion of the first period even led to a worse display of LAs and more artifacts on the activation maps. This may be due to the reduction of the total number of periods, leading to a greater impact of artifacts/measurement errors on the activation maps.

In summary, the deletion the first period can have a positive effect on the activation maps as shown in Figure 15. In this (positive) example, the exclusion of the first TC period has resulted in a clearer and more complete display of the LAs, especially Wernicke's area. The second example in Figure 15 shows a better representation of the average effect that the exclusion of the first TC period had. This example showed that neither the display of the LAs was improved nor the number of artifacts was reduced.

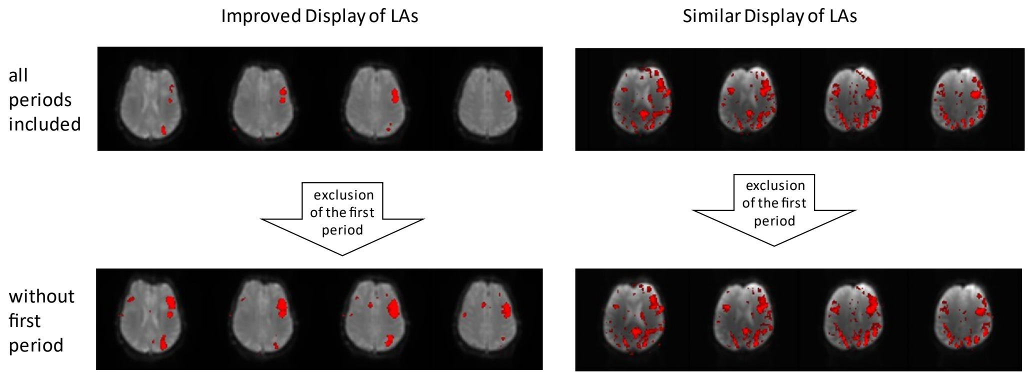


Figure 15: Comparison of an activation map created with all six periods (top) to activation maps created without the first period to show the effect of the exclusion of the first TC period on the simple t-maps. Measured activations are displayed red on the activation maps.
Top: Simple t-maps calculated with all six periods.
Bottom: Simple t-maps calculated without the first period.
Left: Example of a dataset in which the exclusion of the first TC period led to an improvement in the display of activations (especially LAs) on the activation maps.
right: Example of a dataset in which the exclusion of the first period led to similar results regarding the display of activations on the activation maps.

3.3 Number of periods

A thesis was formed based on the earlier observation that showed the negative effect of the use of fewer periods on the display of activations: an increasement of the number of periods may increase the quality of activation maps.

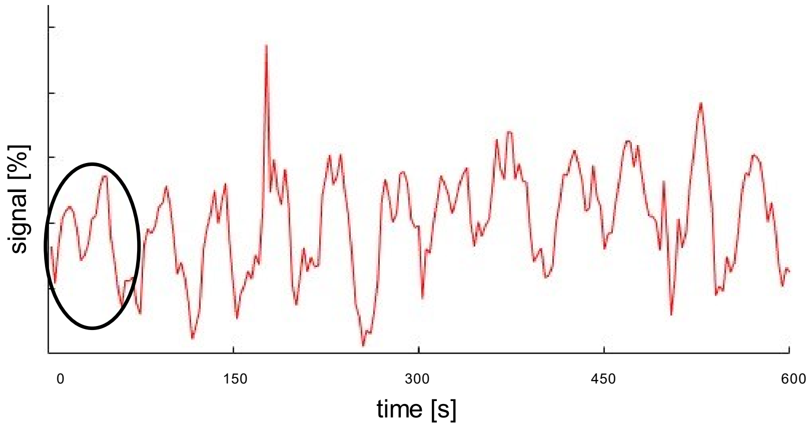
The number of periods was doubled in four scans (12 periods instead of 6 periods) to study this observation, resulting in a total time of 10:00 minutes per scan. The TCs of two extended scans are shown in Figure 16. As expected, the TC of the first period was of poor quality compared to the following periods. Periods 2-10 were consistently of good quality and from period 11 onwards, a reduction of quality appeared in half of the extended scans.

In Figure 16(a) the first period showed irregularities in the form of a dip during the activation plateau. This indicates that the signal in this period was of insufficient quality. In Figure 16(b) the first period was of good quality, but the last two periods (11 and 12) were showed bad signal in form of irregular patterns. The periods can hardly be distinguished and neither the activation phase with the activation plateau nor the pause phase were identifiable.

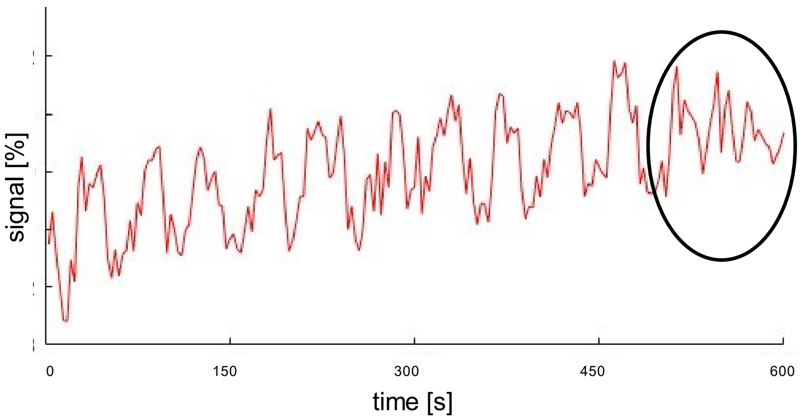
The comparison of the 12-period scans to the 6-period scans, showed, that the signal of the 12-period scan was a lot stronger. This may be of advantage when creating maps with datasets have a to low signal or the maps display the LAs incompletely.

Figure 17 shows the comparison of 6-period activation map to a 12-period activation map of the same subject. In both cases, a language task with acoustic stimulus presentation was used. The 12-period activation map shows a stronger signal in general, leading to an increase of artifacts and unwanted regions (such as the auditory cortex), making the identification and separation of LAs more difficult.

The increase of periods led to more artifacts and a worse display of LAs in all extended activation maps, compared to the datasets of the subjects with six periods that were mostly of good quality with a complete display of the LAs. Six periods per scan is a well-balanced number of periods when working with good datasets, enabling the creation of good quality activation maps, and offering a shorter duration of the scan.



(a)



(b)

Figure 16: Extended TCs with 12 periods.
(a)*: First period with bad signal quality (circle).*
(b)*: Period 11 and 12 with bad signal quality (circle).*

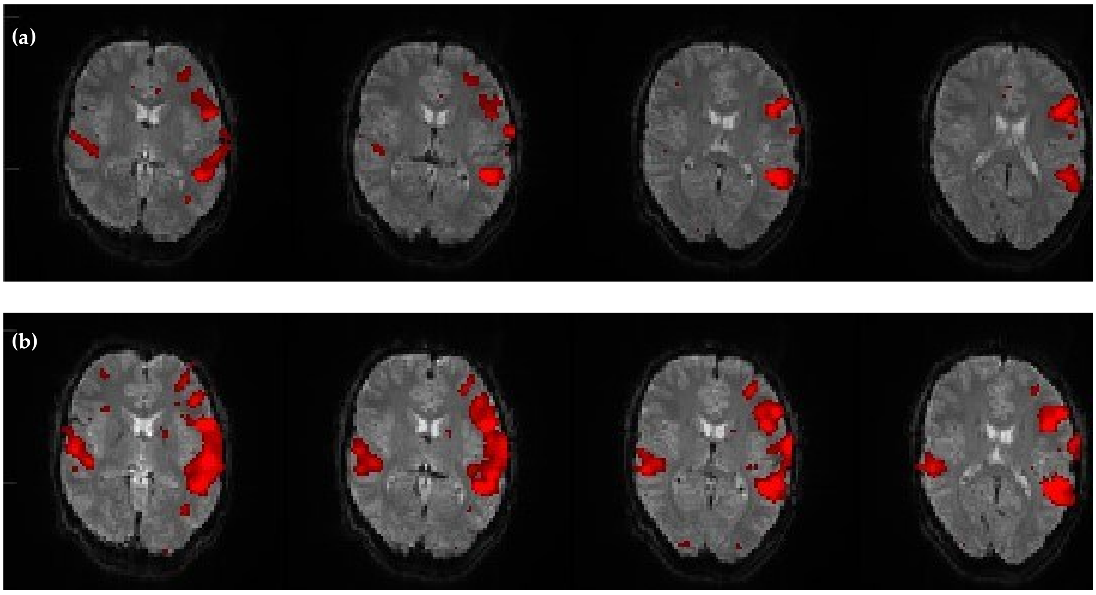


Figure 17: Comparison of two activation maps: one activation map was measured using 6 periods (a) the other activation map was measured using 12 periods (b). Both scans were performed with the same subject. Measured activations are displayed red on the activation maps.

3.4 Rest Block (Activated Pause)

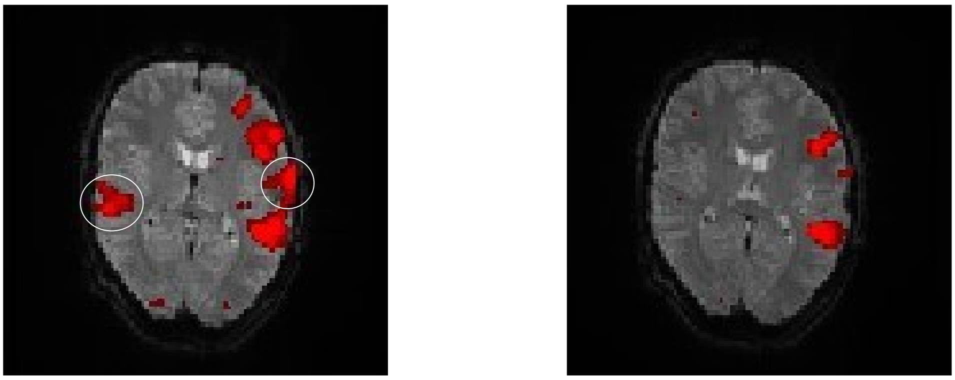
The experimental design of the subject study consisted of an activated pause during the rest-block. During the activated pause the subject/patient executed non-language tasks prompted the same way as the stimulus of the language tasks during the activation-block (visual or acoustic). Scans using an acoustic task and an (acoustic) activated pause were compared to scans that used acoustic tasks and a silent pause, to evaluate the effect of the activated pause on the activation maps. The silent pause was prompted visually with the word "pause" on the screen at the beginning of the rest block, without the stimulation of the auditory cortex. Also, there was no task to be executed during the rest-block, subsequently the subject remained without stimulation for the duration of the rest block (24s) and waited for the start of the next activation block.

The effect of the activated pause compared to the silence pause on the activation maps is shown in Figure 18.

The activation map created with the silent pause (Figure 18(a)) showed a strong activation of the auditory cortex. These activations, due to their proximity to Wernicke's area, have a negative effect on the localization of the LAs. The activation map created with an activated pause (Figure 18(b)) showed less activations in general, especially of the auditory cortex, simplifying the identification of the exact location of Wernicke's area. The ability to identify and localize the LAs is considerably improved using the activated pause compared to the silent-pause map.

The reduced number of activations of the auditory cortex by using the activated pause could be explained by the technique used to analyze the fMRI data. Activations are defined by a significant difference in TCs between activation- and rest-blocks. If, as in this case, the auditory cortex was also stimulated during the rest-block, the TCs of activated voxels in the auditory cortex were expected to be activated during activation- and rest-block. Therefore, less significant differences between activation and rest block can be found, leading to this region not being defined as an activation. This is particularly important for the display of Wernicke's area, because of its proximity to the auditory cortex.

The reduction of the total number of activations could be explained by the fact that the patient/subject remained focused during the rest-block. The wandering of thoughts is reduced due to the distraction by the acoustic stimulation and the motoric tasks during the activated pause. This theory is consistent with the subjects' feedback, as the use of an activated pause was described to be more pleasant as it kept the subjects busy, improving the maintenance of focus and reduced the occurrence of fatigue or loss of concentration.



(**a**) (**b**)

Figure 18: Comparison of activation maps created with silent pause (a) and activated pause (fist movement prompted acoustically during the rest block) (b). Measured activations are displayed red on the activation maps.
(a) Silent pause: This map shows strong activation of the auditory cortex (circled). These activations, due to their proximity, have a negative effect on the localization of the LAs (especially Wernicke's area).
(b) Activated pause: Here, almost no activation of the auditory cortex is visible, and the overall number of artifacts was reduced. The localization of the LAs is considerably improved compared to the silent-pause map.

3.5 Subject feedback

A short interview was conducted with the subject after the language fMRI examination. The subjects were asked about the difficulty, the executability, the presentation and the display of the different language tasks. The duration of the display of the tasks on the screen was also part of the interview to find the duration that provides the best readability of the words and sentences.

All subjects were healthy and cognitively fit and were able to understand or figure out the requirement of the tasks, even if only brief explanations were given. Thus, according to most subjects, a detailed explanation of the procedure and the language tasks prior to the scan was essential for the successful execution of the tasks. Subjectively, uncertainty about the language task or the process of the examination led to confusion and poor execution at the beginning of the tasks. In the case of cognitively impaired patients, the explanation of the language tasks is even more important, as it is generally more challenging to understand and perform the required language tasks for mentally impaired patients.

The difficulty of the tasks was described reasonable. The subjects rated the verb generation task as the easiest and the naming and sentence completion task as more difficult but equally challenging. Occasional problems with the (acoustic) comprehension of longer acoustic tasks (especially with the language task SC) were reported. This problem is due to the measurement-related noise of the MRI scanner and has been solved in most cases by using better headphones and increasing the volume.

Regarding the concentration of the subjects during the scan, a comfortable body position was very important. An uncomfortable position led to a decrease in concentration and an increase of movements, reflecting negatively on the outcome of the activation maps.

At the beginning of every new language task, a cognitive adjustment by the subject to the new task had to take place, leading to a weaker performance of the task. This is consistent with the reduced quality of the first period of the TCs. A decrease in concentration towards the end of the scan was subjectively not noticed. Fatigue was described occasionally during the rest blocks, but the use of an activated pause helped most subjects to stay focused and awake.

3.6 Filter t-Maps

A major part of this thesis was the creation of activation maps using unprocessed fMRI data, to generate a realistic and easy to evaluate representation of the localization of LAs. Parts of the results described in chapter 3.6 have previously been published in Fetscher et al. 2022.

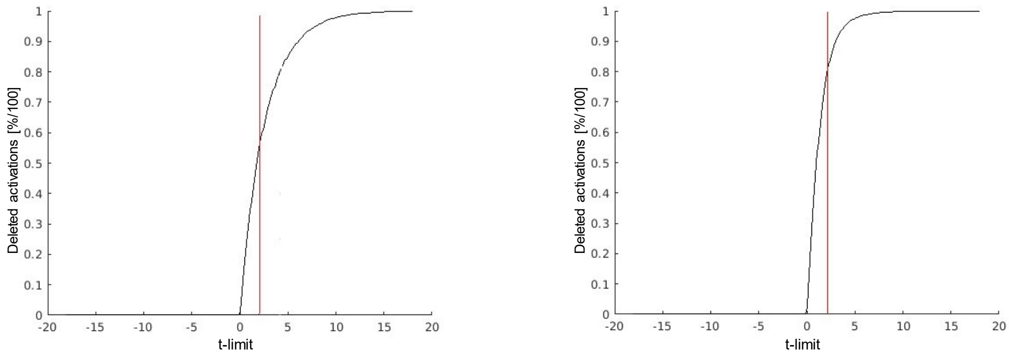
The definition of the activations on the maps was based on a t-test (t-map), as this allows the analyzation of the raw fMRI data without any alternation. For the creation of the t-maps, a universally valid t-value was defined that was used for all used language fMRI datasets.

To highlight the LAs and to simplify the evaluation of the activation maps, HRF-dependent parameters were defined by analyzing the TCs, enabling an improved separation of the activations of the LAs from other activated regions and artifacts. When creating a t-map, all temporal information contained in the TCs is lost, leading to a reduction in the specificity of displayed activations on the activation maps. Therefore, parameters that were based on the temporal information were implemented, enabling the use of the temporal information included in the TCs for the creation of activation maps. Limits were defined for each parameter, to include most of all activated voxels within the LAs and to exclude non-language activations.

The parameters with the defined limits were used like a filter, analyzing every single activated voxel displayed on the t-map, to, in the best case, only display activations of the LAs.

3.6.1 t-Limit

Figure 19 shows the number of voxels that were excluded from the activations maps at a given t-value. Two histograms were created, one demonstrating the number of excluded voxels in the LAs, the other demonstrating the number of excluded voxels on the entire map. Activations within the LAs are mostly more significant, as they mostly were excluded at higher t-limits, compared to all activations of the entire map. This can be seen in the lower slope of the graph representing the excluded activations of the LAs and the higher maximum t-value (t_0), defined as the value where the number of excluded activations goes to 100%, of the excluded activations of the LAs (t_0(LA) ≈ 15; t_0(entire) ≈ 9). It can be concluded that many activations of the LAs were still counted as activations at higher t-values, while many of the non-language activations were excluded. On the histograms, the t-value of 2.2 is highlighted with a red bar. At this t-value, 80% of the activated voxels of the entire map are excluded, while only 55% of the activated voxels within the LAs are excluded. There are 25% less activated voxels excluded in the LAs than on the entire map leading to the highlighting of these areas and the removing of unwanted activations and artifacts. The difference of 25% was the highest possible difference, therefore the general t-limit was set to t = 2.2.



(a) **(b)**

Figure 19: Number of excluded activations at the general t-limit of t = 2.2 (red line).
(a) Number of activated voxels within the LAs, 55% of activated voxels were excluded.
(b) Number of activated voxels of the *entire map, 80% of activated voxels were excluded.*

3.6.2 Filter parameters

Based on the interregional TC differences of different brain regions, a region-dependent analysis of the TCs was used to determine six parameters to distinguish activations of the LAs from other activated regions and artifacts. The six relevant parameters are shown in Figure 20: time-to-onset and time-to-end of activation [s] (measured by the timepoint of the maximum and the minimum slope), the maximum and minimum signal strength [%], and the maximum and minimum slope [%/s].

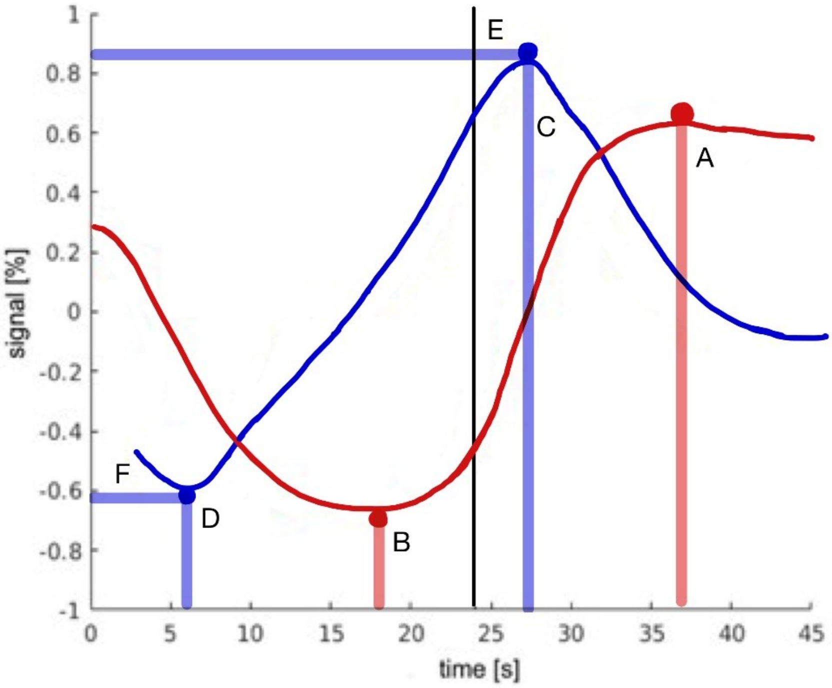


Figure 20: Demonstration of the six parameters on a perfect TC.
The TC (red) and slope of the TC (blue) demonstrates the expected behavior of a voxel in the LAs activated by the language tasks of the experiment. On the x-axis the time of one period is shown, the y-axis describes the signal change in relation to the average signal in percent and the black bar signals the start of the activation block at 24s. The blue or red lines named by capital letters show the value of one of the six parameters. A: Maximum signal strength; B: Minimum signal strength; C: Time of the onset of activation; D: Time of the end of activation; E: Maximum slope; F Minimum slope

3.6.2.1 Signal strength

The first step included the assessment of the data of whole regions, to measure the interregional differences in signal strength and to estimate the value of language activations. The different regions of all fMRI datasets used (n=48) were selected and cumulated, and the average of the region-specific value was determined (Table 9 and Table 10).

The mean values of the maximum and minimum signal strength of language activations were at +/-2%, with Broca's area showing higher (2.22% and -2.27%) and Wernicke's area showing lower (1.78% and -1.89%) values. Activations in the motor cortex tended to show lower values (1.34% and -1.52%), while activations in the auditory cortex (2.51% and -2.5%) and the "frontal artifact" (2.5% and -2.6%) contained higher values. This variability of values made it possible to distinguish activations of these regions from language activations by using a top and bottom signal strength limit in the filter. All areas, except for the "frontal artifact", showed relatively consistent values, based on a low standard deviation and variance.

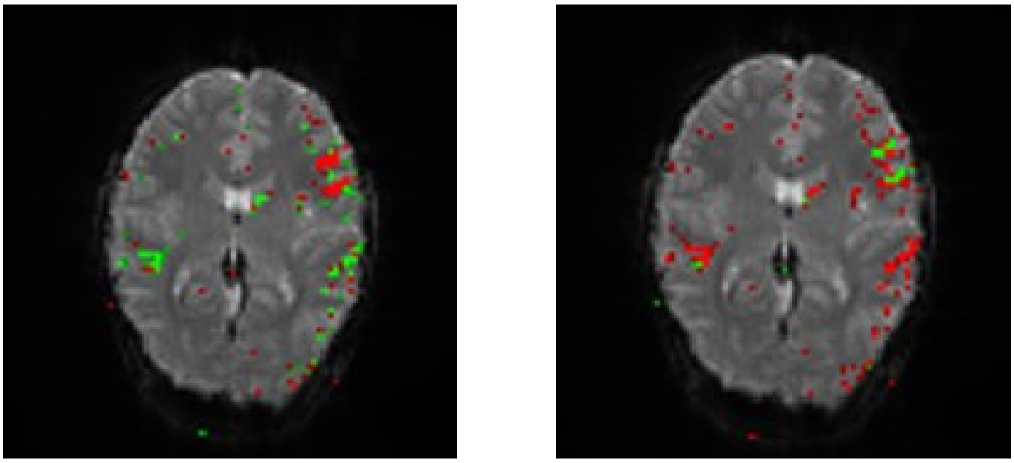
Table 9: Signal strength maximum of activations in different areas.

Region	**Mean**	**Variance**	**Std.-Deviation**
Broca	2,22	0,59	0,77
Wernicke	1,78	0,21	0,46
Language areas	2,01	046	0,68
SMA	1,34	0,33	0,58
Auditory Cortex	2,51	0,48	0,7
Visual Cortex	1,95	0,51	0,72
Frontal Artifact	2,5	2,2	1,49
Ventricular Artifact	1,96	0,35	0,6

Table 10: Signal strength minimum of activations in different areas.

Region	**Mean**	**Variance**	**Std.-Deviation**
Broca	-2,27	0,6	0,78
Wernicke	-1,89	0,26	0,51
Language areas	-2,09	0,47	0,69
SMA	-1,52	0,46	0,68
Auditory Cortex	-2,5	0,57	0,76
Visual Cortex	-1,96	0,49	0,7
Frontal Artifact	-2,6	2,33	1,53
Ventricular Artifact	-1,96	0,41	0,64

In the next step, to be able to estimate the value distribution, the (rounded) average signal strength of Broca's area was implemented as the maximum signal strength limit in the filter (max. signal strength <2.5%). The activation maps created this way were compared to maps with the filter setting of a max. signal strength limit >2.5%, to better demonstrate the voxels excluded by this filter setting (Figure 21). This comparison revealed a pattern in the displayed language activation: Differences in the signal strength between the margin and the center of language activations were observed, showing higher signal strength values in the center of the language activations and lower signal strength values on the sides. Only the edges of language activations were displayed while the center of the activation was excluded on the activation maps with a signal strength limit of <2.5%. The maps with a signal strength limit of >2.5% showed the exact opposite (Figure 21). This finding occurred mostly in the LAs but was also observed in other functional regions. Since the center of the LAs was frequently excluded at this limit, the range of values had to be defined wider to display the entire language activations.

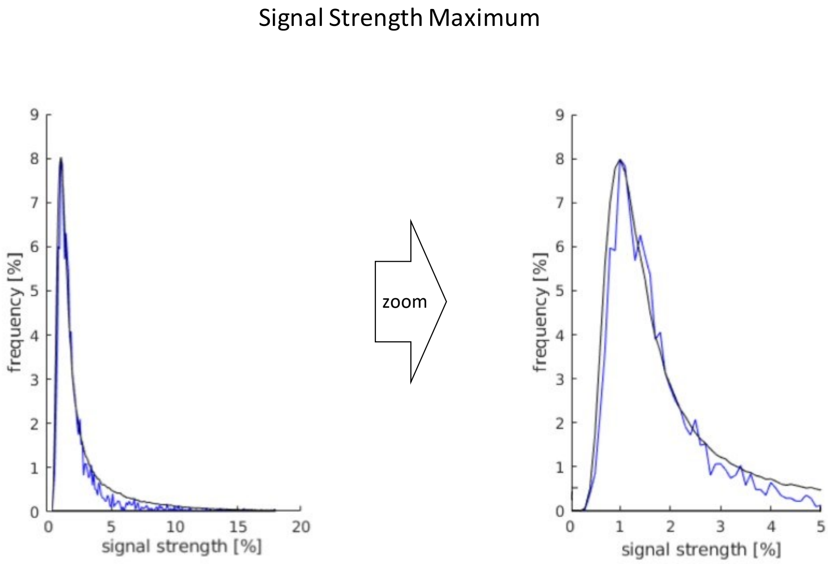


(**a**) (**b**)

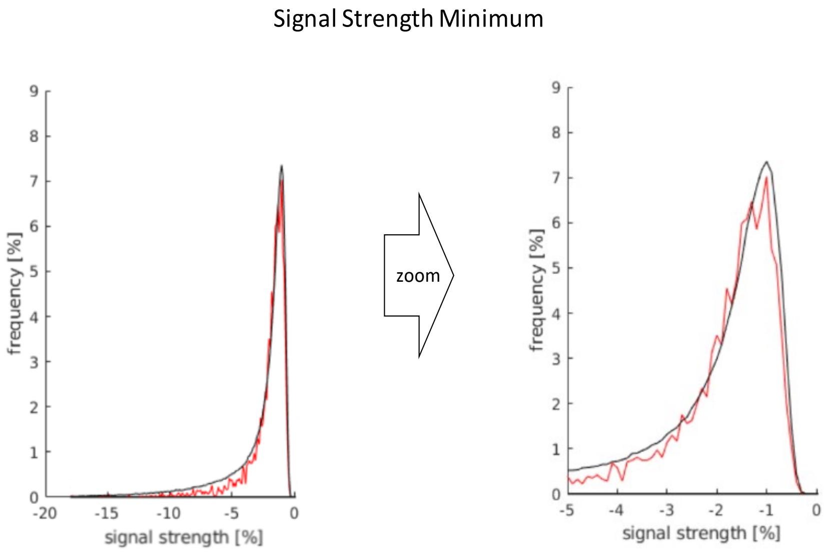
Figure 21: Parameter signal strength applied to a t-map to demonstrate a blood-vessel-dependent effect. Green activations lay within the limit, red voxels show measured activations that were excluded.
(a) Signal strength limit < 2.5%; (b) Signal strength limit > 2.5%.
Voxels in the center of the Broca area show higher signal strength (<2.5%) than the surrounding voxels.

To find the actual range of value for the parameter signal strength, including the signal strength values of the entire LAs, the TC values of each individual activated voxel on the activation maps were analyzed, and the maximum and minimum signal strength of each activated voxel was collected. This single voxel approach is similar to the method used in the filter that functioned with a single voxel TC analysis to inspect the activations. In Figure 22 the distribution of the signal strength minimum/maximum of the single voxels is shown. Two histograms were created, one demonstrating the maximum signal strength of language activations compared to the values of all activated voxels, the other one comparing the minimum signal strength values of language activations to the values of all activated voxels on the entire map.

The number of voxels in the LAs increased slower in the lower maxima/minima range (up to approx. +/-0.5%), compared to activations of the entire map. Between +/-0.5% and +/-5.5%, both groups showed a similar distribution of values. From approx. +/-5.5% on, the number of activations in the LAs showing high maxima/minima decreased faster. This led to the conclusion that most language activations have at their signal strength maximum of between +/-0.5% and +/- 5.5%, while the signal strength maxima/minima of all activations on the map extends over a larger interval.



(a)



(b)

Figure 22: Single voxel analysis of the parameter signal strength maximum (a) and signal strength minimum (b).
On the left, the measured signal strength maxima/minima from 0% - 20%/ -20% - 0% is shown. On the right, a magnification of the histogram is displayed, in which only the relevant range of values from 0%-5%/ -5% - 0% are displayed.
(a) Histogram of the signal strength maxima in the LAs (blue) and the signal strength maxima of the entire map (black).
The number of voxels (in percent) at different signal strength maxima was compared.
(b) Histogram of the signal strength minima in the LAs (red) and the signal strength minima of the entire map (black).
The number of voxels (in percent) at different signal strength minima was compared.

Based on these findings, the filter limit for the parameter signal strength was set to 0,5% – 5.5% for the maximum signal strength and -0,5% - -5.5% for the minimum signal strength, achieving the best results across all maps and deleting as few language activations as possible.

In Figure 23 an example of a filter t-map with only the parameter signal strength applied is shown. With this filter limit most of the LAs were preserved and some artifacts and apparent activations were excluded. However, many artifacts and apparent activations were still displayed, which must be excluded by the other parameters. Despite the widely set limit (+/-5.5%), deletions in the center of Broca's area still occurred in one slice, marked by an arrow in Figure 23. The shape drawn by the excluded voxels resembles a vascular course suggesting the existence of a large blood vessel in this area. These voxels were probably excluded due to a high maximal signal strength, which is typical for blood vessels. This additional information helps the experienced physician to interpret this activation and the spatial extend of activations in this region.

In summary, this parameter is an important component of the filter that can identify high signal strength artifacts and blood vessel dependent effects while preserving language activations.

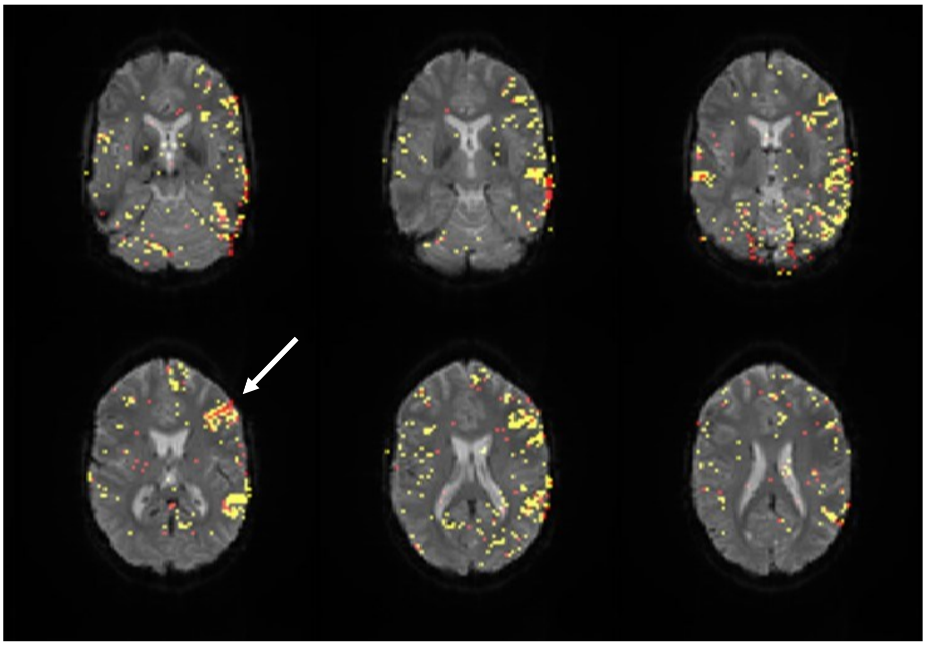


Figure 23: Effect of applying only the parameters max/min signal strength to a filter t-map. Red activations did not match the limit of the parameter and were excluded, yellow activations were accepted. In one slice, a part of Broca's area was deleted due to abnormally high signal strengths (arrow). This is most likely caused by a blood vessel located in this area, based on the shape of the deleted area.

3.6.2.2 Slope

The slope of a TP was calculated as its slope to the second next TP on the TC. The signal of the activations in the LAs was expected to rise quick and distinct at the start of the activation block and decrease quickly after the end of the activation block.

In Table 11 and Table 12 the averaged maximum and minimum slope values of entire regions are shown. These values were used as an orientation and to assess the interregional differences. The LAs, especially Broca's area (2.94%/-2.84%), showed higher values than most other functional areas (on average approx. +/-2.2%). Only the auditory cortex (3.2%/-3.06%) and activations in the frontal brain areas (3.4%/-3.22%) showed higher slope maxima, matching the parameter signal strength (3.6.2.1). A larger number of outliers was seen in activations of the "frontal artifact" (standard deviation >2), while the other regions showed consistent values (standard deviation <1).

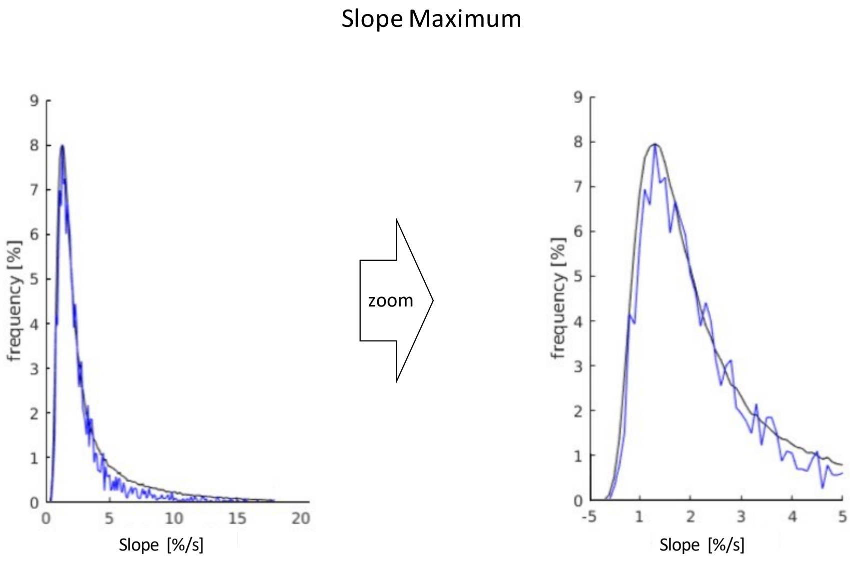
Table 11: Average slope maximum of different regions.

Region	Mean	Variance	Std.-Deviation
Broca	2,94	1,06	1,03
Wernicke	2,38	0,36	0,6
Language areas	2,67	0,79	0,89
SMA	1,86	0,57	0,76
Auditory Cortex	3,2	0,96	0,98
Visual Cortex	2,44	0,67	0,82
Frontal Artifact	3,4	5,53	2,35
Ventricular Artifact	2,54	0,62	0,79

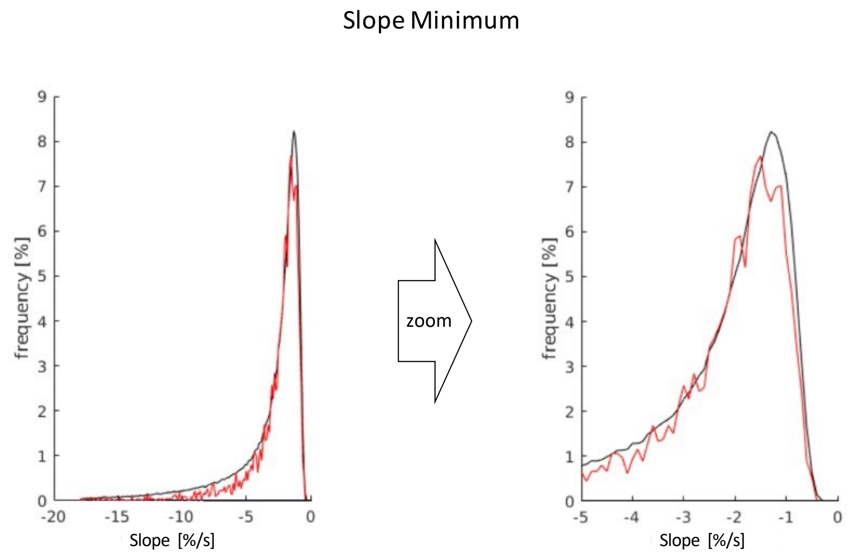
Table 12: Average slope minimum of different regions.

Region	Mean	Variance	Std.-Deviation
Broca	-2,84	1,01	1,01
Wernicke	-2,27	0,4	0,63
Language Areas	-2,57	0,79	0,89
SMA	-1,89	0,99	1,0
Auditory Cortex	-3,06	0,82	0,91
Visual Cortex	-2,38	0,58	0,76
Frontal Artifact	-3,22	4,16	2,04
Ventricular Artifact	-2,54	0,75	0,86

The evaluation of the single voxel analysis of the TCs is presented in Figure 24. The two histograms illustrated the maximum and minimum slope of language activations compared to all activations. The results were mostly congruent with the results of the parameter signal strength. Fewer language activations showed below/above +/-1%, compared to the activations of the entire map. The distribution of both groups was similar and ranged between the slope maxima/minima values of +/-1% and +/-6%. From a slope maximum/minimum value of +/-6%, the number of activations in the LAs decreased faster compared to the activations of the entire map. This led to the conclusion that most language activation had their slope maximum/minimum between +/-1% and +/- 6%, while the maximum/minimum slopes of activations of the entire map extended over a larger interval.



(**a**)



(b)

*Figure 24: Single voxel analysis of the parameter slope maximum **(a)** and slope minimum (b):*
On the left, the measured signal slope maxima/minima from 0% - 20%/ -20% - 0% is shown. On the right, a magnification of the histogram is displayed, in which only the relevant range of values from 0%-5%/ -5% - 0% are displayed.
(a) Histogram of the slope maximum values of language activations (blue) and all activations on the entire map (black).
The number of voxels (in percent) at different slope maxima is displayed (y-axis).
(b) Histogram of the slope minimum values of language activations (blue) and all activations on the entire map (black).
The number of voxels (in percent) at different slope minima is displayed (y-axis).

In contrast to the parameter signal strength, in this parameter unwanted activations and artifacts were mainly be excluded by the lower cutoff value. Language and other functional activations are expected to show a strong and rapid signal rise at the beginning of the activation block and should correspondingly show increased maximum slope values. Also, the expected fast decrease of the signal strength after end of the activation block was visible in the increased slope minimum of the language activations.

The limits for this parameter were set to 1% - 6% for the maximum and -1% - -6% for the minimum slope. The effect of applying this parameter on an activation map is shown in Figure 25. Implementing the parameter slope maximum/minimum with the defined range of values led to the exclusion of artifacts and unwanted activations in other functional regions. The LAs were completely displayed for the most part, even though, as seen in Figure 23, a group of activated voxels in Broca's area, were excluded, most likely caused by a blood vessel. Due to the increased signal strength in this area, the values of the maximum and minimum slope were also increased.

In summary, an improvement of the activation map was achieved by using this parameter and that the display of the LAs was not negatively affected.

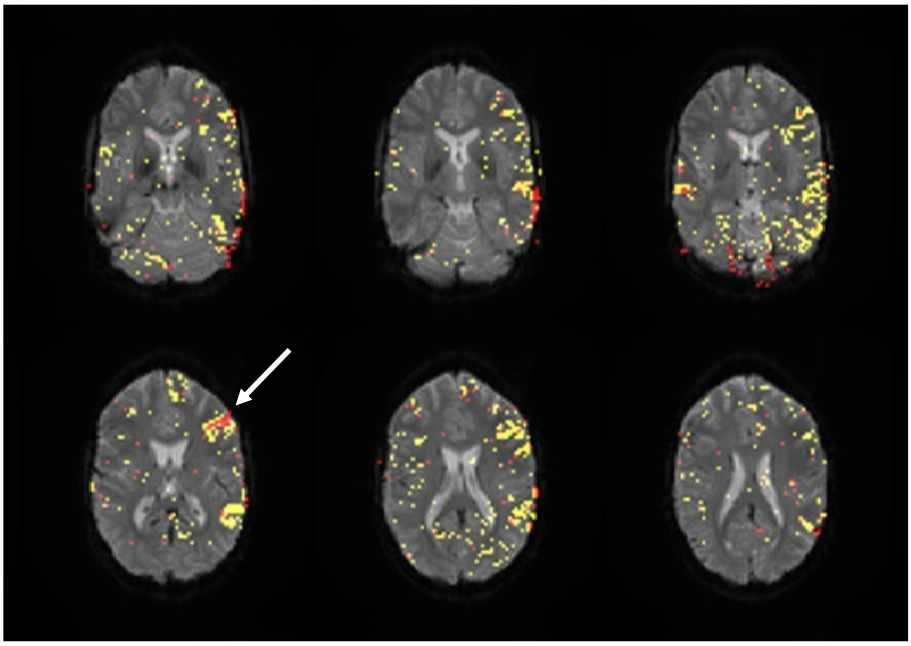


Figure 25: Effect of applying only the parameters max/min slope to a filter t-map. Red activations did not match the limits of the parameters and were excluded, yellow activations were accepted. A few artifacts and unwanted activations were excluded, the LAs mostly preserved. The (most likely) blood vessel-dependent activation in Broca's area (arrow) was also excluded by the parameter, as seen with the parameter signal strength.

3.6.2.3 Timing of maximum and minimum slope

When creating a t-map, all temporal information contained in the TCs is lost. By implementing the parameter timing of maximum and minimum slope, relevant temporal information was reintegrated in the creation of the activation maps. Temporal entities, such as the start and end of activation of stimulated regions, can be predicted, as the timing of important events of the experimental setup used in this study (fMRI in block design) were known. These predictions are based on the timing and duration of the blocks and the hemodynamic behavior of activated brain regions (HRF).

The activation-block started at 24s and ended at 48s. During this period, the LAs were stimulated by the language tasks. Based on the HRF, the time-to-onset (slope maximum) of the activated brain regions was expected to occur 6-8s after the onset of the stimulus. This increase of signal strength is followed by a plateau at an elevated signal, as the stimulus persisted over the entire 24s of the activation-block. The largest increase of signal (the maximum positive slope) was accordingly expected between 24s and 32s.

This prediction is consistent with the results of the single-voxel analysis shown in Figure 26. This histogram shows the number of voxels with a slope maximum at a given time. Activated voxels of the LAs are compared with activated voxels of the entire activation map. The majority language activations had their slope maximum between 27s (TP 9) and 33s (TP 11), with a peak at 27s (TP 9), 3s after the start of activation. This result is consistent with the expected behavior of the HRF, as the slope is measured between a TP and its second next TP, meaning that the slope of most language activations was the highest between 27s and 33s (TP 9 and TP 11). Also, an increased number of language activations had a slope maximum at 24s or 36s (TP 8 and TP 12).

The distribution of the timing of the slope maximum of activated voxels of the entire activation map was similar to the voxels of the LAs, showing an increased number of voxels with a slope maximum between 24s and 36s (time point 6-12,Figure 26). This can be explained by the fact that not only the LAs but also other functional areas, such as the auditory or visual cortex, were stimulated by the tasks. Less increase of signal was expected at the beginning of the activation block in the functional areas, due to the use of the activated pause, during which the functional regions were stimulated as well. This led in a less consistent occurrence of the slope maximum at the beginning of the activation-blocks in these regions and, accordingly, resulted in a lower proportion of voxels that have a slope maximum at the beginning of the activation-block (especially at 27s and 30s (TP 9 and TP 10)).

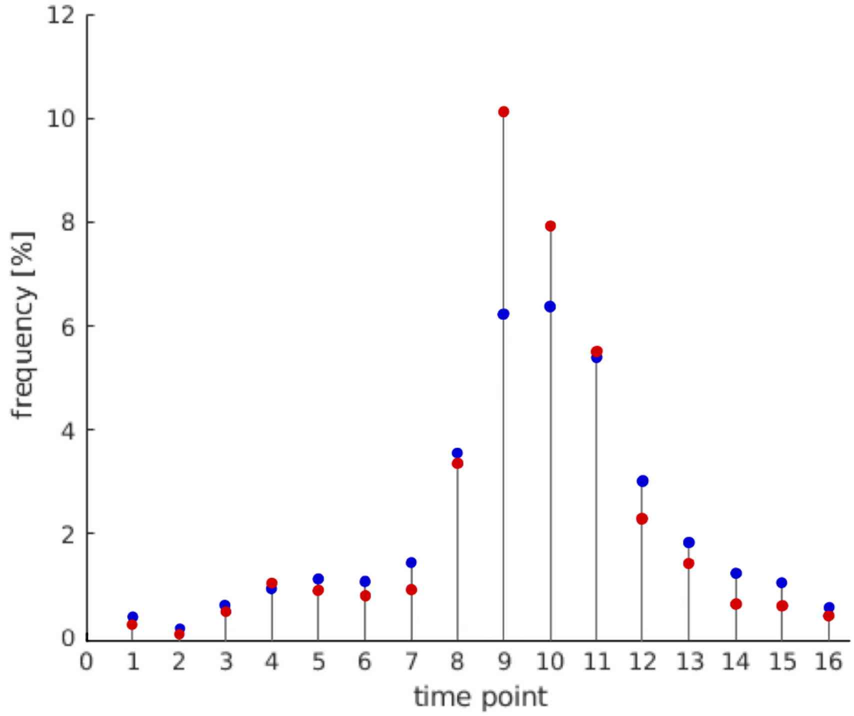


Figure 26: Single voxel analysis of the parameter time-to-onset (timing of slope maximum): Histogram comparing voxels in the LAs (red) to voxels on the entire map (blue).
Both groups showed an increased number of voxels with a slope maximum between timepoint 8 and 12 (24s and 36s), with a peak at time point 9, second 27, 3s after the start of the activation block. In the group of the LAs, this increase is more significant.

The time-to-end of activation (timing of the slope minimum) is expected shortly after the end of the stimulation, at the beginning of the following rest-block. This is consistent with results of the single voxel analysis shown in Figure 27. The slope minimum was at 3s and 9s (TP 1 and TP 3) for most voxels in the LAs. Also, at 48s (TP 16) a slightly increased number of voxels showed their slope minimum, even though this timepoint is still during the activation. However, since the slope describes a period of 6s, the slope minimum occurred between 48s and 9s (in the following rest block).

As before with the timing of the slope maximum, the activations of the entire map showed a similar pattern as the activations in the LAs, but with a proportionally smaller number of voxels at the peaks.

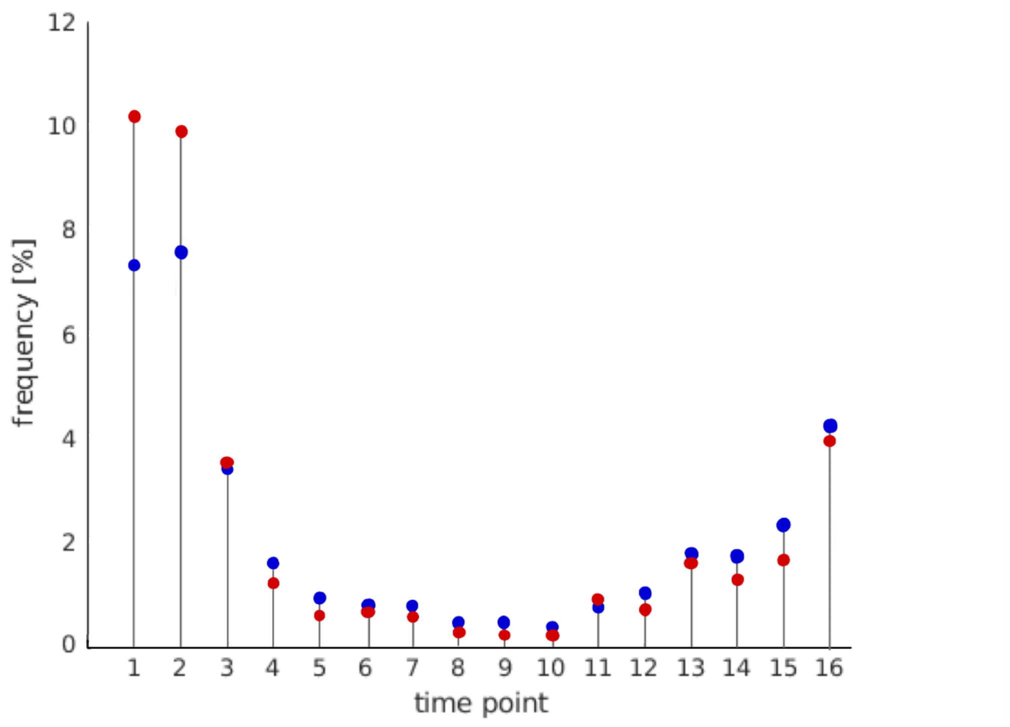


Figure 27: Single voxel analysis of the parameter time-to-end of activation (timing of slope minimum): Histogram comparing voxels in the LAs (red) to voxels on the entire map (blue).

Many voxels in the LAs show a slope minimum between 3s and 9s (TP 1 and TP 3) and 9s, at the beginning of the rest-block, but also at 48s (TP 16), at the end of the activation-block.

In Figure 28 the effect of the parameter time-to-onset/time-to-end on a filter t-map is shown. Only voxels that have a slope maximum between 24s and 32s and the slope minimum between 48s and 6s were allowed on the activation maps. All other voxels were excluded. These values were later used in the filter.

Using this parameter, many activations, that most likely occurred independent from our experiment, were deleted, clearing the map of artifacts in unwanted regions. Furthermore, activations near the LAs were also deleted, identifying those as artifacts and improving the spatial resolution. These artifacts are very difficult to identify and are usually assigned to the LAs, when not using this filter. Overall, more artifacts and apparent activations, including some activations of the functional regions and artifacts, were excluded, compared to the two previously described parameters. This improves the display and identifiability of the LAs.

As with the use of other parameters, some voxels were excluded in the region of a previously identified blood vessel. However, the area of the excluded voxels was more extensive and did not follow the typical shape of the vessel. This illustrates the impact of larger blood vessels on the activation signal and thereby the definition and the display of activations. By choosing wider limits (24s-36s (TP 8-TP 12) and 48s-9s (TP 16 – TP 9), values used for the final filter) the LAs were displayed more completely and the impact of blood vessels was reduced in most activation maps.

In summary, this parameter had a great impact on the creation of the t-maps and the display activated regions. It was the only language tasks that integrated temporal information and information of the experimental setup in the definition and display of activations. This allowed a precise exclusion of activations that are independent of the tasks and stimuli presented during the experiment.

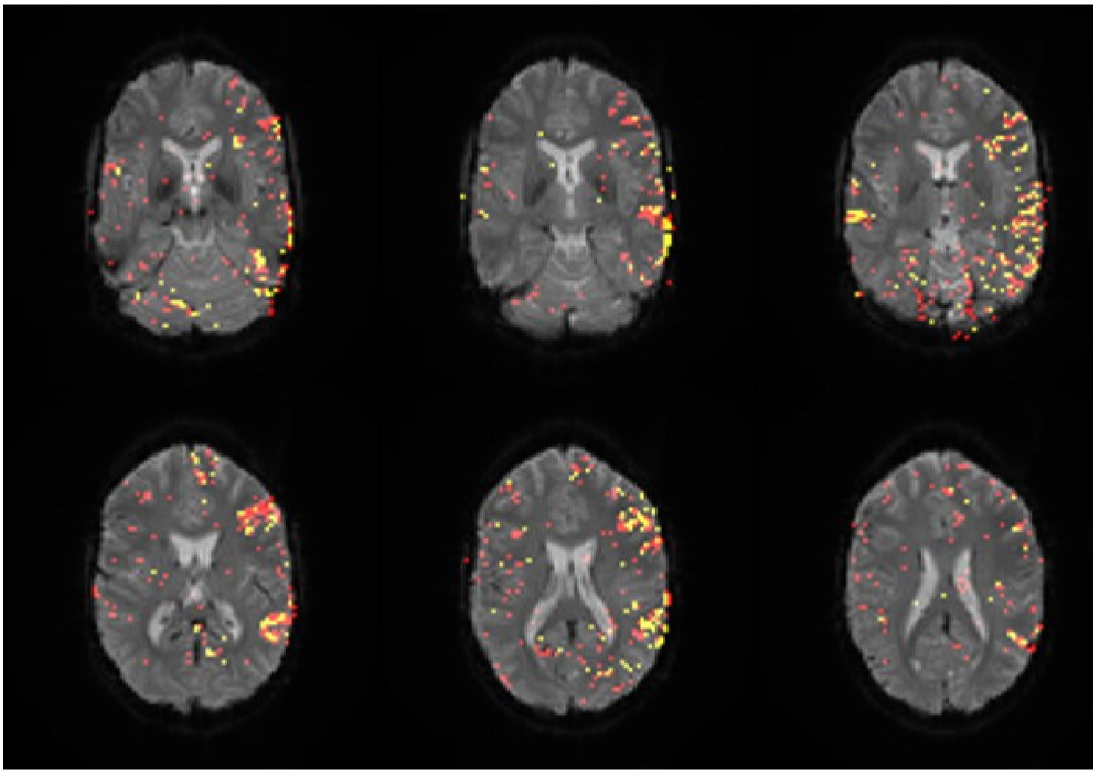
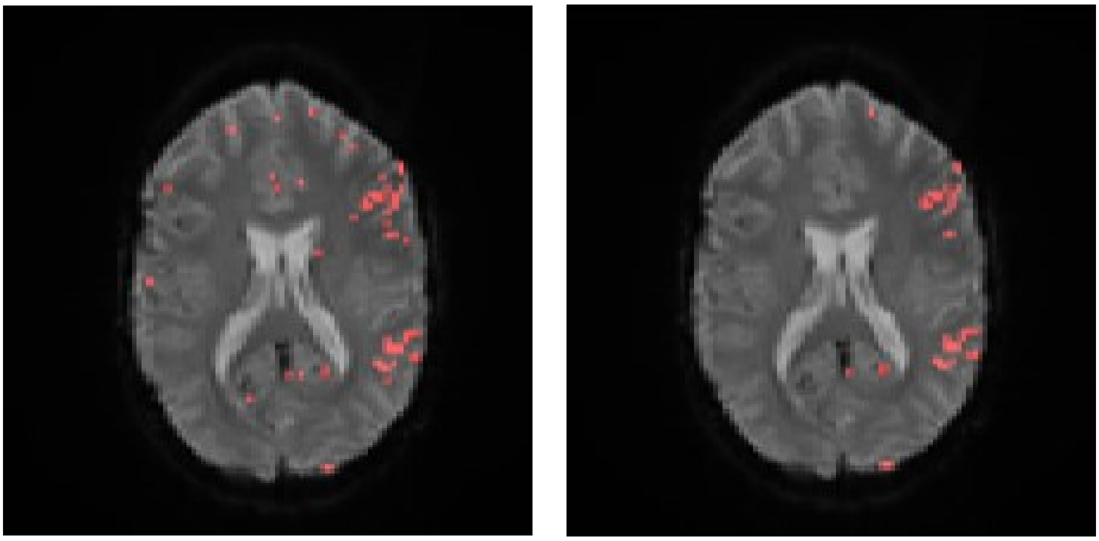


Figure 28: Effect of applying only the parameter time-to-onset/time-to-end of activation (slope maximum/minimum) to a filter t-map. Red activations did not match the limits of the parameters and were deleted, yellow activations were accepted.
Some voxels in Broca's area were excluded, due to the possible blood vessel running in this area.

3.6.3 Cluster limit

Small activations consisting of only a few voxels were displayed on most maps, since no smoothing was used for the creation of the activation maps. To remove these small activations and hereby improve the appearance of the map, a cluster limit was implemented. This cluster limit excluded all activations consisting of less than 3 contiguous voxels. Figure 29 demonstrates the effect of the cluster limit. Small activations distributed over the entire map were excluded by the cluster limit, while the larger activations, mainly in the LAs, remained complete.



(a) (b)

Figure 29: Comparison of a filter t-map with and without cluster limit. Measured activations are displayed red on the activation maps.
(a) Activation map without cluster limit. Small artifacts were displayed in frontal and central regions of the brain.
(b) Activation map with cluster limit. Most of the small artifacts are excluded.

3.6.4 Application of combined parameters

After the appropriate limits for each parameter had been defined (summarized in Table 13), the complete filter consisting of all parameters was applied to the t-maps. Compared to the simple t-maps, the use of the filter led to a general improvement of the activation maps, especially in the reduction of noise and in highlighting the LAs.

Table 13: Limits for each parameter.

Parameter	Bottom limit	Top limit
Max. signal strength	0.5%	5.5%
Min. signal strength	-0.5%	-5.5%
Max. slope	0.5%	6%
Min. slope	0.5%	-6%
Time-To-Onset	24s	36s
Time-To-End	48s	9s

In Figure 30, a comparison between a filter t-map and a simple t-map is shown. The use of the filter led to the exclusion of many apparent activations and artifacts on the whole map, as well as to a more limited display of activations in the LAs. The deletion of apparent activations (i.e., of the functional regions) facilitated the identification of LAs, while the exclusion of activations near the LAs helped to estimate the true spatial extent of the LAs.

Nevertheless, activations outside the LAs, especially in the functional regions, were still displayed frequently. This is due to the allowed limits of each parameter, which had to be defined with a leeway, because of intra- and interpersonal (between different measurements and subjects/patients) and the intraregional (between single voxels within the LAs) differences of the TCs of language activations. The limits of the parameters were chosen in a way that ensured the complete display of the language activations, as this was more important than the exclusion of the activations in the functional regions. The other reason for the persistence of these activations was that the functional regions, just like the LAs, were for the most part directly stimulated by the tasks, or by the presentation of the tasks, leading to similarities in the TCs.

Another similarity of LAs and functional regions was the consistency of the TCs. Therefore, the TC values of the language activations and the functional regions lay mostly within the parameter limits, which is another reason why the functional regions were often displayed on the filter t-maps. Artifacts, on the other hand, show less constant TCs with many outliers and can therefore be well distinguished from LAs. The number of outliers often led to an incomplete display of artifacts and other apparent activations, which was further enhanced using the cluster limit.

Overall, the use of the filter contributed to significant improvements in the display and identification of the LAs and the reduction of artifacts.

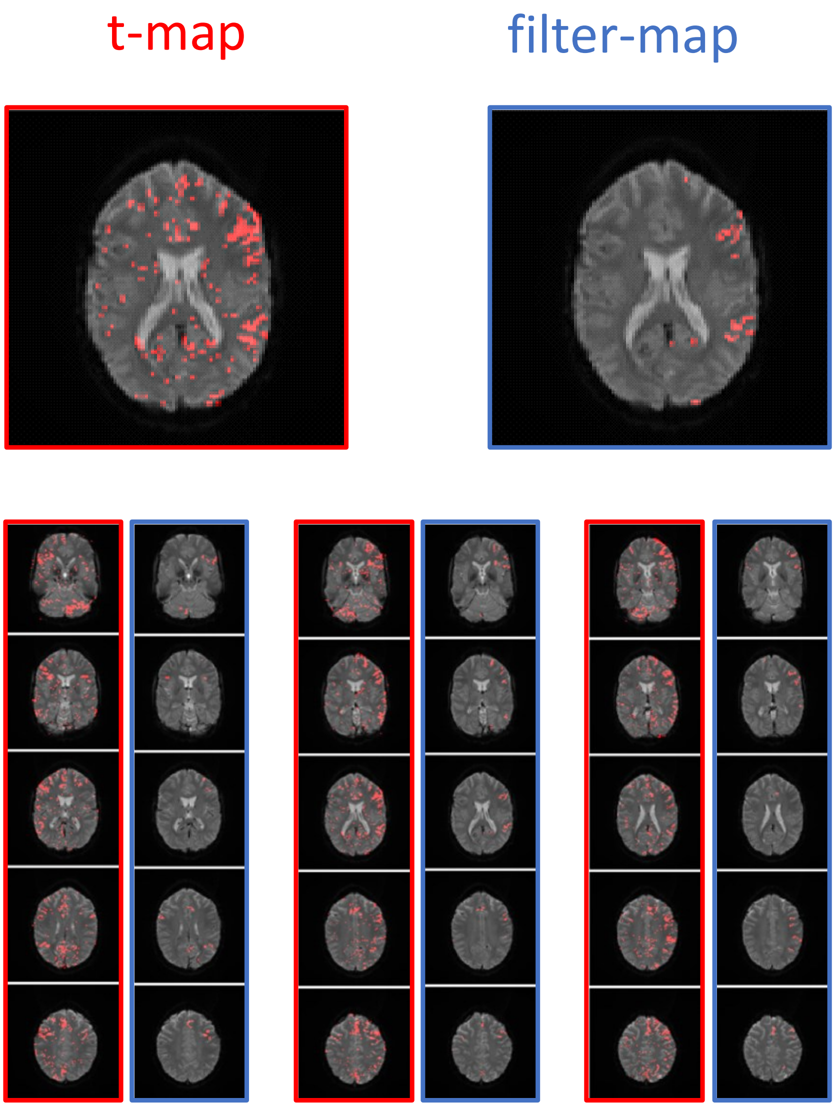


Figure 30: Effect of the filter t-map (blue) compared to simple t-map (red). The application of the filter results in the exclusion of many apparent activations and artifacts. In addition, the spatial extent of the LAs is more clearly separated from the surrounding artifacts, allowing an easier identification and spatial definition of the LAs. Measured activations are displayed red on the activation maps.

3.6.5 *Performance of the t-map*

The use of the filter resulted in a general improvement of the activation maps compared to the simple t-maps. Nevertheless, functional regions and some artifacts were still displayed on the filter t-maps, because the TCs of these activations were similar to the TCs of the LAs. To exclude these activations, the limits of the parameters would have to be defined more restrictively. However, this would have led to the exclusion of activated voxels in the LAs. Since the complete display of the LAs had the highest priority, the display of the functional regions and the remaining artifacts had to be accepted. A promising way to exclude these remaining artifacts is to establish additional parameters.

To evaluate the general performance of the filter t-maps and their relevance to everyday clinical practice, the filter t-maps were compared with GLM-maps (software: syngo.MR Neuro fMRI, Siemens Healthcare GmbH, Germany), that are frequently used in clinical practice. Three aspects were rated in this comparison: Display of Broca's area, display of Wernicke's area and number of artifacts. More detailed information on the rating scale used is described in the methodology. In Figure 31, a comparison of the filter t-map and the GLM-Map is shown, demonstrating improvements in the three rated aspects (display of Broca's area, display of Wernicke's area and signal & artifacts).

Figure 31 a) shows an improved display of Broca's area, as it is well separable from adjacent activations in the filter t-map. In addition, there was a strong reduction of artifacts, both near Broca's area and in the auditory cortex. Figure 31 b) shows an improvement of the display of Wernicke's area that was achieved by using the filter. There are fewer apparent activations adjacent to Wernicke's area, leading to a more precise definition of the boundaries of Wernicke's area. Figure 31 c) focused on the artifacts. There is a strong reduction of artifacts in both auditory and visual cortex on the filter t-map compared to the GLM-map.

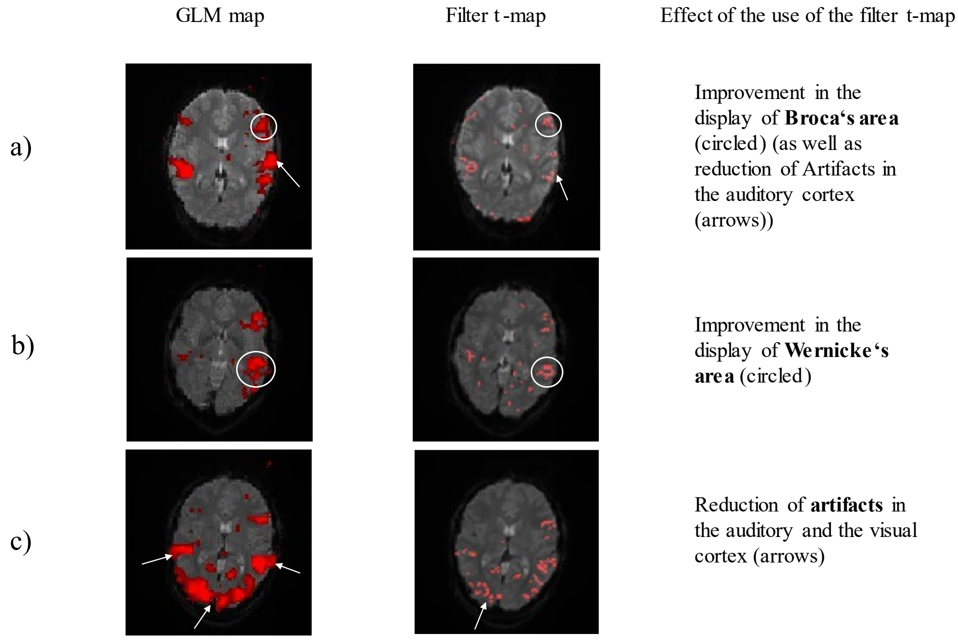


Figure 31: Exemplary comparison of GLM-maps and filter t-maps to demonstrate positive effects of using the filter t-map. Measured activations are displayed red on the activation maps.
a) Effect of the filter on the display of Broca's area.
b) Effect of the filter on the display of Wernicke's area.
c) Effect of the filter on the display of Artifacts.

The first part of the evaluation was done with the fMRI data of the subjects (n=48). The filter t-maps showed better results overall, especially in terms of displaying Wernicke's area and the number of displayed artifacts, as shown in Table 14. The display of Broca's area was similar in both mapping techniques. When looking at the smoothed GLM-maps, it was noticeable that the activated areas appeared larger and more coherent than in the filter t-maps, creating the impression of more consistent activations. This "more pleasant" display of the activations is due to the smoothing function, which was used for the creation of the GLM-maps, as it is frequently used in clinical practice. The problem with this is, however, that the smoothing tends to merge language activations with adjacent activations, making it impossible to draw a line between LAs and artifact. This observation is shown in Figure 32. On the GLM-map, Broca's area (blue circle) is connected to an adjacent activation (arrow), making it difficult to distinguish between these activations and to estimate the spatial extent of Broca's area. On the filter t-map Broca's area (blue circle) is displayed well separated from this activation (arrow). Wernicke's area (orange circle) is equally well displayed on both activation maps.

It became evident, that the display of the truly measured activations, often led to a smaller and less contiguous display of activated areas, which improved the separability of directly adjacent activations. Also, the estimation of the true location and extent of activations was improved by using the filter t-map, making the evaluation of the fMRI data more valid. Overall, the presented method of creating filter t-maps provided high quality of activation maps, by reducing the number and size of false activations, within LAs as well as artifacts.

Table 14: Evaluation of the comparison between the filter t-maps and the GLM-maps using the fMRI data of the subjects.

Score	**Improvement (+1)**	**Similar (0)**	**Deterioration (-1)**	**Total rating [points]**
Broca [number]	8	35	5	+ 3
Wernicke [number]	13	33	1	+ 12
Signal & Artifacts [number]	28	18	2	+ 26

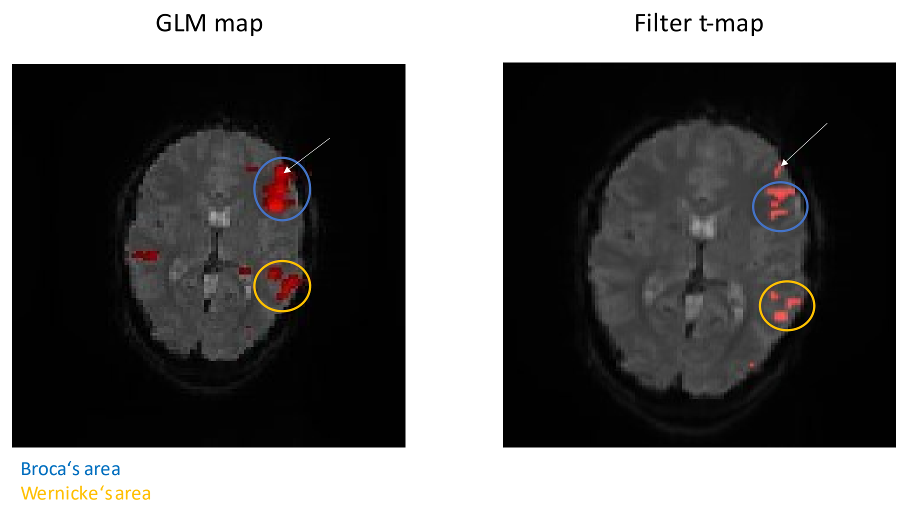


Figure 32: Comparison of the display of the LAs between a GLM-map (left) and a filter t-map (right). Broca's area was marked with a blue circle, adjacent activations were marked with an arrow, Wernicke's area was marked with an orange circle. Measured activations are displayed red on the activation maps.

In the second part of the evaluation, the fMRI data of the patients (patient-data) (n=31) was used to compare the GLM-map and the filter t-map (Table 15). The validity of the evaluation was increased by using this patient-data, compared to the first rating, as the fMRI data of the subject used in the first part of the evaluation were used as reference data for the creation of the filter. Also, the clinical suitability was verified, since the patient-data consisted of original fMRI images from real patients, independent of age, pathology and cognitive or physical condition of the patient, recorded in a clinical setting.

Table 15 shows the results of the evaluation of the activation maps created with the patient-data. The filter t-map was rated higher than the GLM-map, even leading to an improvement in the display of Broca's area. While the display of the LAs (Broca's and Wernicke's) was improved by the filter t-map in about one third of all rated maps, the signal and the reduction of artifacts was improved in about half of all rated maps.

Table 15: Evaluation of the comparison between the filter t-maps and the GLM-maps using the fMRI data of the patients.

Score	**Improvement (+1)**	**Similar (0)**	**Deterioration (-1)**	**Total rating [points]**
Broca [number]	9	20	2	+ 7
Wernicke [number]	9	22	0	+ 9
Signal & Artifacts [number]	15	15	1	+ 14

In summary, the filter t-map increased the quality of the activation maps and improved the display of the LAs. Due to the improved separability between language activations and adjacent artifacts, a more exact estimation of LA's spatial extent was enabled. All results were shown in the subject study as well as in the clinical setting.

4 Discussion

4.1 Language fMRI Examination

The first part of this thesis focused on improving the procedure of the language fMRI examination itself, pursuing two different approaches. On the one hand, settings were found that made the examination more robust and universally applicable. However, the improvements that were achieved by changing individual settings, especially the parameters, were often minor. This would also explain why many different approaches can be used in clinical practice and research (Black et al. 2017) without fundamental differences in the outcome.

The realization that the language tasks, stimuli, and study settings can be changed relatively freely, without much expected compromise in the results, was the basis for the second approach described in this thesis, that individualization of the study, depending on the patient's conditions and impairments, and specification of the study based on the medical question, could lead to an improvement in the results. The possibility to focus on the individual patient makes the examination available to a larger patient population.

4.1.1 Language tasks and stimulus

The comparison of the different language tasks showed that all presented tasks could be used for the creation high quality activation maps, with little clinically relevant differences. This was expected as all language tasks are commonly used by researchers and in hospitals (Black et al. 2017).

However, the language task “naming” stood out for both visual and auditory stimulus presentation, being rated best in the overall rating. Accordingly, this language task could be recommended for a standardized approach of the language fMRI examination. Standardized language fMRI examinations could provide more similarity between the results of different hospitals (Binder et al. 2011), leading to a better comparability of the examinations and data, and improving the quality management and research. The comparison of data is currently complicated, as variable approaches to the examination used worldwide (Black et al. 2017).

On the other hand, a different, more patient-oriented approach regarding the choice of the language tasks and experiment setup could be pursued. As demonstrated, the right choice of language task led to improved results, depending on the area of interest, as some language tasks tended to show better results in displaying Broca’s area (naming-task), whilst others performed better displaying Wernicke’s area (sentence-completion task) (Black et al. 2017). If a language fMRI examination in clinically necessary and the focus lies close to a certain area, the best fitting language task should be chosen to achieve the best results. Also, the right choice of the stimulus presentation led to an improvement of the maps. Visual stimulus presentation was suited better for the display of Wernicke’s area, because the activations of the visual cortex lay not as close to Wernicke's area (compared to the activations of the auditory cortex). Therefore, the language activations were not changed by or confused with the activations of the visual cortex. This approach was also described by (Szaflarski et al. 2017), where, depending on the clinical question, different recommendations were given for the fMRI examination.

However, the examination should not only be adjusted to the clinical question and the region of interest, but also to the individual condition of the patient. Nowadays, most hospitals offer a language fMRI examination with only visual language tasks (Black et al. 2017), leading to severe limitations or exclusion of visually impaired patients. The possibility of an acoustic stimulus presentation enables blind, severely ametropic, and even dyslectic patients to undergo this examination, without compromising the quality of the activation maps and with only a small amount of additional effort.

The cognitive abilities of the patients should also be considered, especially in examinations for preoperative mapping, as patients with brain tumors or other lesions affecting the brain often have cognitive deficits. Cognitively impaired patients might be overchallenged by some of the more complex language tasks, leading to a poor execution of the tasks and consequently to an insufficient display of the LAs. When examining cognitively impaired patients, the choice of a simpler language task (e.g., verb generation task) and an increased time per word for the presentation of the tasks, could help to obtain the best possible results.

4.1.2 Experiment setup

The evaluation of the duration of the scan and the number of periods showed that 6 periods (each consisting of one activation block and one rest block) is the best compromise between duration of the scan and quality of the activation map. Using less than 6 periods led to worse display of the LAs on the activation maps, while increasing the number of periods only led to stronger signal with more artifacts and not to a better display of the LAs. This stands in contrast to a recommendation in literature by (Soares et al. 2016): ”Generic recommendations include […] scan times as long as possible (the more trials the better, with several shorter runs preferred over one long run)[…]”.

The TC of the first period was frequently worse than in the following periods, supporting the theory of the existence of a training-effect, meaning that the execution of the tasks improves after a short adjustment period of the patient/subject. This was also described in literature by (Kay et al. 2008). However, it was not possible to turn this knowledge into an advantage for the creation of the activation maps, as the activation maps that were created with TC period 2-6, leading to worse results compared to the maps created with all six periods. The worse performance was attributed to the reduced number of periods, which led to a decrease of the quality of activation maps. Since these results were only obtained by analyzing the data of healthy subjects, it is possible that the training effect has a different impact when examining cognitively impaired patients. Based on the appearance of the TCs of most of the patient’s fMRI data, it can be assumed that training effect has greater impact on the activation maps, as the TC of the first period was mostly even worse compared the subject data. To be able to make final conclusions on this topic, further studies must be conducted, in which, for example, the language tasks is presented for seven periods, but the actual fMRI scan starts at period two. The extended duration of the examination should not affect the scan or the execution of the tasks negatively, as no sign of a decrease in concentration or performance was found in the TCs of the longer scans (12 periods) or mentioned by the subjects in the interview. Whether this also applies to patients depends on their cognitive and physical condition and must be decided on a case-by-case basis.

The activated pause led to large improvements in the creation activation maps, especially in reducing non-language activations in the functional regions originating from the stimulus and the tasks. The persistence of the stimulus in the rest block resulted in a reduction of activations in the functional regions activated by the stimulus. The motor component (fist) also provided automatic feedback about the functionality of the examination (headphones, monitor, etc.) and the patient's cooperation and mental condition (e.g., awake/asleep). Lastly, based on the subject’s feedback, the activated pause helped maintaining concentration and distracted the subjects, resulting in maintenance of focus and less mental wandering.

The use of the activated pause, or active control, was reported to be used in about in about 1/5th of fMRI examinations, sometimes matching the stimulus with the presentation of acoustic white noise or the display of visual patterns (Benjamin et al. 2018a; Szaflarski et al. 2017).

A detailed explanation to the patient/subjects prior to the fMRI examination about the language tasks and the procedure of the language fMRI examination in general is necessary for a successful fMRI examination (also mentioned by (Benjamin et al. 2017)). In the feedback interview, almost all subjects emphasized the importance of the explanation, as inadequate instructions led to high levels of uncertainty, irritation, and subjective poor performance of the tasks, visible in the quality of the activation maps.

Although the instruction and training of the patients is mandatory in most clinics, the quality and accuracy of explanations varies. In everyday clinical practice, due to a lack of time and personnel, explanations and training often are only carried out to a limited extent as it is assumed that patients are getting used to the tasks during the examination. This statement is consistent with the responses of physicians who were consulted on this issue.

Based on the TC analysis it was seen that the quality of the fMRI data of the patients, especially the first period, was worse compared the data of the subjects that had a detailed explanation before the scan. This may be to some extend due the differences in the cognitive condition between the healthy subjects and the patients, but it can be concluded that especially in the case of cognitively impaired patients, the instructions and explanation before the examination are important.

4.2 Analysis and evaluation of the language fMRI data

The second part of this study consisted of the analysis and evaluation of the data and the creation of activation maps. This part of the study has previously been published in Fetscher et al. 2022. Therefore, this part of the study overlaps to some extend with the published methods, results and conclusions drawn therefrom.

Simple t-maps were created and improved by a TC-dependent filter. With the filter and the three chosen parameters (signal strength, signal slope, time-to-onset, and time-to-end of activation) many artifacts and experiment independent activations were excluded, resulting in improved activation maps and a good display of the LAs. More difficult was the identification and differentiation of functional regions that also were stimulated directly by the language tasks or by the stimulus (such as the visual, auditory, and motor cortex). Although the interregional differences described in the literature (Lindquist et al. 2009; Lindquist und Wager 2007; Aguirre et al. 1998; Bielczyk et al. 2017) were found, they were too small and inconsistent to completely separate functional regions from the LAs. Reasons for these dissimilarities in significance and measurability of the interregional differences could be differences in the setup of the experiment (single-task activation vs. block-design), in the analysis of the signal (single HRF vs. HRF averaged over 6 periods), or in the single-voxel analysis approach used here.

Due to the single voxel approach, the limits of the parameters had to be defined wider to be able to guarantee the complete display of LAs, which was the highest priority in the creation of the filter t-map. The reason were intraregional differences of single voxels in activated regions, both LAs and other functional regions and observed interpersonal differences, as described in the literature (Benjamin et al. 2018a) and intra personal (between measurements) differences. Therefore, activations in the regions relevant for the reception and processing of the language tasks were still displayed on the filter t-maps, as these could not be sufficiently identified and excluded by the filter. These activations of the functional regions were also displayed on the GLM-maps, indicating that the reliable identification and exclusion of these activations is very difficult and only conditionally relevant in clinical practice. However, the filter helped to separate the LAs from its adjacent activations more clearly and thus improve the evaluation of the activation maps.

A possibility to exclude the activations in the functional regions would be to find parameters that are even more specific for language activations, or to adjust the activated pause during the rest block in a way that it completely resembles the language tasks except for the actual language production.

The different parameters and the possibility to adjust them individually provides a lot of additional information. The finding that the signal strength decreases from the center to the edge of a language activation was used to improve the estimation of the spatial extent of the language activations and with the parameter timing of maximum/minimum slope activations independent of the experiment were identified.

Furthermore, the course larger blood vessels in critical locations, such as Broca's area, could in some cases be traced on the filter t-maps. Information about the course of blood vessels is relevant because it can alter the signal of surrounding activations (Menon et al. 1993; Ugurbil 2016). As a result, an actual language activation may be altered because of proximity to a blood vessel in a way that it cannot be identified as a language activation. Conversely, the TC of a voxel in a region not activated by the language tasks may be altered to resemble a language activation leading to the voxel being displayed as an activation on the map. The information about the course of blood vessels can enhance the correct assessment of the spatial extent of activations and reveal the reason for deleted voxels in the LAs. When using other evaluation methods, this specific information is mostly lost, which, in the worst case, could lead to an incorrect identification and localization of the LAs.

4.3 Recommendations and improvements for the clinical use case

This thesis provided a holistic approach to improve the language fMRI examination, including the creation and appropriate choice of language tasks, the right experiment setup, and the creation of spatially precise and realistic filter t-maps, including their evaluation. There was possible improvement in almost every part of the language fMRI examination. The clinically most relevant improvements found in this study have been summarized in the following to achieve the greatest benefit for everyday clinical practice.

Prior to the language fMRI examination, it must be decided which language task and stimulus should be used. It seemed like a patient- and question-dependent approach should lead to the best results. All the used language tasks evaluated here, which were frequently used in everyday clinical practice (Black et al. 2017), can be exchanged without clinically relevant differences in the results. This gave the possibility to adjust the examination to the patient's conditions and impairments, making the examination more robust and available to a larger population (e.g., the use of acoustic language tasks for visually impaired patients). The language tasks should be adjusted to the medical question, as the selection of the appropriate tasks led to an improved display of the respective region of interest (Black et al. 2017; Gaillard et al. 2004; Połczyńska et al. 2017).

However, the patient-specific selection of language tasks requires more time and resources, which is often not practicable for clinics. This is the reason why many clinics offer only a few language tasks for the language fMRI examination (independent of medical question or the patient’s condition). Since it is difficult to define a "universal language task", there is no definite guideline and accordingly a variety of language tasks are used in the clinic (Black et al. 2017). Based on this study, the "naming" task was best suited for universal use, as it achieved overall the best results. To further improve the outcome of the examination, it is recommended to combine more than one language task (Gaillard et al. 2004).

When creating the language tasks, it was important to choose the rest block that fits the stimulus, or even the language task itself. In this study it was found that the design of the rest block has a strong impact on the results, in reducing the displayed artifacts, and on the concentration and comfort of the patient. The use of an activated pause using the same stimulus presentation as the language task, is recommended.

Another very important part of the examination was the training of patients. It led to a great improvement in fMRI data and activation maps with only little additional effort. A detailed explanation of the language tasks and the procedure of the examination should be performed prior every language fMRI examination, as it increases the patient's comfort and compliance and, as a result, can lead to a higher rate of successful measurements.

After the language fMRI examination has successfully been completed, the activation maps must be created and evaluated, including the identification and localization of the LAs. In everyday clinical practice, activation maps are mostly created using the software of major companies that are included in the software of the MRI scanner (e.g., Syngo.via, Siemens, Germany). These GLM-based activation maps are suitable for regular use in the clinic and generate solid and sufficiently accurate activation maps. However, to generate these maps, various (pre-)processing steps are applied to the fMRI data to correct inaccuracies, motion, and artifacts and to improve the readability and appearance of the map. These processing steps alter the fMRI data, which can result in the display of an inaccurate or even incorrect location of language activations. Especially the use of smoothing blurs the boundaries of activations, leading to the merging of different activations with the language activations, resulting in a loss of information (such as the location of blood vessels) and errors in the evaluation. This alternation of the data is very difficult to detect and can lead to misinterpretations, especially in spatially critical preoperative questions.

For this reason, a spatially close-to-reality and comprehensible approach for the creation of activation maps was generated, based on simple technology (t-map), little processing of the data and the avoiding of smoothing and other processing steps, altering the raw fMRI data as little as possible. With a single-voxel TC analysis based on language activation and experiment specific parameters, most artifacts were removed from the t-maps and the LAs were highlighted. Furthermore, additional information about the activations and the map based on the TCs, such as the course of blood vessels, were obtained. The technology used in the filter t-map and the processing of the data is comprehensible without excessive technical understanding, making it easier to assess the data correctly. In addition, the parameters, providing easy-to-understand information about the TCs, can be easily changed, exchanged, or extended to adjust the filter to more specific questions.

The clinical suitability of the filter t-map was demonstrated by using the data of patients and comparing them with GLM-map used in clinical practice created with the same data. The additional and multifaceted information obtained through the filter t-map is of great importance for the correct identification of LAs, especially when working with neurosurgical or neurooncological patients, where anatomical abnormalities are common.

It was possible to create activation maps with comprehensible technical methods, generating good activation maps without much alteration of the original fMRI data, providing a realistic display of the location of the language activations, and thereby allowing important clinical statements to be made. Nevertheless, definitive statements about the true location of the LAs can only be made with an intraoperative ECS examination. An ECS examination was not available due to the limited resources available for this study and, above all, due to the disproportionately high risk of this examination compared to the relevance of the expected findings.

5 Conclusion

In clinical practice, fMRI is mainly used to display functional brain regions such as the language areas. During the fMRI examination, patients are presented with cognitive tasks, here language tasks, during the MRI-scan, to activate the relevant brain regions. Activated regions can be identified and displayed based on the changing the blood flow od the activated brain regions (through the HRF), based on the BOLD effect.

Despite the widespread use of the fMRI examination, difficulties are regularly encountered in performing the scan and handling the tasks, processing the data or evaluating the activation maps and making spatially accurate statements about the location of the respective brain regions. This study therefore consisted of a holistic approach, analyzing every part of the language fMRI examination, to improve the examination and the results of the activation maps.

There was a prospective part, which consisted of a subject study, investigating the examination itself, i.e., the language tasks and the design of the experiment. The generally most robust task (naming task) was found, but equally relevant it was determined that there are only minor differences between the (frequently used) tasks. To improve the examination and the results on the activation maps, the tasks should rather be selected depending on the individual conditions of the patients and the medical question. More important points, which led to an improvement of the examination, were the use of an "activated pause", the correct number of activation and pause blocks and a sufficient training of the subjects/patients prior to the examination about the procedure of the examination and the tasks used.

Secondly, there was a retrospective part dealing with the processing of the fMRI data, i.e., the creation and analysis of the activation maps. Here, simple t-maps were created and enhanced using a filter. What was special about the filter t-maps was the omission of the (pre-)processing steps that are usually performed in clinical practice when using GLM maps. Since the true measured location of activations could be transferred unchanged to the activation maps, the relevant regions were spatially more realistically displayed, language activations were better separated from adjacent activations, and even additional artifacts and other non-language activations could be identified and excluded. The filter was based on a single-voxel analysis that evaluated the time courses of activations based on the parameters. Activations were thereby defined as true activation and displayed on the activation map or defined as false activation and excluded. The filter parameters (max/min. signal strength, max/min signal slope, and time at start/end of activation) were chosen based on interregional differences by comparing the averaged time courses of different activated regions. The allowed ranges of values of these parameters were defined by single-voxel analysis of the activations and elaboration of the interregional differences.

Finally, based on the answers found and after verifying the function of the filter t-maps, it was possible to provide a holistic recommendation for improving the language fMRI examination.

Zusammenfassung

Um funktionelle Hirnregionen wie das Sprachzentrum darzustellen, wird im klinischen Alltag hauptsächlich das fMRI verwendet. Zur Durchführung dieser fMRI-Untersuchung werden den Patienten während der Messung kognitive Aufgaben, wie in diesem Falle „Language Tasks“ präsentiert, um so die gesuchten Hirnregionen zu aktivieren. Durch Änderung der Durchblutung aktivierter Hirnregionen (HRF), basierend auf dem BOLD-Effekt, können so aktivierte Regionen identifiziert und dargestellt werden.

Trotz der weit verbreiteten Verwendung dieser Untersuchung treten regelmäßig Schwierigkeiten bei der Messung und der Bearbeitung der Aufgaben, der Verarbeitung der Daten oder der Auswertung der Aktivierungskarten und das Treffen räumlich exakter Aussagen zur Lage der gesuchten Hirnregionen. Deshalb wurde die (Sprach-) fMRI-Untersuchung in dieser Studie in ihrer Gesamtheit untersucht, um so, durch diesen ganzheitlichen Ansatz, die (Sprach-)fMRI-Untersuchung und die dabei erhaltenen Ergebnisse zu verbessern. Hierfür wurde die Studie in zwei Teile geteilt.

Es gab einen Prospektiven Teil, der mithilfe eine Probandenstudie die Untersuchung, also die Language Tasks und den Aufbau des Experiments, untersuchte. Hierbei konnte zum einen der allgemein robusteste Task (Naming Task) gefunden werden, es wurde allerdings auch festgestellt, dass sich die Unterschiede der (häufig verwendeten) Tasks in Grenzen halten und, um die Untersuchungsergebnisse zu verbessern, die Tasks eher abhängig von den individuellen Voraussetzungen der Patienten und der medizinischen Fragestellung ausgewählt werden sollten. Weitere wichtige Punkte, welche zu einer Verbesserung der Untersuchung führten, waren die Verwendung einer „Aktivierungen Pause“, die richtige Anzahl an Antivierungs- und Pause-Blöcken und eine suffiziente Aufklärung über den Ablauf der Untersuchung und die verwendeten Tasks.

Ergänzend gab es noch einen retrospektiven Teil, der sich mit der Verarbeitung der fMRI-Daten, also der Erstellung und Auswertung der Aktivierungskarten, befasste. Hier wurden einfache t-Maps erstellt und mithilfe eines Filters verbessert. Das spezielle an den Filter-t-Maps war das Weglassen der (Vor-)Verarbeitungsschritte, die bei der Verwendung von GLM-Maps normalerweise im klinischen Alltag durchgeführt werden. Da die tatsächlich gemessene Lage von Aktivierungen somit unverändert auf die Aktivierungskarten übertragen werden konnten, wurden die gesuchten Regionen räumlich realistischer

dargestellt werden, die Sprachaktivierungen besser von angrenzenden Aktivierungen getrennt werden und sogar Artefakte und weitere nicht-Sprach-Aktivierungen identifiziert und ausgeschlossen werden. Der Filter basierte auf einer Single-Voxel-Analyse, welche die Zeitverläufe von Aktivierungen auf Grundlage der Parameter bewertete und als echte Aktivierung definierte und somit erlaubte, oder als falsch definierte und ausschloss. Die Filterparameter (Max./min. Signalstärke, max./min Signalsteigung und Zeit bei Start/Ende der Aktivierung) wurden, basierend auf interregionalen Unterschieden, durch den Vergleich der Zeitverläufe verschiedener aktivierter Regionen ausgewählt. Die erlaubten Wertebereiche dieser Parameter wurden durch eine Single-Voxel-Analyse der Aktivierungen und die Ausarbeitung der interregionalen Unterschiede definiert.

Auf Grundlage der gefundenen Antworten und nach Verifizierung der Funktion der Filter-t-Maps konnte abschließend eine ganzheitliche Empfehlung zur Verbesserung der Sprach-fMRI-Untersuchung ausgearbeitet werden.

Figure Index

Table Index

References

Aguirre GK, Zarahn E, D'esposito M (1998) The variability of human, BOLD hemodynamic responses. NeuroImage 8 (4), 360–369.

Benjamin CF, Dhingra I, Li AX, Blumenfeld H, Alkawadri R, Bickel S, Helmstaedter C, Meletti S, Bronen RA, Warfield SK, Peters JM, Reutens D, Połczyńska MM, Hirsch LJ, Spencer DD (2018a) Presurgical language fMRI: Technical practices in epilepsy surgical planning. Human brain mapping 39 (10), 4032–4042.

Benjamin CF, Dhingra I, Li AX, Blumenfeld H, Alkawadri R, Bickel S, Helmstaedter C, Meletti S, Bronen RA, Warfield SK, Peters JM, Reutens D, Połczyńska MM, Hirsch LJ, Spencer DD (2018b) Presurgical language fMRI: Technical practices in epilepsy surgical planning. Human brain mapping 39 (10), 4032–4042.

Benjamin CF, Walshaw PD, Hale K, Gaillard WD, Baxter LC, Berl MM/Polczynska M, Noble S, Alkawadri R, Hirsch LJ, Constable RT, Bookheimer SY (2017) Presurgical language fMRI: Mapping of six critical regions. Human brain mapping 38 (8), 4239–4255.

Benson DF, Ardila A (1996) Aphasia. A clinical perspective. New York, Oxford University Press.

Bielczyk NZ, Llera A, Buitelaar JK, Glennon JC, Beckmann CF (2017) The impact of hemodynamic variability and signal mixing on the identifiability of effective connectivity structures in BOLD fMRI. Brain and behavior 7 (8), e00777.

Binder JR, Desai RH, Graves WW, Conant LL (2009) Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. Cerebral cortex (New York, N.Y. : 1991) 19 (12), 2767–2796.

Binder JR, Gross WL, Allendorfer JB, Bonilha L, Chapin J, Edwards JC, Grabowski TJ, Langfitt JT, Loring DW, Lowe MJ, Koenig K, Morgan PS, Ojemann JG, Rorden C, Szaflarski JP, Tivarus ME, Weaver KE (2011) Mapping anterior temporal lobe language areas with fMRI: a multicenter normative study. NeuroImage 54 (2), 1465–1475.

Black DF, Vachha B, Mian A, Faro SH, Maheshwari M, Sair HI, Petrella JR, Pillai JJ, Welker K (2017) American Society of Functional Neuroradiology-Recommended fMRI Paradigm Algorithms for Presurgical Language Assessment. AJNR. American journal of neuroradiology 38 (10), E65-E73.

Brownsett SLE, Ramajoo K, Copland D, McMahon KL, Robinson G, Drummond K, Jeffree RL, Olson S, Ong B, Zubicaray G (2019) Language deficits following dominant hemisphere tumour resection are significantly underestimated by syndrome-based aphasia assessments. Aphasiology 33 (10), 1163–1181.

Burnstine TH, Lesser RP, Hart J, Uematsu S, Zinreich SJ, Krauss GL, Fisher RS, Vining EP, Gordon B (1990) Characterization of the basal temporal language area in patients with left temporal lobe epilepsy. Neurology 40 (6), 966–970.

Caputi F, Spaziante R, de Divitiis E, Nashold BS Jr. (1995) Luigi Rolando and his pioneering efforts to relate structure to function in the nervous system. Journal of Neurosurgery 83, 5, 933-937.

Catani M, Mesulam M (2008) The arcuate fasciculus and the disconnection theme in language and aphasia: history and current state. Cortex; a journal devoted to the study of the nervous system and behavior 44 (8), 953–961.

Chen G, Saad ZS, Adleman NE, Leibenluft E, Cox RW (2015) Detecting the subtle shape differences in hemodynamic responses at the group level. Frontiers in neuroscience 9, 375.

Chen, JE, Glover GH (2015) Functional Magnetic Resonance Imaging Methods. Neuropsychology review 25 (3), 289–313.

Coale A. J. (1984). The demographic transition. Pakistan development review, 23(4), 531–552.

Dronkers NF, Plaisant O, Iba-Zizen MT, Cabanis EA (2007) Paul Broca's historic cases: high resolution MR imaging of the brains of Leborgne and Lelong. Brain : a journal of neurology 130 (Pt 5), 1432–1441.

D'Souza MM, Sharma R, Tripathi M, Panwar P, Jaimini A, Mondal A (2011) Novel positron emission tomography radiotracers in brain tumor imaging. Indian Journal of Radiology and Imaging 21 (3), 202–208.

Duffau H (2016) Long-term outcomes after supratotal resection of diffuse low-grade gliomas: a consecutive series with 11-year follow-up. Acta neurochirurgica 158 (1), 51–58.

Eliassen JC, Boespflug EL, Lamy M, Allendorfer J, Chu WJ, Szaflarski JP (2008) Brain-mapping techniques for evaluating poststroke recovery and rehabilitation: a review. Topics in stroke rehabilitation 15 (5), 427–450.

Evans AC (2006) The NIH MRI study of normal brain development. NeuroImage 30 (1), 184–202.

Fetscher L, Batra M, Klose U (2022) Improved localization of language areas using single voxel signal analysis of unprocessed fMRI data. In: Front. Radio 2, Artikel 997330.

Fowler JS, Volkow ND, Kassed CA, Chang L (2007) Imaging the addicted human brain. Science & Practice Perspectives 3 (2), 4–16.

Friedman L, Glover GH (2006) Report on a multicenter fMRI quality assurance protocol. Journal of magnetic resonance imaging : JMRI 23 (6), 827–839.

Gaillard WD, Balsamo L, Xu B, Grandin CB, Braniecki SH, Papero PH, Weinstein S, Conry J, Pearl PL, Sachs B, Sato S, Jabbari B, Vezina LG, Frattali C, Theodore WH (2002) Language dominance in partial epilepsy patients identified with an fMRI reading task. Neurology 59 (2), 256–265.

Gaillard WD, Balsamo L, Xu B, McKinney C, Papero PH, Weinstein S, Conry J, Pearl PL, Sachs B, Sato S, Vezina LG, Frattali C, Theodore WH (2004) fMRI language task panel improves determination of language dominance. Neurology 63 (8), 1403–1408.

Garrett MC, Pouratian N, Niau LM (2012) Use of language mapping to aid in resection of gliomas in eloquent brain regions. Neurosurgery clinics of North America 23 (3), 497–506.

Goman AM, Lin FR (2016) Prevalence of Hearing Loss by Severity in the United States. American journal of public health 106 (10), 1820–1822.

Herbst KG, Humphrey C (1981) Prevalence of hearing impairment in the elderly living at home. The Journal of the Royal College of General Practitioners 31 (224), 155–160.

Hiroshima S, Anei R, Murakami N, Kamada K (2014) Functional localization of the supplementary motor area. Neurologia medico-chirurgica 54 (7), 511–520.

Hong N, Yoo H, Gwak HS, Shin SH, Lee SH (2013) Outcome of surgical resection of symptomatic cerebral lesions in non-small cell lung cancer patients with multiple brain metastases. Brain Tumor Research and Treatment 1 (2), 64–70.

Javed K, Reddy VM, Das J, Wroten M, 2022. Neuroanatomy, Wernicke Area, StatPearls Publishing, StatPearls, Treasure Island (FL).

Kauhanen ML, Korpelainen JT, Hiltunen P, Määttä R, Mononen H, Brusin E, Sotaniemi KA, Myllylä VV (2000) Aphasia, depression, and non-verbal cognitive impairment in ischaemic stroke. Cerebrovascular diseases (Basel, Switzerland) 10 (6), 455–461.

Kay KN, David SV, Prenger RJ, Hansen KA, Gallant JL (2008) Modeling low-frequency fluctuation and hemodynamic response timecourse in event-related fMRI. Human brain mapping 29 (2), 142–156.

Klinke R, Pape HC/Silbernagl S, 2005, Physiologie. 5. Aufl., Thieme, Stuttgart.

Krauss GL, Fisher R, Plate C, Hart J, Uematsu S, Gordon B, Lesser RP (1996) Cognitive effects of resecting basal temporal language areas. Epilepsia 37 (5), 476–483.

Kundu B, Penwarden A, Wood JM, Gallagher TA, Andreoli MJ, Voss J, Meier T, Nair VA, Kuo JS, Field AS, Moritz C, Meyerand ME, Prabhakaran V (2013) Association of functional magnetic resonance imaging indices with postoperative language outcomes in patients with primary brain tumors. Neurosurgical focus 34 (4), E6.

Lacroix M, Abi-Said D, Fourney DR, Gokaslan ZL, Shi W, DeMonte F, Lang FF, McCutcheon IE, Hassenbusch SJ, Holland E, Hess K, Michael C, Miller D, Sawaya R (2001) A multivariate analysis of 416 patients with glioblastoma multiforme: prognosis, extent of resection, and survival. Journal of neurosurgery 95 (2), 190–198.

Laska AC, Hellblom A, Murray V, Kahan T, Arbin M (2001) Aphasia in acute stroke and relation to outcome. Journal of Internal Medicine 249 (5), 413–422.

Laures-Gore JS, Dotson VM, Belagaje S (2020) Depression in Poststroke Aphasia. American journal of speech-language pathology 29 (4), 1798–1810.

Lemée JM, Berro DH, Bernard F, Chinier E, Leiber LM, Menei P, Ter Minassian A (2019) Resting-state functional magnetic resonance imaging versus task-based activity for language mapping and correlation with perioperative cortical mapping. Brain and behavior 9 (10), e01362.

Lindquist MA, Meng Loh J, Atlas LY, Wager TD (2009) Modeling the hemodynamic response function in fMRI: efficiency, bias and mis-modeling. NeuroImage 45 (1 Suppl), S187-98.

Lindquist MA, Wager TD (2007) Validity and power in hemodynamic response modeling: a comparison study and a new approach. Human brain mapping 28 (8), 764–784.

Maberley DAL, Hollands H, Chuo J, Tam G, Konkal J, Roesch M, Veselinovic A, Witzigmann M, Bassett K (2006) The prevalence of low vision and blindness in Canada. Eye (London, England) 20 (3), 341–346.

Menon RS, Ogawa S, Tank DW, Uğurbil K (1993) Tesla gradient recalled echo characteristics of photic stimulation-induced signal changes in the human primary visual cortex. Magnetic resonance in medicine 30 (3), 380–386.

Mueller SG, Weiner MW, Thal LJ, Petersen RC, Jack C, Jagust W, Trojanowski JQ, Toga AW, Beckett L (2005) The Alzheimer's disease neuroimaging initiative. Neuroimaging clinics of North America 15 (4), 869-77, xi-xii.

Ogawa S, Lee TM (1990) Magnetic resonance imaging of blood vessels at high fields: in vivo and in vitro measurements and image simulation. Magnetic resonance in medicine 16 (1), 9–18.

Ogawa S, Lee TM. Kay AR, Tank DW (1990) Brain magnetic resonance imaging with contrast dependent on blood oxygenation. Proceedings of the National Academy of Sciences of the United States of America 87 (24), 9868–9872.

Ogawa S (2012) Finding the BOLD effect in brain images. NeuroImage 62 (2), 608–609.

Paolucci S, Antonucci G, Pratesi L, Traballesi M, Lubich S, Grasso MG (1998) Functional outcome in stroke inpatient rehabilitation: predicting no, low and high response patients. Cerebrovascular diseases (Basel, Switzerland) 8 (4), 228–234.

Pascolini D, Mariotti SP (2012) Global estimates of visual impairment: 2010. The British journal of ophthalmology 96 (5), 614–618.

Perthen JE, Lansing AE, Liau J, Liu TT, Buxton RB (2008) Caffeine-induced uncoupling of cerebral blood flow and oxygen metabolism: a calibrated BOLD fMRI study. NeuroImage 40 (1), 237–247.

Petrella JR, Shah LM, Harris KM, Friedman AH, George TM, Sampson JH, Pekala JS, Voyvodic JT (2006) Preoperative functional MR imaging localization of language and motor areas: effect on therapeutic decision making in patients with potentially resectable brain tumors. Radiology 240 (3), 793–802.

Połczyńska M, Japardi K, Curtiss S, Moody T, Benjamin C, Cho A, Vigil C, Kuhn T, Jones M, Bookheimer S (2017) Improving language mapping in clinical fMRI through assessment of grammar. NeuroImage. Clinical 15, 415–427.

Raichle ME, Martin WR, Herscovitch P, Mintun MA, Markham J (1983) Brain blood flow measured with intravenous H2(15)O. II. Implementation and validation. Journal of nuclear medicine : official publication, Society of Nuclear Medicine 24 (9), 790–798.

Ritaccio AL, Brunner P, Schalk G (2018) Electrical Stimulation Mapping of the Brain: Basic Principles and Emerging Alternatives. Journal of clinical neurophysiology : official publication of the American Electroencephalographic Society 35 (2), 86–97.

Roux FE, Draper L, Köpke B, Démonet JF (2010) Who actually read Exner? Returning to the source of the frontal "writing centre" hypothesis. Cortex; a journal devoted to the study of the nervous system and behavior 46 (9), 1204–1210.

Sanai N, Mirzadeh Z, Berger MS (2008) Functional outcome after language mapping for glioma resection. The New England journal of medicine 358 (1), 18–27.

Sawaya R (1999) Extent of resection in malignant gliomas: a critical summary. Journal of neuro-oncology 42 (3), 303–305.

Schäffler L, Lüders HO, Morris HH, Wyllie E (1994) Anatomic distribution of cortical language sites in the basal temporal language area in patients with left temporal lobe epilepsy. Epilepsia 35 (3), 525–528.

Soares JM, Magalhães R, Moreira PS, Sousa A, Ganz E, Sampaio A, Alves V, Marques P, Sousa N (2016) A Hitchhiker's Guide to Functional Magnetic Resonance Imaging. Frontiers in neuroscience 10, 515.

Swatzyna RJ, Kozlowski GP, Tarnow JD (2015) Pharmaco-EEG: A Study of Individualized Medicine in Clinical Practice. Clinical EEG and Neuroscience 46 (3), 192–196. https://doi.org/10.1177/1550059414556120.

Szaflarski JP, Gloss D, Binder JR, Gaillard WD, Golby AJ, Holland SK, Ojemann J, Spencer DC, Swanson SJ, French JA, Theodore WH (2017) Practice guideline summary: Use of fMRI in the presurgical evaluation of patients with epilepsy: Report of the Guideline Development, Dissemination, and Implementation Subcommittee of the American Academy of Neurology. Neurology 88 (4), 395–402.

Tang Y, Wang X, Wang J, Huang W, Gao Y, Luo Y, Lu Y (2015) Prevalence and Causes of Visual Impairment in a Chinese Adult Population: The Taizhou Eye Study. Ophthalmology 122 (7), 1480–1488.

Ugurbil K (2016) What is feasible with imaging human brain function and connectivity using functional magnetic resonance imaging. Philosophical transactions of the Royal Society of London. Series B, Biological sciences 371 (1705).

Wada J, Rasmussen T (2007) Intracarotid injection of sodium amytal for the lateralization of cerebral speech dominance. 1960. Journal of neurosurgery 106 (6), 1117–1133.

Appendix

Erklärungen zum Eigenteil

Die vorliegende Dissertation wurde in der Abteilung Diagnostische und Interventionelle Neuroradiologie des Universitätsklinikums Tübingen unter Betreuung von Herrn Prof. Dr. rer. nat. Uwe Klose und, von ärztlicher Seite aus, Frau Dr. Marion Batra durchgeführt.

Die Konzeption der Studie erfolgte durch Herrn Prof. Dr. rer. nat. Uwe Klose.

Nach meiner Einarbeitung durch Herrn Prof. Dr. rer. nat. Uwe Klose wurden die MRT-Messungen sämtlicher Probanden eigenständig von mir durchgeführt. Frau Dr. med. Marion Batra stand bei klinischen und anatomischen und radiologischen Fragen zur Verfügung. Die Datenauswertung erfolgte durch mich mit Unterstützung durch Herrn Prof. Dr. rer. nat. Uwe Klose. Die statistische Auswertung erfolgte eigenständig durch mich.

Die aus dieser Arbeit hervorgegangene Publikation wurde von mir verfasst und von Herrn Prof. Dr. rer. nat. Uwe Klose, Frau Dr. med. Marion Batra gelesen und kommentiert.

Dieses Manuskript dieser Dissertation wurde von Herrn Prof. Dr. rer. nat. Uwe Klose korrigiert.

Ich versichere, das Manuskript selbständig verfasst zu haben und keine weiteren als die von mir angegebenen Quellen verwendet zu haben.

Veröffentlichungen

Teile der vorliegenden Arbeit wurden publiziert in:

Fetscher, Leonard; Batra, Marion; Klose, Uwe (2022): Improved localization of language areas using single voxel signal analysis of unprocessed fMRI data. In: Front. Radio 2, Artikel 997330. DOI: 10.3389/fradi.2022.997330.

Danksagung

Zunächst möchte ich mich bei Prof. Dr. rer. Nat. Uwe Klose für die gute, geduldige und jederzeit verfügbare Unterstützung während der gesamten Arbeit bedanken.

Weiter möchte ich mich bei Dr. Marion Batra für die Unterstützung und das Teilen neuroradiologischen Wissens bedanken.

Ein großer Dank geht außerdem an alle Probanden, die sich Zeit für diese Studie genommen haben und ohne Beschwerden die benötigte Zeit in der engen MRT-Röhre verbracht haben. Ohne euch wäre das nicht möglich gewesen.

Danken möchte ich auch dem IZKF-Promotionskolleg der medizinischen Fakultät Tübingen, welches mir den Einstieg in diese Arbeit durch das Teilen von Wissen und mit finanzieller Förderung erleichtert hat.

Zu guter Letzt danke ich auch meiner Familie, meinen Eltern, meinen Schwestern und meiner Oma, welche mir bei dieser Arbeit mit größtmöglicher Unterstützung und bedingungsloser Liebe geholfen haben.