

# The effect of fMRI task combinations on determining the hemispheric dominance of language functions

Eini Niskanen · Mervi Könönen · Ville Villberg · Mikko Nissi · Perttu Ranta-aho ·  
Laura Säisänen · Pasi Karjalainen · Marja Äikiä · Reetta Kälviäinen · Esa Mervaala ·  
Ritva Vanninen

Received: 1 May 2011 / Accepted: 7 September 2011 / Published online: 20 September 2011  
© The Author(s) 2011. This article is published with open access at Springerlink.com

## Abstract

**Introduction** The purpose of this study is to establish the most suitable combination of functional magnetic resonance imaging (fMRI) language tasks for clinical use in determining language dominance and to define the variability in laterality index (LI) and activation power between different combinations of language tasks.

E. Niskanen · M. Nissi · P. Ranta-aho · P. Karjalainen  
Department of Applied Physics, University of Eastern Finland,  
Kuopio, Finland

E. Niskanen (✉) · M. Könönen · R. Vanninen  
Department of Clinical Radiology, Kuopio University Hospital,  
Kuopio, Finland  
e-mail: Eini.Niskanen@kuh.fi

M. Könönen · L. Säisänen · E. Mervaala  
Department of Clinical Neurophysiology,  
Kuopio University Hospital,  
Kuopio, Finland

V. Villberg · M. Äikiä · R. Kälviäinen  
Department of Neurology, Kuopio University Hospital,  
Kuopio, Finland

R. Kälviäinen  
Institute of Clinical Medicine, Neurology,  
University of Eastern Finland,  
Kuopio, Finland

L. Säisänen · E. Mervaala  
Institute of Clinical Medicine, Clinical Neurophysiology,  
University of Eastern Finland,  
Kuopio, Finland

R. Vanninen  
Institute of Clinical Medicine, Clinical Radiology,  
University of Eastern Finland,  
Kuopio, Finland

**Methods** Activation patterns of different fMRI analyses of five language tasks (word generation, responsive naming, letter task, sentence comprehension, and word pair) were defined for 20 healthy volunteers (16 right-handed). LIs and sums of  $T$  values were calculated for each task separately and for four combinations of tasks in predefined regions of interest. Variability in terms of activation power and lateralization was defined in each analysis. In addition, the visual assessment of lateralization of language functions based on the individual fMRI activation maps was conducted by an experienced neuroradiologist.

**Results** A combination analysis of word generation, responsive naming, and sentence comprehension was the most suitable in terms of activation power, robustness to detect essential language areas, and scanning time. In general, combination analyses of the tasks provided higher overall activation levels than single tasks and reduced the number of outlier voxels disturbing the calculation of LI.

**Conclusions** A combination of auditory and visually presented tasks that activate different aspects of language functions with sufficient activation power may be a useful task battery for determining language dominance in patients.

**Keywords** Language lateralization · Functional magnetic resonance imaging · fMRI

## Introduction

Determining hemispheric language dominance is essential during the preoperative planning of neurosurgical resection to minimize the risk of postoperative verbal

deficits. The clinical gold standard for lateralizing language function in neurosurgical planning is the Wada test (intracarotid amobarbital test) [1]. However, due to its invasive nature and the effect of ultrashort lasting barbiturates, conclusions from the Wada test are sometimes uncertain. Therefore, alternative non-invasive methods, such as functional magnetic resonance imaging (fMRI), have been utilized. Selecting the optimal combination of tasks for clinical language fMRI and the interpretation of the results are, however, not straightforward.

Several different language tasks have been used in fMRI, as reviewed by Engström et al. [2]. The hemispheric dominance of language functions in fMRI has usually been assessed by using the laterality index (LI), which is relatively concordant with the Wada test [3–5], although differences have been reported as well [6]. However, the limitation of fMRI for lateralizing language function in clinical practice is that fMRI reveals not only essential but also contributory areas in language processing. Previous studies have shown that a combination analysis of several language tasks is more robust than a single task for the detection of essential language areas, and also less sensitive to supplementary areas [7–11]. Furthermore, LIs based on a combination of several tasks are shown to be more congruent with the Wada test or electrocortical stimulation [12, 13]. However, different tasks have seldom been compared quantitatively and thus there is no consensus regarding which tasks should be used in the clinical setting.

In a recent review article, criteria for assessing the usefulness of fMRI language paradigms were defined [14]. According to the criteria, the paradigm should (1) lateralize to the left hemisphere in healthy, right-handed adults, (2) create robust activation, (3) be in concordance with the lateralization measured with other techniques, such as the Wada test, and (4) produce activations in particular target area depending on the surgical target.

In this study, we evaluated five language tasks with healthy volunteers to determine the optimal combination of tasks for clinical use. According to the aforementioned criteria, an optimal combination should cover different aspects of language processing and detect the essential language areas robustly. Furthermore, since our subjects were healthy volunteers, the optimal combination should lateralize to the left hemisphere especially in the right-handed individuals. For this purpose, the task variability in activation power and LI of different tasks and their combinations was examined. In addition, to assess clinical feasibility, the laterality of language functions based on activation maps was assessed by an experienced clinical neuroradiologist.

## Materials and methods

### Subjects

A total of 20 healthy subjects (11 females; mean age 36 years, range 18–59) participated in this study. Subjects were recruited from the partners of epilepsy patients and from the university and hospital staff. The majority of the subjects were right-handed ( $n=16$ ), two were left-handed, and two were ambidextrous (the Edinburgh Handedness Inventory, [15]). Detailed subject demographics are shown in Table 1. The study was approved by the local ethics committee and the subjects gave their written informed consent.

### fMRI tasks

Five different fMRI tasks were used: word generation (WGEN) [16–19], letter task (LET) [11, 20], responsive naming (RNAM) [6, 12, 21], sentence comprehension (SCOMP) [3, 22, 23], and word pair (WP). The tasks were chosen so that the brain regions generally thought to be specialized in different aspects of language processes were activated as widely as possible. Furthermore, both auditory and visual stimuli were included to ensure that suitable language tasks were also available for patients with poor eyesight or hearing problems. A conventional block design was used in all tasks. Details of the task paradigms are presented in Table 2.

In the WGEN task, the active condition was to covertly generate different words starting with a given letter. In the control condition, subjects were shown a fixation cross and were instructed to stop generating words and instead perform a simple finger tapping task with the right and left index fingers.

In the LET task, line drawings of common objects were presented and the task was to press one of two buttons depending on whether the name of the presented object contained a phoneme [i] or not. The control condition consisted of abstract line drawings with smaller images that either were (press the “yes” button) or were not (press the “no” button) a part of the larger image.

In the RNAM task, written descriptions of common nouns (“a long yellow fruit”) were presented. The subjects were instructed to read the description and to covertly name it (“a banana”). The control condition was to look at varying rows of hashes (##### ##).

In the SCOMP task, short spoken sentences, either semantically correct or incorrect, were played to the subjects through headphones. The task was to decide whether the sentence was correct or not, and to press one of two buttons accordingly. The control condition consisted of short series of tones; the subjects were instructed to press

**Table 1** Subject demographics

Subject #	Age (years)	Gender	Handedness	Education (years)
1	18	Female	Right	11
2	27	Female	Right	16
3	27	Male	Right	16
4	52	Female	Right	15
5	26	Female	Right	16
6	23	Male	Right	14
7	38	Female	Right	16
8	38	Male	Ambidextrous	16
9	59	Female	Left	13
10	58	Male	Right	9
11	44	Male	Right	11
12	34	Female	Right	16
13	31	Male	Right	11
14	47	Male	Right	13
15	42	Female	Right	12
16	18	Female	Right	11
17	19	Male	Left	11
18	40	Female	Right	12
19	42	Male	Ambidextrous	12
20	40	Female	Right	16

Subjects' age, gender, handedness, and years of education are presented. Minimum length of education (9 years) means that the subject had only basic education, while maximum (16 years) means that the subject obtained an academic degree

the “yes” button if there were two high pitch tones in the series and the “no” button if there were none or only one high pitch tone.

In the WP task, subjects heard pairs of spoken nouns and their task was to covertly name the category to which both of the nouns belonged (i.e., “a ring” and “a bracelet” would both belong to the category “jewelry”). In the control condition, subjects heard pairs of tones and indicated by a button press whether the last tone was higher than the first one.

The Presentation software package (Neurobehavioral Systems, Inc., USA) was used for stimulus delivery. Visual stimuli were projected onto a transparent screen and subjects looked at the screen via a mirror attached to the

head coil. Auditory stimuli were delivered to the subjects through MRI-compatible headphones (NordicNeuroLab AS, Bergen, Norway). Responses were recorded using an MRI-compatible response interface (NordicNeuroLab AS).

#### MRI scanning

Subjects were scanned using a 1.5-T MRI scanner (Siemens Magnetom Avanto, Siemens AG, Germany). Before the functional tasks, a 3D Magnetization Prepared Rapid Acquisition GRE T1-weighted image was acquired for anatomical reference (TR 1,980 ms, TE 3.09 ms, flip angle 15°, matrix 256×256, 176 contiguous sagittal slices with isotropic 1 mm<sup>3</sup> voxels). All functional images consisted of

**Table 2** Structures of the task paradigms

Name of the task	Modality	Structure of the task paradigm				Structure of a block	
		Number of active blocks	Duration (min)	TR (ms)	Response	Content	Length
Word generation (WGEN)	Visual	4	5	3,100	None	One letter or motor task	10 Scans
Letter task (LET)	Visual	5	7	3,100	“Yes/no”	6 Pictures or puzzles	10 Scans
Responsive naming (RNAM)	Visual	5	10.5	3,530	None	10 Descriptions or hashes	14 Scans
Sentence comprehension (SCOM)	Auditory	5	8.5	3,070	“Yes/no”	8 Sentences or tone series	13 Scans
Word pair (WP)	Auditory	4	7	3,100	None	8 Pairs of words or tones	13 Scans

A conventional block design, in which the control condition alternated with the active condition, was used in all tasks. The first and the last block were always a control condition block

36 contiguous slices with a voxel size of  $3 \times 3 \times 3$  mm<sup>3</sup> acquired parallel to the AC–PC plane with a gradient-echo echo-planar imaging sequence (TE 50 ms, flip angle 90°, matrix 64×64). Four additional scans were acquired before each task to ensure stability of the image intensity level. These additional scans were excluded from the analysis. The total scanning time, not including patient positioning, pauses between the tasks, or instructions, was approximately 50 min.

#### Data analysis

fMRI data was analyzed with the SPM5 software package (the Wellcome Trust Center for Neuroimaging, UCL, London, UK, [www.fil.ion.ucl.ac.uk/spm/](http://www.fil.ion.ucl.ac.uk/spm/)) running under Matlab R2007a (The Mathworks Inc., Natick, MA, USA). Preprocessing of the images included motion correction, correction of the acquiring time difference between slices, normalization to the standard ICBM152-space, and finally, spatial smoothing with a Gaussian kernel of 6 mm FWHM. Statistical analysis was performed on a voxel-by-voxel basis using the general linear model [24]. Data of all tasks were entered into the same design as separate sessions to enable analyses of the different task combinations. Nine statistical analyses were performed for each subject, all comparing the active condition and the control condition with a *t* test: one analysis for each language task separately, combination analyses of the visual and auditory tasks, and a combination analysis of all tasks. Finally, based on the results of the single task analyses, an additional combination analysis of the WGEN, RNAM, and SCOMP tasks was performed.

It has been shown that LI corresponds better with the Wada test if only language areas are used in the LI calculation [25, 26]. Therefore, the following regions of interest (ROIs) were defined: (1) Broca ROI, consisting of Brodmann's area (BA) 44, BA45, and BA47; (2) Wernicke ROI, consisting of BA21, BA22, BA37, BA39, and BA40; and (3) combined ROI, including both the Broca ROI and the Wernicke ROI complemented with BA46, Heschl gyrus, and the hippocampus. The ROIs were defined using the atlases in WFU PickAtlas [27, 28].

To explore the activation power and laterality of the activation between the analyses in more detail, two scatter plot analyses were performed. First, the sum of positive *T* values within the combined ROI was calculated for both the left and right hemispheres for each subject. These sums were further normalized to the maximum of the sums over the subjects and analyses to enable the cross-analysis comparison. The normalized sums of positive *T* values represent the intensity of the activations. The left hemisphere normalized sums were then plotted against the corresponding right hemisphere sums and group mean and standard

deviations were calculated for each analysis. Second, to assess the extent of the activations within the combined ROI, the number of activated voxels ( $p < 0.005$ , uncorrected) was calculated for the left and right hemisphere. Again, the numbers of activated voxels on the left hemisphere were plotted against the numbers of activated voxels on the right hemisphere and group mean and standard deviations were calculated. In both of the scatter plot analyses, the farther off the dot is from the lower left corner, the stronger the activation has been. If the dot is located near the  $R = L$  line, there has been activation on both the left and right hemisphere, whereas if the dot is near either *y*- or *x*-axis, the activation has been concentrated mostly on only one of the hemispheres.

The laterality index was calculated as  $LI = [(L - R)/(L + R)]$ , where *L* and *R* are the number of voxels surviving the threshold in these ROIs on the left and the right hemisphere, respectively. The threshold was defined for each subject individually as 80% of the maximum *T* value of the combined ROI [26]. Language dominance was defined based on the LIs, with  $LI > 0.1$  indicating left hemisphere dominance,  $LI < -0.1$  indicating right hemisphere dominance, and  $-0.1 \leq LI \leq 0.1$  indicating bilateral language dominance.

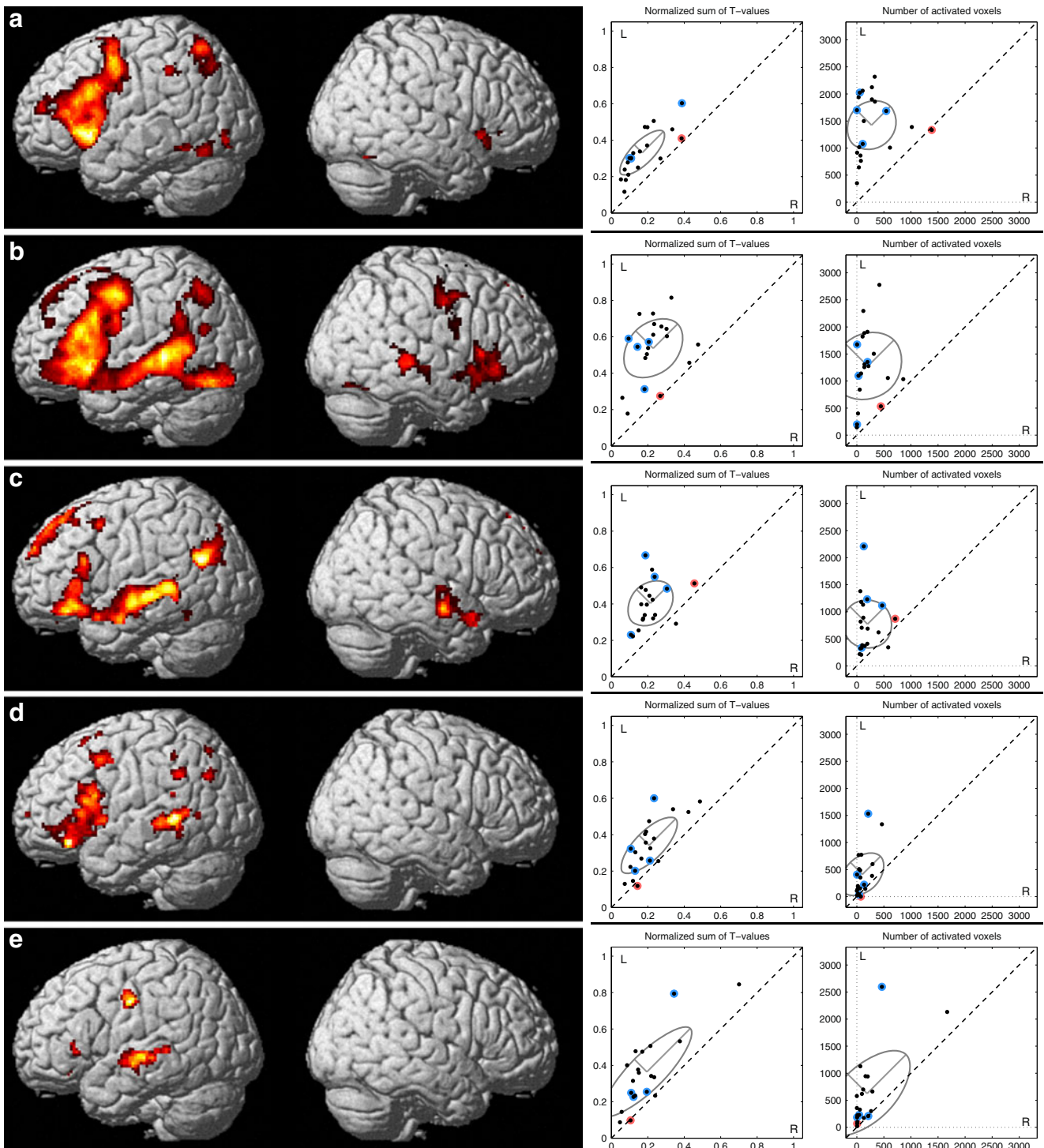
#### Visual assessment of language lateralization

The individual fMRI results for each task and the different task combinations were further evaluated visually by an experienced neuroradiologist blinded to the calculated LI results. The activation threshold was chosen for each subject and analysis individually excluding clearly false positives (e.g., in the midbrain or near the eyes) but including logical activations. Language lateralization was classified separately for the whole brain, Wernicke's area, and Broca's area using a five-step dominance scale: definitely left, probably left, bilateral or questionable, probably right, and definitely right dominance. If there were not enough active voxels or there were some other uncertainties in the activations, then laterality was not assessed. The observer was told that the proportion of the non-right-handed individuals was notably higher than in a normal population.

## Results

#### Activation levels and laterality

The overall fMRI activation level varied not only between subjects, but also between tasks and analyses within subject. Figure 1 illustrates the group results of each task using the threshold of  $p < 0.05$  (false discovery rate (FDR) corrected) and corresponding scatter plots of the sums of



**Fig. 1** Group results of single tasks. **a** Word generation (WGEN), **b** responsive naming (RNAM), **c** sentence comprehension (SCOMP), **d** letter task (LET), and **e** word pair (WP). The results are presented using the threshold of  $p < 0.05$  (FDR corrected). The scatter plots on the left are calculated as normalized sums of positive  $T$  values and on the right as number of activated voxels ( $p < 0.005$ , uncorrected) within the combined ROI for each task. The standard deviation ellipses

visualize the group means (central point) and variations in activation power ( $2 \times \text{SD}$  on major axis along the  $L = R$  line) and in lateralization ( $2 \times \text{SD}$  on the other major axis, perpendicular to the  $L = R$  line). The non-right-handed subjects are highlighted in blue in the scatter plots. The red dot in the scatter plots highlights a subject with highly variable laterality according to the LI and visual inspection (subject #14)

the  $T$  values and of the number of activated voxels (Fig. 1a, WGEN; Fig. 1b, RNAM; Fig. 1c, SCOMP; Fig. 1d, LET; Fig. 1e, WP). The mean and standard deviations of both of the sums of the  $T$  values and number of activated voxels are illustrated with an ellipse of radii of  $2 \times SD$  in each scatter plot in Fig. 1 and presented in Table 3. The standard deviation along the  $L = R$  line ( $SD_{act}$ ) represents the variation in total activation power and the standard deviation perpendicular to that ( $SD_{lat}$ ) represents the variation in laterality. When the tasks were examined separately, the RNAM produced the strongest activation (highest group mean in terms of both intensity and extent). The  $SD_{act}$  was largest in the WP thus reducing the robustness of its activation power. The variation in laterality was smallest in the WGEN task and the SCOMP task had the smallest variation in total activation power. To be able to select tests to the optimal task battery, we used the criteria proposed by Binder [14]. Three tasks were chosen based on the robustness of the activation power and laterality, and capability to activate the language network as widely as possible: WGEN, RNAM, and SCOMP. Among the visual tasks, RNAM and WGEN produced both strong activations with clear left hemisphere lateralization. The LET task was excluded because of its poor activation power. The SCOMP task was selected as the auditory task since it produced much stronger activation with less variation than the WP task.

After comparing the five individual tasks, we performed comparisons between different combinations of tasks: (a) all tasks, (b) visual tasks, (c) auditory tasks, and (d) an optimal combination of visual and auditory tasks, i.e., WGEN, RNAM and SCOMP. Group results of activated areas in all four combination analyses are illustrated in

Fig. 2 using the threshold of  $p < 0.05$  (FDR corrected) along the similar scatter plots as for the single task analyses. The mean and standard deviations of the sums of the  $T$  values and of the number of activated voxels are illustrated with an ellipse of radii of  $2 \times SD$  in each scatter plot in Fig. 2 and further presented in Table 3. At the group level, the combination of all five tasks (Fig. 2a) and the combination of WGEN, RNAM, and SCOMP (Fig. 2d) produced almost identical results in terms of activated areas but the standard deviations  $SD_{lat}$  and  $SD_{act}$  were smaller in the combination of WGEN, RNAM, and SCOMP.

#### Laterality indices

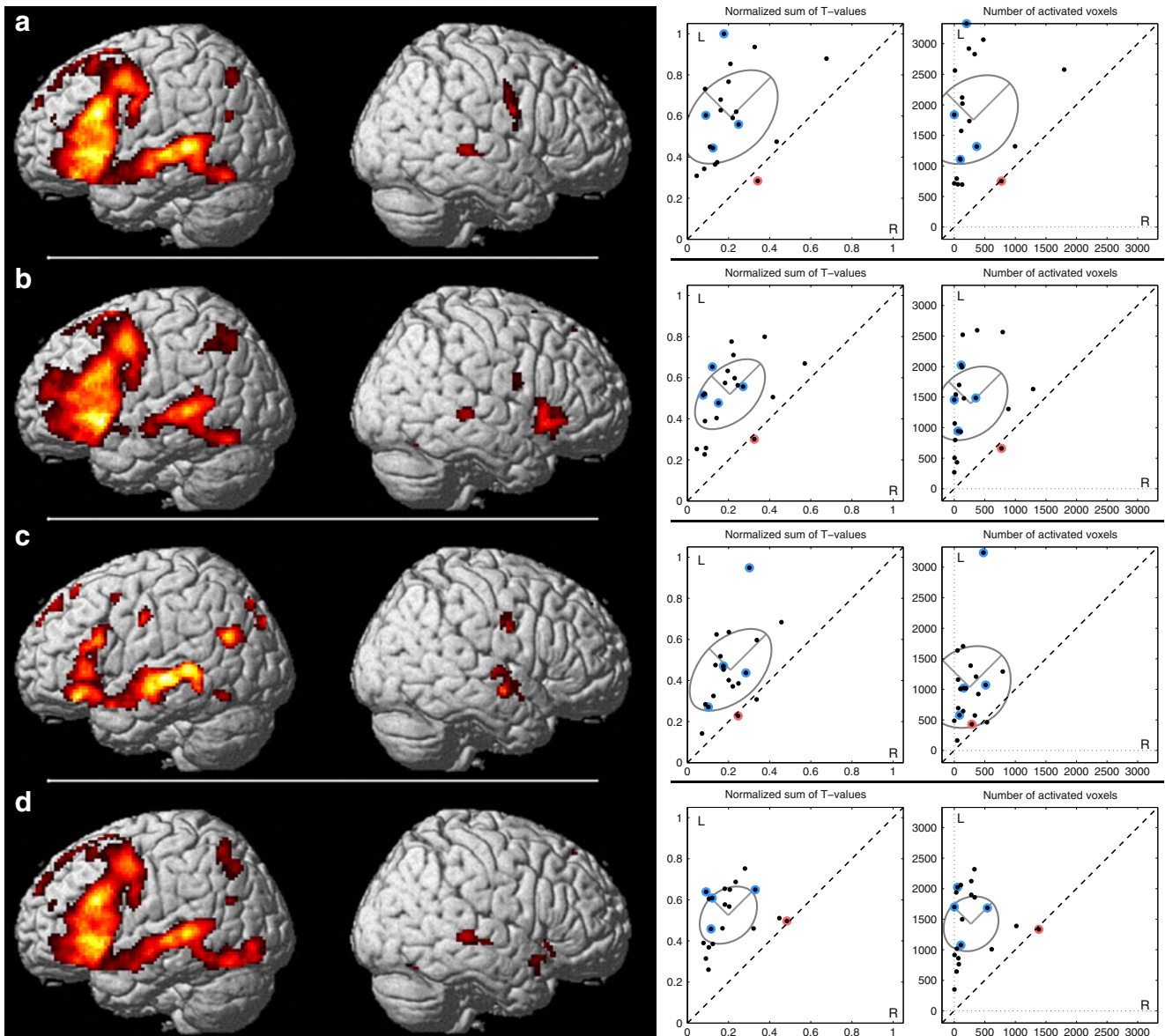
The LIs for each subject and for each analysis are presented in Tables 4 and 5. The subjects are assorted in the tables based on the sum of the LIs so that the subjects with the most congruent left hemisphere dominance are presented first. Only 12 of the subjects (two left-handed, one ambidextrous) were classified as left hemisphere dominant in all analyses and ROIs. For the rest of the subjects ( $n=8$ , 1 ambidextrous), the LI varied from clear left hemisphere dominance to bilateral to clear right hemisphere dominance depending on the task, the analysis, and the ROI. Figure 3 illustrates especially intense hemispheric and regional variation between tasks in a single subject (subject #14).

The RNAM task produced clear left hemisphere dominance for 19 of 20 subjects, all ROIs included, while the largest deviation in language dominance between subjects was found in the LET task. When the LI was calculated based on the Wernicke ROI only, 13 subjects were classified as left hemisphere dominant in all analyses; if the analysis was done using the Broca ROI only, 15

**Table 3** The distributions of the activations on the left and right hemisphere

Analysis	Normalized sum of positive $T$ values				Number of activated voxels			
	Group mean, left hemisphere	Group mean, right hemisphere	$SD_{lat}$	$SD_{act}$	Group mean, left hemisphere	Group mean, right hemisphere	$SD_{lat}$	$SD_{act}$
WGEN	0.33	0.17	0.03	0.08	711	193	152	281
RNAM	0.54	0.23	0.07	0.09	1,365	230	303	357
SCOMP	0.40	0.21	0.05	0.07	841	224	241	238
LET	0.34	0.21	0.04	0.10	462	123	164	249
WP	0.36	0.19	0.05	0.17	681	214	252	543
All tasks	0.59	0.21	0.09	0.14	1,841	341	319	439
Visual tasks	0.32	0.21	0.06	0.11	1,476	294	263	370
Auditory tasks	0.45	0.21	0.07	0.12	1,112	279	313	388
WGEN, RNAM, SCOMP	0.52	0.20	0.06	0.08	1,509	299	226	262

The table presents the group means of the normalized sums of positive  $T$  values and of the number of activated voxels ( $p < 0.005$ , uncorrected) for each analysis for left and right hemispheres.  $SD_{lat}$  is the standard deviation of the lateralization, i.e., the smaller the  $SD_{lat}$ , the more congruent is the lateralization of the analysis within the group.  $SD_{act}$  is the standard deviation of the total activation power of the analyses



**Fig. 2** Group results of combination analyses. **a** All tasks in the same analysis, **b** visual tasks (WGEN, LET, RNAM), **c** auditory tasks (SCOMP, WP), and **d** combination of WGEN, RNAM, and SCOMP. The results are presented using the threshold of  $p < 0.05$  (FDR corrected). The scatter plots on the left are calculated as normalized sums of positive  $T$  values and on the right as number of activated voxels ( $p < 0.005$ , uncorrected) within the combined ROI for each task.

The standard deviation ellipses visualize the group means (central point) and variations in activation power ( $2 \times SD$  on major axis along the  $L = R$  line) and in lateralization ( $2 \times SD$  on the other major axis, perpendicular to the  $L = R$  line). The non-right-handed subjects are highlighted in blue in the scatter plots. The red dot in the scatter plots highlights a subject with highly variable laterality according to the LI and visual inspection (subject #14)

subjects were left hemisphere dominant. Only one subject had exactly the same LI ( $LI = 1.00$ ) in all analyses and ROIs (subject #19).

Visually assessed language lateralization

The language lateralization assessments conducted by a neuroradiologist are presented in Tables 4 and 5, along with the calculated LIs. The observer’s confidence on interpre-

tation of laterality was greater when assessing the laterality jointly from the Broca and Wernicke areas, when compared to the separate analyses of these functional areas. When the five individual tasks were assessed separately, the visual interpretation provided variable and even contradictory results (Tables 4 and 5). However, when assessing the combinations of tasks the interpretation of lateralization became more congruent and the observer’s confidence on the results increased. When the ROI-based LI calculations

**Table 4** Calculated laterality indices and visual assessments for combination analysis of all tasks; visual tasks; auditory tasks; and combined WGEN, RNAM, and SCOMP for all subjects using three different ROI combinations: combination ROI, Wernicke ROI, and Broca ROI

Subj.	All tasks						Visual tasks (WGEN, LET, RNAM)			Auditory tasks (SCOMP, WP)			WGEN, RNAM, SCOMP		
	Comb. ROI	Wernicke	Broca	Comb. ROI	Wernicke	Broca	Comb. ROI	Wernicke	Broca	Comb. ROI	Wernicke	Broca	Comb. ROI	Wernicke	Broca
	3	1	1	1	1	1	1	1	1	1	1	1	1	1	1
4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
9 <sup>a</sup>	1	1	1	1	1	1	1	pL	1	NA	1	pL	1	1	NA
11	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
13	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
16	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
17 <sup>a</sup>	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
18	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
19 <sup>b</sup>	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
20	1	1	1	1	1	0.92	1	1	1	1	1	1	1	1	1
6	1	1	1	1	1	1	0.89	0.82	1	1	1	1	pL	1	B
5	1	1	pL	1	1	1	1	1	1	1	1	1	1	0.66	B
12	1	1	1	1	1	1	0.78	pL	0.63	B	1	1	1	1	1
10	1	1	1	1	1	1	0.80	pL	0.80	0.67	NA	1	1	1	1
2	0.71	0.30	1	1	1	1	1	1	1	1	1	1	1	0.85	1
8 <sup>b</sup>	1	1	1	1	1	1	0.74	pL	0.74	B	0.33	pL	1	1	1
15	1	1	1	1	1	1	-0.25	-0.25	-0.25	1	1	1	1	1	1
1	1	1	0.90	1	1	1	0.44	0.42	0.42	-0.67	1	1	1	1	1
7	0.70	0.64	B	0.57	0.79	0.14	-0.18	-0.18	-0.18	1	0.94	1	0.94	1	pL
14	-1	-1	B	0.32	B	0.92	0.25	pL	0.25	pL	1	NA	0.84	B	1.00

Visual assessment of hemispheric dominance is denoted with letters if it was not clear left hemisphere dominance: *pL* probable left, *B* bilateral, *R* right, *NA* not enough active voxels for evaluation. The subjects are presented in descending order based on the sum of the LIs. LIs of 1.00 or -1.00 are presented as 1 or -1, respectively, to improve the readability of the table

<sup>a</sup> Subject is left-handed

<sup>b</sup> Subject is ambidextrous



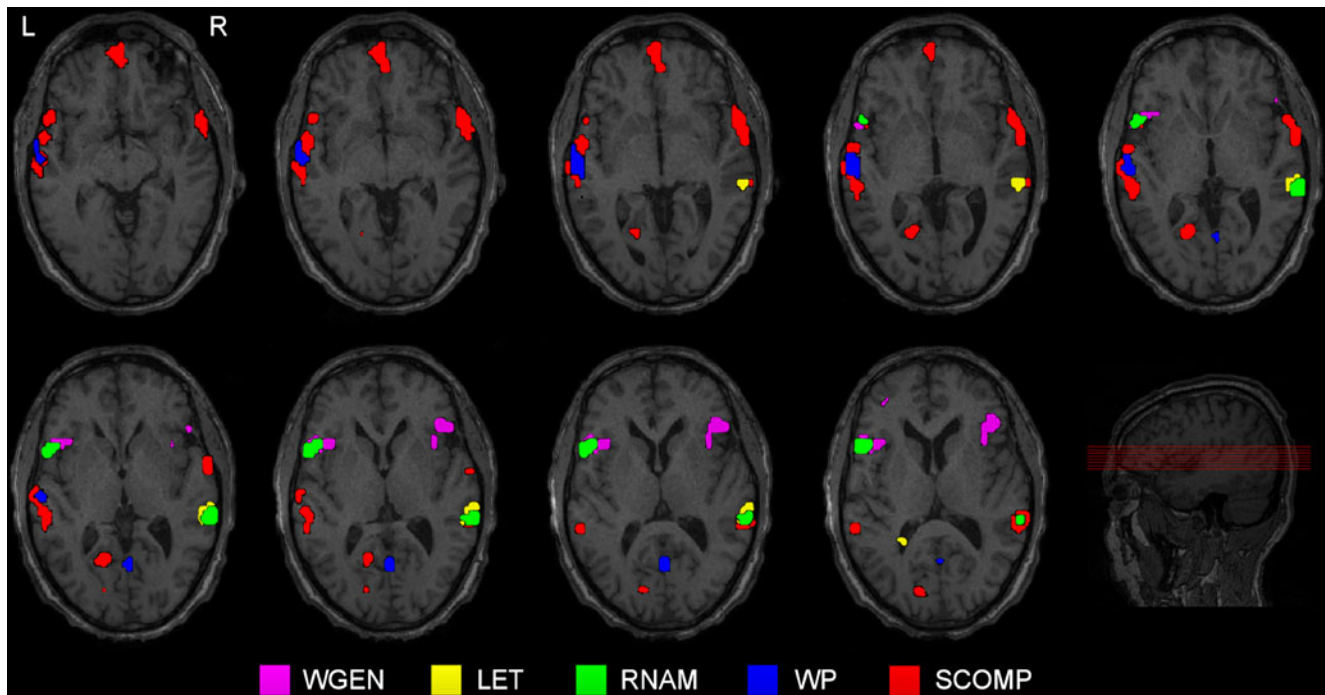
**Table 5** Calculated laterality indices and visual assessments for each task and subject using three different ROI combinations: combination ROI, Wernicke ROI, and Broca ROI

Subj.	Word generation (WGEN)			Letter task (LET)			Responsive naming (RNAM)			Word pair (WP)			Sentence comprehension (SCOMP)		
	Comb. ROI	Wernicke	Broca	Comb. ROI	Wernicke	Broca	Comb. ROI	Wernicke	Broca	Comb. ROI	Wernicke	Broca	Comb. ROI	Wernicke	Broca
3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
4	1	pL	1	0.81	0.13	pL	1	1	1	1	0.52	pL	1	0.88	1
9 <sup>a</sup>	1	1	1	0.29	0.27	1	1	pL	1	NA	1	pL	1	pL	1
11	1	1	1	0.57	0.00	1	1	1	1	1	0.95	1	1	1	1
13	1	1	1	1	1	1	1	1	1	1	1	1	1	0.42	pL
16	1	pL	1	1	0.85	pL	1	1	1	1	1	1	1	1	1
17 <sup>a</sup>	0.33	pL	1	1	1	1	1	1	1	1	1	1	1	1	1
18	1	1	1	0.84	0.68	pL	1	1	1	1	0.90	1	1	1	1
19 <sup>b</sup>	1	1	1	1	1	pL	1	1	1	1	1	1	1	1	1
20	1	1	0.95	0.40	0.22	pL	1	1	1	1	1	1	1	1	1
6	1	1	1	0.66	0.47	1	1	1	1	pL	1	1	0.75	pR	0.72
5	1	-1	1	1	1	1	1	1	1	1	1	1	0.94	B	1
12	1	1	1	1	0.92	1	1	1	1	1	1	pL	1	0.33	B
10	1	1	1	1	1	pL	1	1	1	pL	0.50	pL	1	0.33	B
2	1	1	1	1	1	0.95	1	1	1	1	1	1	1	1	1
8 <sup>b</sup>	0.94	0.85	1	-0.29	-0.30	-1	pL	1	1	1	-0.23	B	-0.19	pL	-1
15	1	pL	1	1	1	1	1	1	1	pL	-0.20	pL	1	0.31	0.44
1	0.67	0.67	1	0.12	0.07	B	1	1	1	1	0.39	B	-0.24	-0.75	B
7	0.05	B	0.36	0.69	0.46	B	1	1	1	0.13	B	0.26	B	0.65	0.78
14	0.66	B	-1	1	pR	-1	pR	-1	NA	-0.16	pL	1	pL	1	1

Visual assessment of hemispheric dominance is denoted with letters if it was not clear left hemisphere dominance: *pL* probably left, *B* bilateral, *R* right, *pR* probably right, *NA* not enough active voxels for evaluation. The subjects are presented in the same order as in Table 3. Lis of 1.00 or -1.00 are presented as 1 or -1, respectively, to improve the readability of the table

<sup>a</sup> Subject is left-handed

<sup>b</sup> Subject is ambidextrous



**Fig. 3** An illustrative case of large variation and contradictory results in activations and language laterality between tasks in a single subject (subject #14). The general activation level varied between the tasks

and the visual assessments of the combinations of tasks were compared, the results were contradictory in one individual.

## Discussion

This study compared five different language tasks to determine hemispheric language dominance and overall activation power, and to establish the most suitable task combination for clinical use. The tasks were compared using the criteria proposed in a recent review article [14]. According to these criteria, the task should (1) lateralize to the left hemisphere in healthy, right-handed adults, (2) create robust activation, (3) be in concordance with the lateralization measured with other techniques, such as the Wada test, and (4) produce activations in particular target area depending on the surgical target. The purpose of this study was to find the most optimal combination of tasks to be used in defining the language lateralization in clinical patients. Since the total MRI scanning time of these five tasks was relatively long (50 min), we wanted to limit the number of tasks with healthy volunteers before applying them to patients. The use of healthy volunteers naturally limits the applicability of the above mentioned third criterion, i.e., the comparison of the result with other techniques such as the Wada test. Therefore, the selection of the tasks for the optimal task battery was performed

too much to use a single threshold for all tasks. Therefore, the threshold was set separately for each task by the neuroradiologist, based on whole brain activation maps

using the three remaining criteria. In the present study, a combination of WGEN, RNAM, and SCOMP was chosen to form the optimal task battery. The combination includes both visual and auditory tasks and the single tasks activate different parts of the language network. In the present study, they all produced relatively robust and strong activations and the pattern of the activation lateralized to the left hemisphere both in the right-handed subjects and the total study population.

The first criterion considers the use of healthy right-handed volunteers. In our study, the majority of the subjects ( $n=16$ ) were right-handed, but four non-right-handed subjects were studied as well originally to be able to test heterogeneous results and to reflect clinical reality. All these four non-right-handed subjects, however, showed clear left hemisphere dominance and they were therefore included in the group analyses.

According to the fourth criterion, the fMRI task should produce activations in the surgical target area. Since the surgical target area varies between the patients, a combination of several different tasks activating different aspects of the language network should be used. Furthermore, it has been shown previously that a combined analysis of several tasks produces more reliable results in language lateralization than only single tasks [6, 8, 10, 11] and the use of both auditory and visual stimuli has been shown to produce less LI variability [29]. Our results agree with these prior findings. There was less LI variation between and within

subjects when combined analyses of tasks were used than with single tasks. Moreover, using both auditory and visual tasks in the combination further decreased variability. The clinical observer's confidence on the visual assessment of lateralization improved when a combination of tasks was used instead of individual tasks. Although the variation measured with  $SD_{\text{lat}}$  and  $SD_{\text{act}}$  was in general slightly higher in the combination analyses than in single task analyses, the overall activation power was higher in the combination analyses thus favoring the use of combination analyses. However, it should be noted that when tasks with relatively poor activation power (LET, WP) were included in the combined analysis, the variability between subjects and within subjects increased.

Since the laterality index is a common method to define the laterality, we wanted to see how much the LI varies depending on the analysis. There are several ways to calculate the LI [30, 31]. It has been shown that the LI depends on the threshold, as higher thresholds produce higher LIs [7, 8]. However, too high a threshold limits the number of voxels in the LI calculation, possibly biasing the results. Furthermore, the same threshold for all subjects is unreliable unless a range of thresholds is used, hence the use of an LI curve has been proposed [32, 33]. Moreover, since the LI curve may not always be easy to interpret, a subject-specific threshold has been proposed instead of a fixed threshold or an LI curve, especially in clinical work [26]. Therefore, all LIs in this study were calculated based on individually set thresholds defined as 80% of the maximum  $T$  value within the combined ROI. This method is simple, and takes into account the individual variation in overall activation level. It also balances the number of voxels used in LI calculation between subjects, enabling the comparison of different tasks and analysis methods.

However, individually adjusted thresholds as used in the LI calculation are not suitable in comparing the activation power of the different analyses between subjects. Therefore, a lower, fixed threshold ( $p < 0.005$ , uncorrected) was chosen to define the extent of activation i.e. number of activated voxels. In the scatter plot analysis, to define the intensity of the activation, we chose to use the sum of positive  $T$  values. Inclusion of all the positive values in the analysis takes into account the variation in the overall activation level between the subjects while also providing information of the variation between the hemispheres. It should be noted, however, that although different thresholds were used in this article, they all produce the same conclusion in selecting the optimal task battery.

Comparison of the visually assessed language lateralization and the calculated LI was performed to emphasize the importance of using a combination of several tasks. In the clinics, the fMRI results are often analyzed only visually

mainly because visual assessment is usually faster than complex quantitative analysis. Results of the visual assessments based on computerized combination analyses were more congruent than those based on single tasks. Visual assessment of the activations' laterality cannot be omitted, as it can interpret possible atypical language organization in cases of large variability in LIs, regionally and between tasks. The variable and sometimes contradictory results in visual assessment of the individual tasks advocate combination analyses of multiple tasks for clinical use.

The larger variation in activation between subjects in auditory tasks may partly result from the tasks' control condition, which has been shown to have an effect on LI [34]. In general, it is recommended that the control condition should be as similar as possible to the active condition but with no language processing. However, the tones used in the control conditions in this study were acoustically simpler than words. For some subjects, this situation may have left some primary auditory activation visible in the results, increasing bilateral activity in the temporal lobe. Some subjects may also have utilized verbal strategies in the control condition. Another possible cause for the large variability in the LIs in auditory tasks is possible misalignment of the atlas-based ROIs and normalized anatomy. To be able to use predefined ROIs, the individual data must be normalized to the standard space. The normalization procedure does not always guarantee complete alignment of every anatomical region. Therefore, the voxels actually belonging to the primary auditory cortex may have been falsely taken into calculation.

The largest and inevitable limitation of our study is that we cannot validate the laterality provided by the fMRI using other techniques such as the Wada test because we used healthy volunteers. Therefore, our next goal is to repeat the study with Wada-tested patients selected for an operative treatment.

## Conclusions

We have compared five fMRI language paradigms to choose an optimal combination for clinical use. The combination of WGEN, RNAM, and SCOMP tasks seems optimal in terms of activation power, robustness to detect essential language areas, and scanning time. Furthermore, we have shown that LI is highly dependent on the selected language task, and whether it is calculated based on a single task or on a combination of several tasks. For clinical purposes, the use of both auditory and visual stimuli should be advocated, since using only one type of stimulus may bias the results. However, only tasks that are proven to activate the cognitive network of interest should be included in the combination analysis. Determining

language lateralization is often challenging and both single tasks and combinations of tasks should be used to arrive at a comprehensive understanding of individual cortical language areas.

**Acknowledgments** This study was partly funded by the Academy of Finland (grant 123579), and Kuopio University Hospital EVO funding (grant 5041716).

**Conflict of interest** We declare that we have no conflicts of interest.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

## References

- Wada J, Rasmussen T (2007) Intracarotid injection of sodium amytal for the lateralization of cerebral speech dominance. 1960. *J Neurosurg* 106:1117–1133
- Engström M, Ragnehed M, Lundberg P, Söderfeldt B (2004) Paradigm design of sensory-motor and language tests in clinical fMRI. *Neurophysiol Clin* 34:267–277
- Carpentier A, Pugh KR, Westerveld M, Studholme C, Skrinjar O, Thompson JL et al (2001) Functional MRI of language processing: dependence on input modality and temporal lobe epilepsy. *Epilepsia* 42:1241–1254
- Woermann FG, Jokeit H, Luerding R, Freitag H, Schulz R, Guertler S et al (2003) Language lateralization by Wada test and fMRI in 100 patients with epilepsy. *Neurology* 61:699–701
- Binder JR, Swanson SJ, Hammeke TA, Morris GL, Mueller WM, Fischer M et al (1996) Determination of language dominance using functional MRI: a comparison with the Wada test. *Neurology* 46:978–984
- Gaillard WD, Balsamo L, Xu B, McKinney C, Papero PH, Weinstein S et al (2004) fMRI language task panel improves determination of language dominance. *Neurology* 63:1403–1408
- Ramsey NF, Sommer IE, Rutten GJ, Kahn RS (2001) Combined analysis of language tasks in fMRI improves assessment of hemispheric dominance for language functions in individual subjects. *NeuroImage* 13:719–733
- Rutten GJ, Ramsey NF, van Rijen PC, van Veelen CW (2002) Reproducibility of fMRI-determined language lateralization in individual subjects. *Brain Lang* 80:421–437
- Harrington GS, Buonocore MH, Farias ST (2006) Intrasubject reproducibility of functional MR imaging activation in language tasks. *AJNR Am J Neuroradiol* 27:938–944
- de Guibert C, Maumet C, Ferré J, Jannin P, Biraben A, Allaire C et al (2010) FMRI language mapping in children: a panel of language tasks using visual and auditory stimulation without reading or metalinguistic requirements. *NeuroImage* 51:897–909
- Wilke M, Pieper T, Lindner K, Dushe T, Holthausen H, Krägeloh-Mann I (2010) Why one task is not enough: functional MRI for atypical language organization in two children. *Eur J Paediatr Neurol* 14:474–478
- Pouratian N, Bookheimer SY, Rex DE, Martin NA, Toga AW (2002) Utility of preoperative functional magnetic resonance imaging for identifying language cortices in patients with vascular malformations. *J Neurosurg* 97:21–32
- Arora J, Pugh K, Westerveld M, Spencer S, Spencer DD, Todd CR (2009) Language lateralization in epilepsy patients: fMRI validated with the Wada procedure. *Epilepsia* 50:2225–2241
- Binder JR (2011) Functional MRI is a valid noninvasive alternative to Wada testing. *Epilepsy Behav* 20:214–222
- Oldfield RC (1971) The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9:97–113
- Cuenod CA, Bookheimer SY, Hertz-Pannier L, Zeffiro TA, Theodore WH, Le Bihan D (1995) Functional MRI during word generation, using conventional equipment: a potential tool for language localization in the clinical environment. *Neurology* 45:1821–1827
- Friedman L, Kenny JT, Wise AL, Wu D, Stuve TA, Miller DA et al (1998) Brain activation during silent word generation evaluated with functional MRI. *Brain Lang* 64:231–256
- Gaillard WD, Hertz-Pannier L, Mott SH, Barnett AS, LeBihan D, Theodore WH (2000) Functional anatomy of cognitive development: fMRI of verbal fluency in children and adults. *Neurology* 54:180–185
- Deblaere K, Backes WH, Hofman P, Vandemaele P, Boon PA, Vonck K et al (2002) Developing a comprehensive presurgical functional MRI protocol for patients with intractable temporal lobe epilepsy: a pilot study. *Neuroradiology* 44:667–673
- Wilke M, Lidzba K, Staudt M, Buchenau K, Grodd W, Krägeloh-Mann I (2006) An fMRI task battery for assessing hemispheric language dominance in children. *NeuroImage* 32:400–410
- Gaillard WD, Pugliese M, Grandin CB, Braniecki SH, Kondapaneni P, Hunter K et al (2001) Cortical localization of reading in normal children: an fMRI language study. *Neurology* 57:47–54
- Constable RT, Pugh KR, Berroya E, Mencl WE, Westerveld M, Ni W et al (2004) Sentence complexity and input modality effects in sentence comprehension: an fMRI study. *NeuroImage* 22:11–21
- Sanjuan A, Forn C, Ventura-Campos N, Rodriguez-Pujadas A, Garcia-Porcar M, Belloch V et al (2010) The sentence verification task: a reliable fMRI protocol for mapping receptive language in individual subjects. *Eur Radiol* 20:2432–2438
- Worsley KJ, Friston KJ (1995) Analysis of fMRI time-series revisited—again. *NeuroImage* 2:173–181
- Spreer J, Arnold S, Quiske A, Wohlfarth R, Ziyeh S, Altenmüller D et al (2002) Determination of hemisphere dominance for language: comparison of frontal and temporal fMRI activation with intracarotid amytal testing. *Neuroradiology* 44:467–474
- Chlebus P, Mikl M, Brázdil M, Pazourková M, Krupa P, Rektor I (2007) fMRI evaluation of hemispheric language dominance using various methods of laterality index calculation. *Exp Brain Res* 179:365–374
- Maldjian JA, Laurienti PJ, Kraft RA, Burdette JH (2003) An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *NeuroImage* 19:1233–1239
- Maldjian JA, Laurienti PJ, Burdette JH (2004) Precentral gyrus discrepancy in electronic versions of the Talairach atlas. *NeuroImage* 21:450–455
- Suarez RO, Whalen S, O’Shea JP, Golby AJ (2008) A surgical planning method for functional MRI assessment of language dominance: influences from threshold, region-of-interest, and stimulus mode. *Brain Imag Behav* 2:59–73
- Seghier ML (2008) Laterality index in functional MRI: methodological issues. *Magn Reson Imaging* 26:594–601
- Wilke M, Lidzba K (2007) LI-tool: a new toolbox to assess lateralization in functional MR-data. *J Neurosci Methods* 163:128–136
- Suarez RO, Whalen S, Nelson AP, Tie Y, Meadows ME, Radmanesh A et al (2009) Threshold-independent functional

- MRI determination of language dominance: a validation study against clinical gold standards. *Epilepsy Behav* 16:288–297
33. Ruff IM, Petrovich Brennan NM, Peck KK, Hou BL, Tabar V, Brennan CW et al (2008) Assessment of the language laterality index in patients with brain tumor using functional MR imaging: effects of thresholding, task selection, and prior surgery. *AJNR Am J Neuroradiol* 29:528–535
34. Hund-Georgiadis M, Lex U, von Cramon DY (2001) Language dominance assessment by means of fMRI: contributions from task design, performance, and stimulus modality. *J Magn Reson Imaging* 13:668–675