

ORIGINAL
RESEARCH

S. Rau
G. Fesl
P. Bruhns
P. Havel
B. Braun
J.-C. Tonn
J. Ilmberger

Reproducibility of Activations in Broca Area with Two Language Tasks: A Functional MR Imaging Study

BACKGROUND AND PURPOSE: Functional MR imaging (fMRI) is rapidly evolving and claims to complement or even substitute intraoperative mapping (IOM) of language functions. However, little is known about the reproducibility of imaging data in the language domain. The aim of our study was to assess the reproducibility of activations for 2 widely used paradigms: naming and word generation. Individual analysis was focused on the Broca area and the left insula.

MATERIALS AND METHODS: We examined 13 healthy right-handed subjects in 3 sessions with fMRI. Two conditions were assessed: overt naming and overt naming plus noun generation. The same stimuli were used in all of the sessions. A random-effects analysis was performed to analyze whole-brain activation on a group level. For the regions of interest, the number of voxels classified as active were counted for each subject, and individual reproducibility coefficients were calculated over sessions.

RESULTS: For the naming condition, the random-effects analysis did not reveal significant activations in the specified regions; small individual activations were not reproducible. For the combined task, all of the subjects showed activations in the Broca area that were more extensive and reproducible than in the naming task. Activations in the insula were only poorly reproducible.

CONCLUSION: Naming is an approved task in IOM but does not identify the Broca area with fMRI in a reproducible way. Priming may have affected our results, but the use of a combined task, in which naming is paired with noun generation, improves the reproducibility of activations and is also suitable for IOM.

The identification of language areas in patients undergoing brain surgery is a major clinical challenge. One very crucial language area is that of Broca in the left inferior frontal gyrus; it represents a “language epicenter” within the language network¹ and is involved in a variety of different linguistic tasks.^{2–7} In our view, the gold standard for the identification of essential language areas in neurosurgical patients is intraoperative mapping (IOM) by direct cortical stimulation as used by various groups,^{8–10} but other techniques, such as functional imaging, are rapidly evolving. Before imaging data can be applied within the framework of neurosurgical planning (eg, for determining the extent of a resection), data on the reproducibility of the language paradigms should be available. However, only a few reports exist on the reproducibility or test-retest reliability of functional imaging results with respect to language processes. Brannen et al¹¹ used a phonologic generation task in patients and found a mean reproducibility ratio of 37% for activations in Brodmann areas 9, 44, 45, and 46. Rutten et al¹² used naming, verb generation, and antonym generation tasks and reported data for summed up (frontal and posterior) language regions; for the single tasks, the per-

centage of overlapping voxels ranged from 10% to 30% depending on task and statistical threshold. Reproducibility increased to approximately 40% of overlapping voxels when tasks were combined. Fernández et al¹³ focused on lateralization effects using a semantic decision task; regarding voxel-by-voxel analysis, they report a reproducibility of approximately 45% for whole-brain activations. Otzenberger et al¹⁴ compared verb generation, lexical decision, and word listening tasks and found, on the basis of a conjunction analysis, the most consistent activations for the first 2 tasks, especially in the frontal lobe. A very recent study¹⁵ used a single trial design and phonologic generation and reports that approximately 45% of the voxels in the Broca area were activated in more than half of the trials.

Assessing the reproducibility of activations found with language paradigms in a repeated-measures design poses a special problem with regard to selection of stimuli. In IOM, the most often used task during cortical stimulation is picture naming, whereas in functional imaging a variety of language tasks have been applied, with a very prominent type of task being word generation. In the context of IOM, the patient has to be familiar with the material presented to him in the operating room. The patient is confronted with the same material before and during the operation, and this material may even be customized in cases where there are (slight) aphasic disturbances preoperatively. If new material were presented during the intraoperative stimulation process, effects of unfamiliarity with the material and effects of stimulation could not be distinguished. With respect to functional imaging, however, priming or repetition effects are well known from neurolinguistic imaging experiments^{16–19}; these priming effects may influence reproducibility of results when the same stimulus material is used

Received February 23, 2006; accepted after revision January 3, 2007.

Departments of Physical Medicine and Rehabilitation (S.R., J.I.), Neuroradiology (G.F., P.B., B.B.), and Neurosurgery (P.H., J.-C.T.), University Clinics of Munich, Munich, Germany.

This study was supported by the Bayerische Forschungsförderung (grant 475/01) and BrainLab AG (Heimstetten, Germany).

The data used in this study are part of the unpublished doctoral dissertation by Dr Rau, Medical Faculty of the Ludwig-Maximilians-University, Munich, Germany.

Address correspondence to Josef Ilmberger, Department of Physical Medicine and Rehabilitation, University Clinics of Munich, Marchioninistr 15, 81377 Munich, Germany; e-mail: Josef.Ilmberger@med.uni-muenchen.de

DOI 10.3174/ajnr.A0581

repeatedly. To avoid this, one could use new stimulus material (matched with regard to frequency, complexity, familiarity, etc) in each measurement; however, this strategy is not useful in the clinical context of IOM, which is used in a routine way by us and other groups. Even if a paradigm using new material in each repeated session would provide better reproducibility of activations, such data could not be used for validating functional MR imaging (fMRI) data by IOM or in the clinical context of planning neurosurgical interventions, because preoperative familiarity with the stimulus material is a basic requirement for the reasons stated above.

In this study, we thus tested the reproducibility of 2 language paradigms, naming and a combination of naming and noun generation, with fMRI in repeated measurements using the same material in all of the sessions, being well aware of the fact that priming or repetition phenomena may influence the experimental outcome. Earlier studies have not documented well the choice of the stimulus material with regard to familiarity. Our tasks were designed with regard to IOM in which naming has to be the core task; intraoperatively, we would rather not rely on a pure generation task for reasons given below. For voxel counts and reproducibility measures, we focused on activation patterns found in the Broca area and the left insula; the insula was included because possible activation reduction in the Broca area in repeated measures may be accompanied by increasing activation in the insula, possibly reflecting the “engagement of the insula in the automation of verbal tasks.”¹⁶

Materials and Methods

Experiment 1

Subjects. We examined 13 healthy subjects (age, 18–40 years; mean age, 24 years; 6 women and 7 men) in 3 sessions at intervals of 3–35 days (mean, 9 days). The subjects gave written informed consent before the beginning of the study, which was performed in accordance with the 1964 Declaration of Helsinki and approved by the local ethics committee. All of the subjects were right-handed according to the Edinburgh handedness inventory²⁰ (ranging from +71 to +100; mean, +91).

Data Acquisition. fMRI was performed on a 1.5T clinical system (Vision; Siemens, Erlangen, Germany) equipped with a gradient booster system and a head volume radio-frequency coil. To minimize head movements, head pads and a forehead strap were used.

Imaging was performed as a block design experiment using a blood oxygen level–dependent sensitive single-shot gradient echo-planar imaging (EPI) sequence^{21–23} with the following parameters: TR at 4 seconds, TE at 60 ms, matrix size of 64×64 , 32 sections, FOV at 240 mm (rectangular), voxel size at 3.75 mm^3 , intersection gap at 0.75 mm, and flip angle at 90° . Imaging sections were oriented along and parallel to the bicommissural plane²⁴ and centered to cover the entire brain.

Stimuli and Paradigm. The visual stimuli consisted of pictures of objects and were presented through MR-compatible glasses using the Presentation 0.60 software (Neurobehavioral Systems, Albany, Calif) that also provided a trigger to synchronize the scanner with the presentation computer. The pictures of objects were gray-scale–shaded images from the “Snodgrass and Vanderwart Like Objects” corpus (Bruno Rossion, Brown University, Providence, RI, and University of Louvain, Louvain, Belgium, and Gilles Pourtois, Tilburg University,

Tilburg, the Netherlands; <http://www.cog.brown.edu/~tarr/stimuli.html>) and were chosen from the semantic categories animals, fruits and vegetables, tools, and household articles. The subjects once practiced the task outside the scanner like patients do before the IOM in the operating room. For this first training session and the following 3 measurement sessions, the same visual stimuli were used. An examination cycle consisted of 2 runs with 24 active and 25 baseline phases (each 20 seconds) for each run. During 12 of the active phases, subjects had to name aloud the objects presented (naming). During the other half of the active phases, subjects again named aloud the objects presented and were required to generate additionally and overtly a noun beginning with the same letter as the object name (naming plus phonologic generation). For this generation task, subjects were instructed to avoid stereotyped responses if possible.

In both conditions, each object was presented for 4 seconds. Between the 2 alternating active blocks, one baseline block was presented, during which the subjects just had to look at random dot images (rest condition). Every run consisted of 245 scans adding up to a total length of 16 minutes per run.

Data Analysis. Functional data were analyzed using SPM99 software (Wellcome Department of Cognitive Neurology, Institute of Neurology, London, UK; <http://www.fil.ion.ucl.ac.uk/spm>) implemented in Matlab 5.2 (Mathworks, Sherborne, Mass). The first baseline at the beginning of each run was extended so that the first 5 volumes were always discarded from analysis to avoid T1-related relaxation effects. To reduce head movement artifacts, the remaining EPI volumes of each series were realigned to the first functional volume by rigid body transformation and by using the sinc interpolation method. Using the bilinear interpolation, the realigned data were normalized into the Montreal Neurologic Institute (MNI) averaged brain, which is based on 152 individual brains and spatially smoothed with a gaussian kernel of 8 mm.^{25,26} Spatial smoothing was applied to attenuate high-frequency noise, thus increasing the signal intensity-to-noise ratio. Statistical analysis was performed using the principles of the general linear model (GLM)²⁷; into the GLM, a boxcar function convolved with the hemodynamic response function was incorporated. The evaluation of the functional data was carried out with a statistical threshold of $P = .05$ corrected for multiple comparisons, and the GLM was applied to the concatenation of the 2 runs in each experiment.

To get an overview of the whole brain activation common to all of the subjects, a random-effects analysis was used over the group of 13 subjects (second-level GLM analysis using a 1-sample t test) as a first step. For both the group study and the analysis of the individual data, the Automated Anatomical Labeling tool²⁸ was used for the classification of the anatomic regions according to the coordinates of the activated voxels. This assignment is based on the MNI single-subject brain, which was spatially coregistered and normalized to the averaged MNI brain as reference. Customized software tools were used to identify and count significantly activated voxels for each single subject in and over sessions. From these data, the reproducibility coefficient for all 3 of the sessions as $r^{ijk} = 3 \times V^{ijk}_{\text{overlap}} / (V^i + V^j + V^k)$ was calculated, with V^{ijk}_{overlap} representing the number of voxels being classified as active in all 3 of the sessions and V^i , V^j , and V^k representing the numbers of voxels being classified as active in individual sessions.²⁹

This individual analysis of each subject was carried out for the following 2 contrasts: naming minus rest (naming) and naming plus generation minus rest (naming plus generation). In this report we will focus on the data found for the opercular and triangular parts of the left inferior frontal gyrus and the left insula.

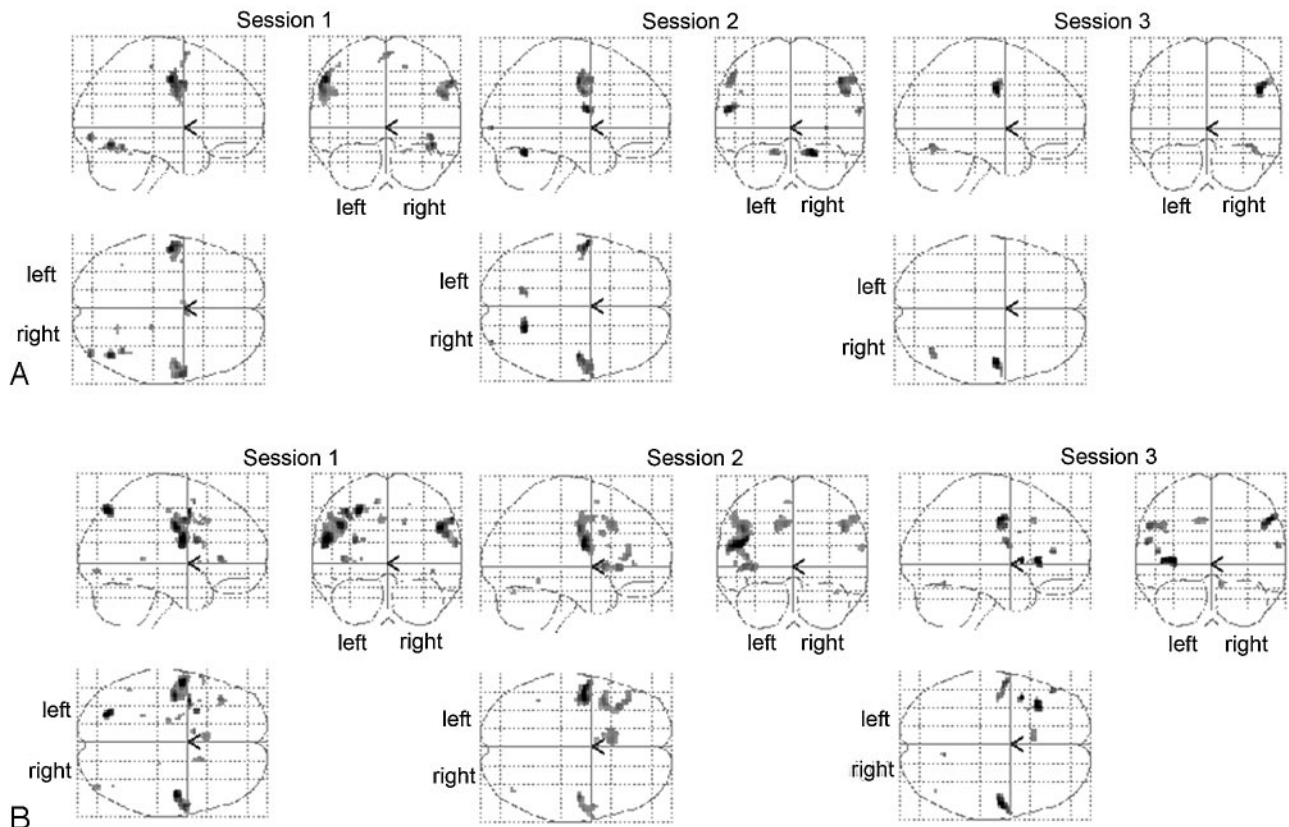


Fig 1. Glass-brain maps of the random-effects analysis. For labeling of the activated areas see Table 1. Axial and coronal view: *left*: left hemisphere; *right*: right hemisphere; sagittal view: whole-brain activation is visible.

A, Naming — rest.

B, Naming plus generation — rest.

For the visualization of all of the reproducible voxels and all of the voxels classified as active in the Broca area and the left insula we created reproducibility maps for both contrasts using customized software tools (Fig 3). To show reproducible voxels, the number of times a voxel was significantly active was counted for each subject and session and was then color coded. The possible maximum value for a single voxel could be 39 if the voxel was significantly activated in every subject ($n = 13$) and every session ($n = 3$). The voxels classified as active in any session or subject are displayed in green; thus, also those voxels are displayed, which were active only once in 1 single subject.

Experiment 2

Because possible activation differences in the Broca area between the naming and the naming plus generation condition in experiment 1 could simply be interpreted as the difference of activation in producing 1 versus 2 words, we conducted a second, exploratory experiment in which the naming plus generation condition was replaced by an overt generation task to control for this possibility. In this new condition, subjects again saw pictures of objects and had to produce 1 noun beginning with the same letter as the name of the object. The technical equipment was the same as in experiment 1. Six healthy subjects (age, 21–30 years; mean age, 25.3 years; 2 women and 4 men) were examined; these subjects had not participated in experiment 1.

Data Acquisition. The EPI sequence consisted of the same parameters as experiment 1. As in experiment 1, an examination cycle consisted of 2 runs with 24 active and 25 baseline phases (each 20 seconds) for each run. During 12 of the active phases, the same naming

task as in experiment 1 was performed (overt naming). During the other half of the active phases, subjects should generate a noun beginning with the same letter as the object name (phonologic generation); overt naming was not required. In both conditions, each object was presented for 4 seconds. Between the 2 alternating active blocks, 1 baseline block, as described in experiment 1, was always presented. Every run again consisted of 245 scans adding up to a total length of 16 minutes per run.

Data Analysis. Functional data were analyzed using SPM99 software (see experiment 1) and performing a single-subject analysis. The evaluation of the functional data was carried out with a statistical threshold of $P = .05$ corrected for multiple comparisons, and the GLM was applied to the concatenation of the 2 runs in each experiment.

Customized software tools were used to count the number of significantly activated voxels in the opercular and triangular part of the left inferior frontal gyrus for each subject individually. Voxel counting was carried out for the following 2 contrasts: naming minus rest and phonologic generation minus rest.

Results

Experiment 1

Random-Effects Analysis. The results of the random-effects analysis are summarized in Table 1 and displayed in Fig 1. For the contrast naming, significant activations in the precentral and postcentral gyri were found bilaterally, reflecting the

Table 1: Functional MR imaging activations in the group study

Variable	Orbital Part Left/Right	Triangular Part Left/Right	Opercular Part Left/Right	Insula Left/Right	Middle Frontal Gyrus Left/Right	Superior Frontal Gyrus Left/Right	Precentral Gyrus Left/Right	Postcentral Gyrus Left/Right	Rolandic Operculum Left/Right	SMA Left/Right
Naming minus rest										
Session 1							X/X	X/X		X/X
Session 2							X/X	X/X	X/X	
Session 3							/X	/X		
Naming plus generation minus rest										
Session 1		X/	X/	X/	X/	X/	X/X	X/X		X/
Session 2	X/	X/	X/	X/		X/	X/X	X/X	X/X	X/
Session 3		X/	X/	X/		X/	X/X	X/X	/X	
Variable	Middle Cingulum Left/Right	Hippocampus Left/Right	Fusiform Gyrus Left/Right	Inferior Temporal Left/Right	Superior Temporal Pole Left/Right	Inferior Parietal Lobule Left/Right	Superior Parietal Lobule Left/Right	Inferior Occipital Gyrus Left/Right	Middle Occipital Gyrus Left/Right	
Naming minus rest										
Session 1			X/X					/X		
Session 2									/X	
Session 3			/X					/X		
Naming plus generation minus rest										
Session 1	X/X	X/		X/		X/	X/	/X		
Session 2	X/			X/	X/					
Session 3	X/		/X							

Note:—X indicates at least 1 significant activation ($P < .05$, corrected); X/, left side of the brain; /X, right side of the brain.

motor components of the overt speech responses. For the contrast naming plus generation, there was consistent activation over all 3 of the sessions in both the opercular and triangular parts of the left inferior frontal gyrus, the left insula, the left middle cingulum, and the precentral and postcentral gyri bilaterally. Other brain areas were activated only inconsistently over sessions.

Single-Subject Analysis. Analysis of individual subjects in terms of voxel counts will be focused on the left opercular and triangular parts of the inferior frontal gyrus and the left insula. In these regions, activations were small in the naming condition (Table 2) and not present in every subject, indicating that this condition does not activate the regions in a robust way. Because consistent activations over all of the sessions were found in only 4 subjects in the opercular part, in no subject in the triangular part of the Broca area, and in only 2 subjects in the insula, reproducibility on a mean group level was low (Table 3). The number of active voxels decreased significantly in the left opercular part from session 2 to 3 ($P < .05$) and in the insula from session 1 to 3 and from session 2 to 3 ($P < .05$). This means that the little activation found in some subjects grew even smaller over time.

For the contrast naming plus generation, data analysis showed very different results. In this task, all of the subjects showed activations in the left opercular and triangular part, and most subjects also showed activations in the insula (Fig 2). In all of the areas, the numbers of activated voxels were much larger than in the naming condition. In the opercular part, the number of voxels decreased significantly from session 1 to 2 and from session 2 to 3 ($P < .05$). In the triangular part, the

number of voxels decreased significantly from session 2 to 3 ($P < .05$). In the insula, the number of voxels did not change significantly over sessions. Nevertheless, left-sided activation in the opercular and triangular part of the Broca area was fairly consistent over sessions, which is reflected in the number of common voxels and rather high reproducibility coefficients (Table 3). Activations in the insula, on the other hand, showed only poor reproducibility on a mean group level.

Figure 3 shows reproducibility maps for both contrasts. It may be clearly seen that reproducibility was highest in the naming plus generation condition.

Experiment 2

In experiment 2, in which only 1 word was produced overtly in both the naming and the phonologic generation condition, there was the same marked difference in the extent of activations between the 2 conditions, which was also present in experiment 1. However, for the naming condition, only a few activations were found (left opercular part: median, 35 voxels and range, 0–82 voxels; left triangular part: median, 0 voxels and range, 0–4 voxels) and generation resulted in extensive patterns of activation (left opercular part: median, 540 voxels and range, 38–760 voxels; left triangular part: median, 637 voxels and range, 38–1563 voxels). These differences were significant ($P < .05$). The results of this analysis indicate that the differences in activation between the contrast naming minus rest and the contrast naming plus generation found in experiment 1 are independent of the amount of motor output, that is, of overtly producing 1 versus 2 words.

Table 2: Median and range of the numbers of all significantly activated voxels

Brain Region	Session 1	Session 2	Session 3
Naming minus rest			
Left opercular part	6 (0–154)	5 (0–133)	0 (0–69)
Left triangular part	0 (0–335)	0 (0–53)	0 (All 0)
Left insula	0 (0–177)	0 (0–155)	0 (0–44)
Naming plus generation minus rest			
Left opercular part	549 (57–778)	508 (102–796)	282 (50–692)
Left triangular part	1006 (131–1749)	1150 (213–1772)	698 (3–1602)
Left insula	258 (0–751)	206 (0–866)	147 (0–773)

Table 3: Median and range of the numbers of common voxels between all sessions and the reproducibility coefficients between all sessions

Contrast and Regions	Median (Range)
Number of common voxels	
Naming minus rest	
Left opercular part	0 (0–42)
Left triangular part	0 (all 0)
Left insula	0 (0–13)
Naming plus generation minus rest	
Left opercular part	244 (21–523)
Left triangular part	543 (3–918)
Left insula	15 (0–376)
Reproducibility coefficients	
Naming minus rest	
Left opercular part	0.00 (0.00–0.60)
Left triangular part	0.00 (All 0 or missing value)
Left insula	0.00 (0.00 or missing to 0.13)
Naming plus generation minus rest	
Left opercular part	0.49 (0.15–0.82)
Left triangular part	0.48 (0.00–0.67)
Left insula	0.07 (0.00–0.56)

Discussion

In our study, we assessed the reproducibility (test-retest reliability) of activations in the Broca area and the left insula for an object naming and a combined naming/noun generation fMRI paradigm. Both tasks are often used in fMRI for the localization of language functions, and naming is the standard task in IOM. The aim of our study was to explore the reproducibility of activations in the Broca area and the insula within a clinical framework, in which fMRI data complement IOM or are to be validated with IOM. The choice of the stimuli sets is critical in this situation: the patient has to be familiar with the stimuli used in the operating room to avoid uncertainties with respect to stimulation results and, not less important, to reduce psychologic distress for the patient. We thus used the same stimulus sets during all 3 of the measurements, accepting the potential occurrence of priming effects.

Random-effects analysis for the naming condition did not reveal activation of the Broca area or of the insula, but individual data show that some subjects showed activations in these regions, particularly in the opercular part of the left inferior frontal gyrus. However, these activations were small and could not be reproduced in a consistent way, resulting in very low test-retest reliability. This low reproducibility may be expected in paradigms that do not yield robust activations. Different levels of analysis (group versus individual) and low test-

retest reliability may explain earlier contradictory results regarding the involvement of the Broca area in the naming process. For example, Indefrey and Levelt³⁰ described in a meta-analysis that only 5 of 9 studies reported activations in the posterior part of the frontal inferior gyrus during naming tasks. In addition, the degree of the subjects' familiarity with the stimuli is an important factor for the amount of activation found during imaging: if, as in our study, subjects had been trained to perform the task before the imaging sessions, reduced activity in the inferior frontal gyrus would be expected from the beginning because of priming effects. This reduction of activity with repeated exposure to the same stimuli has been described in several naming studies.^{16,18} In addition, van Turennout et al¹⁶ have described that the reduction of activity in the inferior frontal lobe was accompanied by an increase of activation in the left insula, which they interpret as "a form of procedural learning involving a reorganization in brain circuitry that leads to more efficient name retrieval in response to a specific object." This effect, however, seems not to be stable over time, as the poor reproducibility of activations in the insula in our experiment indicates; further studies are needed to explore the processes after priming in repeated measurements. Taken together, we do not think that naming paradigms activate the Broca area in a reproducible way, especially not in a clinical setting when the patient has to be familiar with the stimuli.

A very different picture emerged for the naming plus generation condition. All of the subjects showed activations in the opercular and triangular parts of the left inferior frontal gyrus over all of the sessions, and in nearly all of the subjects, the left insula was activated, too. In addition, these activations were much more extended than in the contrast naming. One first obvious explanation for the stronger and more consistent activation in the contrast naming plus generation would be the different motor output in the 2 conditions: whereas in the naming condition only 1 word had to be produced, the production of 2 words was required in the naming plus generation task. To test this explanation, we carried out the second experiment, in which subjects had to produce only 1 word in both conditions. Again, we found little activation in the Broca area during naming and strong activation during noun generation. It is, thus, unlikely that the difference in activation patterns in the first experiment was because of the various lengths of the required motor output.

We cannot be sure about the contribution of the naming part in the naming plus generation task to the total activation pattern of this task; it is possible that a generation task alone would have yielded similar results. Other studies using phonologic generation^{11,15} have already shown that this paradigm activates frontal areas in a reproducible way. However, the use of a generation task alone is not acceptable for IOM; it can be performed as a separate task in addition to naming,³¹ but this is time consuming. We think that our approach to the dilemma (naming gives inconsistent activations but is indispensable in IOM and generation gives consistent activations but is insufficient in IOM) of combining both tasks is a satisfactory and practical solution for clinical purposes.

The 2 tasks used in our experiment differ in other aspects, such as task switching, working memory involvement, and response selection demands; it was not the goal of our study to

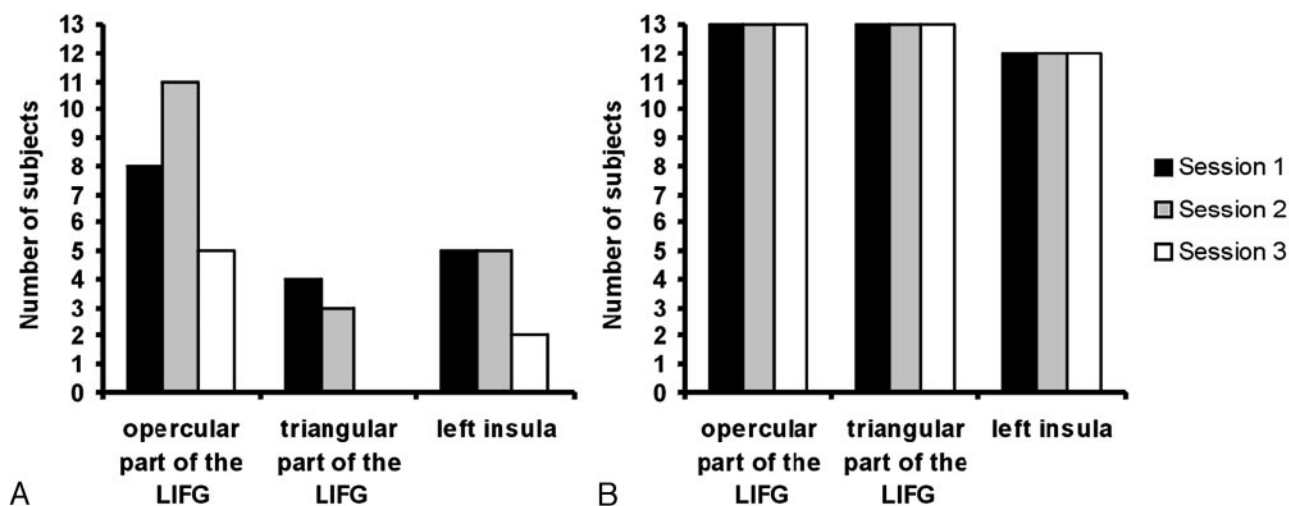


Fig 2. Number of subjects with activation in the opercular part, triangular part, and the insula of the left hemisphere. LIFG indicates left inferior frontal gyrus.

A, Naming — rest.

B, Naming plus generation — rest.

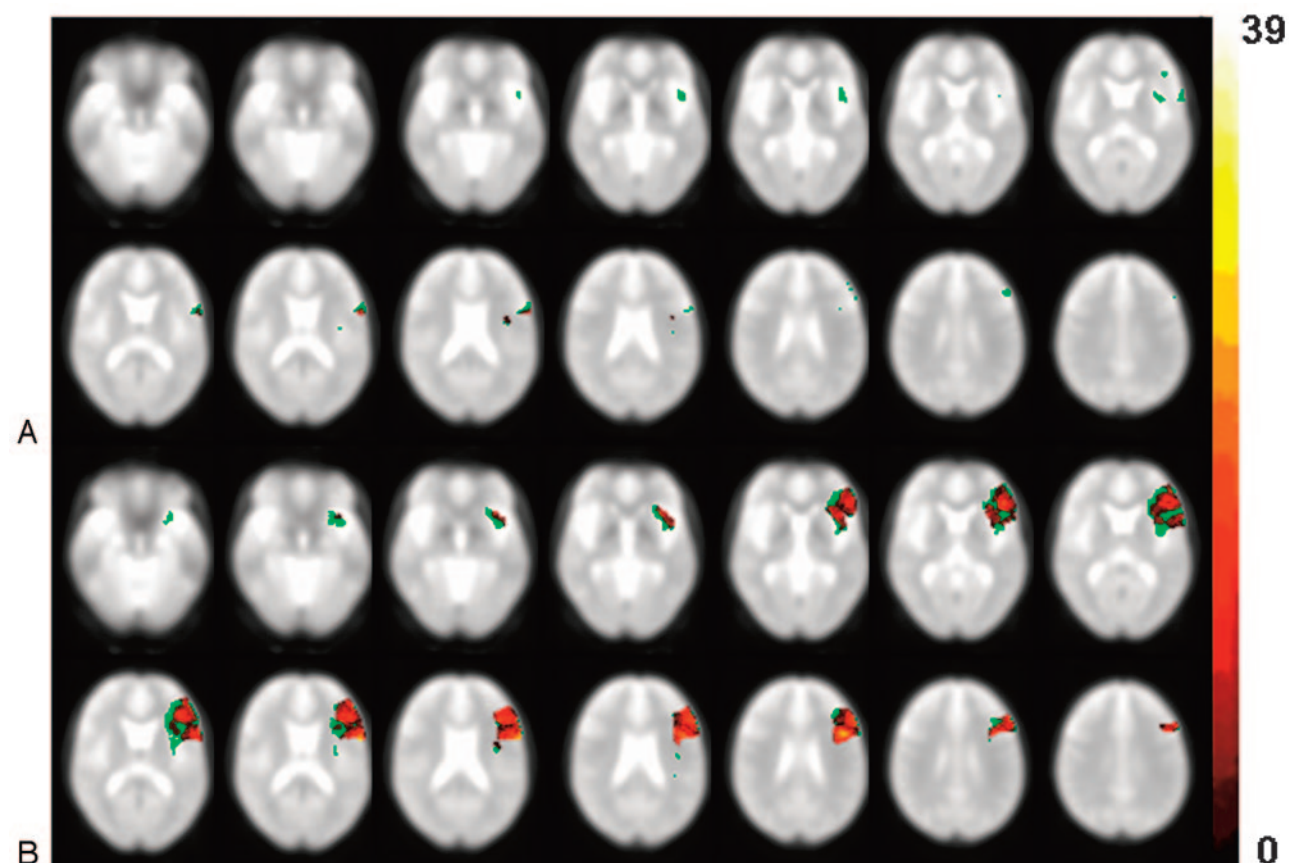


Fig 3. Reproducibility maps for the contrasts (A) naming — rest and (B) generation plus naming — rest. Voxels classified as active in any session or subject (green) and common voxels (color coded with respect to frequency of occurrence according to the color scale). Only activation in the regions of interest (Broca area and left insula) is displayed.

separately identify the neural correlates of these mechanisms. Selectional demand of the respective task, however, may be the most important aspect with regard to the involvement of the Broca area. Although they used a semantic generation task, our findings are similar to those of Etard et al,³² who found no activation of the Broca area when subjects performed a nam-

ing task. However, the Broca area was activated when subjects had to generate verbs corresponding with visual presented objects. They comment on their results as follows³²: "Absence of the Broca area activation during naming could be related to the fact that this region is engaged in the selection of semantic knowledge among competing alternatives. During naming

task the subjects had a single possible label once they identified the object, whereas they had to select one verb among several possibilities during the generation task.” The aspect of selectional demand has also been stressed by Kan and Thompson-Schill³³ within a naming paradigm. They report more activation in the left frontal inferior gyrus when subjects had to name pictures with low name agreement (high selectional demand) in contrast with the naming of objects with high name agreement (low selectional demand).

The necessity to choose between competing response alternatives is a core characteristic of generation tasks, and selection processes may be one of the basic functions in which the Broca area is essentially involved. These processes of selection can, of course, also be influenced by priming effects. Raichle et al³⁴ first demonstrated the dramatic effects of practice in a verb generation task, which resulted in a decrease of activation in the left prefrontal and cingular cortex: “In effect, the practice condition could not be distinguished from the simple repeat noun condition.” Practicing a task reduces competition among response alternatives, but selectional demand can be kept high also with repeated items as demonstrated by Thompson-Schill et al.³⁵ Activity in the left inferior frontal gyrus was reduced only in repetition trials with reduced competition; with increased competition during repetition, activity in that area increased. In our experiment, it seems probable that the instruction to not repeat responses kept selectional demands high during repeated measurements and minimized priming or practice effects. Still, the decrease in the extent of activations over sessions in the inferior frontal gyrus may represent effects of learning, but reproducibility was considerably higher than in the naming condition and similar to that reported in comparable studies.^{11,12} Brannen et al¹¹ found a proportion of repeatedly activated voxels of 37% for a phonologic generation task in patients, the volume of interest including Brodmann areas 9, 46, 44, and 45. Rutten et al¹² found a similar percentage of common voxels for a verb generation and a naming task (24% and 21%), but their volume of interest included posterior language areas, which may explain the lower values in comparison with our results. More important, however, is the outcome of combining the data of the different language tasks (a third antonym generation task was included): in this case, the percentage of overlapping voxels rose to 40%, which is why these authors strongly propagate the use of combined task analysis (CTA). “The CTA targets brain areas that relate to task performance . . . , but are not specifically associated with an individual task, thus aiming more selectively at indispensable, critical language areas than individual task analysis.”¹² We think that the melting of 2 tasks into one as in the naming plus generation condition and the use of noun generation has even further advantages: first, it is more economic and thus can also be used in the operating room where testing time is very restricted; and, second, responses can be judged as right or wrong unambiguously. In our experience in the operating room,³⁶ a generation task alone is not sufficient for IOM, because it is often difficult to judge the correctness of the response (eg, how should the patient’s response “eat” to the stimulus “singing bird” be classified: wrong answer, idiosyncrasy, culturally defined correct answer, or physiologically defined correct answer because the patient is hungry? Moreover, it is difficult even for healthy persons to

generate an appropriate response for each stimulus in a whole series without hesitations and in the temporal frame of a few seconds). Of course, with the use of combined tasks, no differentiation between linguistic subprocesses is possible, but the main goal in a clinical setting is to obtain stable patterns of activation with imaging or to reliably map eloquent sites in specific brain areas to minimize the risk of postoperative deficits. To obtain this goal, more research is needed on the topic of reproducibility of functional imaging data, especially in the cognitive domain, before the application of these data in neurosurgery. In comparison with earlier work, our data confirm that naming is not a suitable paradigm for activating the Broca area consistently and may explain earlier contradictory results. It seems to us a very crucial point to address the factor of familiarity with the stimulus material in future work. With respect to generation tasks, we could replicate the findings that a generative task component enhances reproducibility of activations in the Broca area. Our approach, in contrast to earlier reports, was custom tailored to the needs of IOM. At present, imaging data at best should be viewed as a complementary source of information other than the electrophysiologic mapping techniques. “Fair” or “good” reproducibility of data is not enough in neurosurgical decision-making. The use of combined tasks, developed for specific target brain areas, may help to improve the applicability of functional imaging in the clinical context.

Conclusion

In this study, we tested the reproducibility of fMRI data for 2 widely used language paradigms, naming and word generation. We focused our analysis on activations found in one of the classical language centers, the Broca area. Only the generation task, used in combination with naming, yielded activation data that were reproducible to a certain degree. Naming, while being a reliable and valid task in IOM of language areas by cortical stimulation, should not be used for preoperative identification of the Broca area with fMRI. Activations in this area are more reproducibly achieved in task settings with generative components.

References

1. Mesulam MM. From sensation to cognition. *Brain* 1998;121:1013–52
2. Amunts K, Weiss PH, Mohlberg H, et al. Analysis of neural mechanisms underlying verbal fluency in cytoarchitectonically defined stereotaxic space—the roles of Brodmann areas 44 and 45. *Neuroimage* 2004;22:42–56
3. Demb JB, Desmond JE, Wagner AD, et al. Semantic encoding and retrieval in the left inferior prefrontal cortex: A functional MRI study of task difficulty and process specificity. *J Neurosci* 1995;15:5870–78
4. Poldrack RA, Wagner AD, Prull MW, et al. Functional specialization for semantic and phonological processing in the left inferior prefrontal cortex. *Neuroimage* 1999;10:15–35
5. Mazoyer BM, Dehaene S, Tzourio N, et al. The cortical representation of speech. *J Cogn Neurosci* 1993;5:467–79
6. Meyer M, Friederici AD, von Cramon DY. Neurocognition of auditory sentence comprehension: event-related fMRI reveals sensitivity to syntactic violations and task demands. *Brain Res Cogn Brain Res* 2000;9:19–33
7. Stowe LA, Haverkort M, Zwarts F. Rethinking the neurological basis of language. *Lingua* 2005;115:997–1042
8. Duffau H, Lopes M, Arthius F, et al. Contribution of intraoperative electrical stimulations in surgery of low grade gliomas: a comparative study between two series without (1985–96) and with (1996–2003) functional mapping in the same institution. *J Neurol Neurosurg Psychiatry* 2005;76:845–51
9. Ojemann GA. Cortical organization of language. *J Neurosci* 1991;11:2281–87
10. Ilmberger J, Eisner W, Schmid U, et al. Performance in picture naming and word comprehension: evidence for common neuronal substrates from intraoperative language mapping. *Brain Lang* 2001;76:111–18

11. Brannen JH, Badie B, Moritz CH, et al. **Reliability of functional MR imaging with word-generation tasks for mapping Broca's area.** *AJNR Am J Neuroradiol* 2001;22:1711–18
12. Rutten GJM, Ramsey NF, van Rijen PC, et al. **Reproducibility of fMRI-determined language lateralization in individual subjects.** *Brain Lang* 2002;80:421–37
13. Fernández G, Specht K, Weis S, et al. **Intrasubject reproducibility of presurgical language lateralization and mapping using fMRI.** *Neurology* 2003;60:969–75
14. Otzenberger H, Gounot D, Marrer C, et al. **Reliability of individual functional MRI brain mapping of language.** *Neuropsychology* 2005;19:484–93
15. Mayer AR, Xu J, Pare-Blagoev J, et al. **Reproducibility of activation in Broca's area during covert generation of single words at high field: a single trial fMRI study at 4 T.** *Neuroimage* 2006;32:129–37
16. Van Turennout M, Bielamowicz L, Martin A. **Modulation of neural activity during object naming: Effects of time and practice.** *Cereb Cortex* 2003;13:381–91
17. Wagner AD, Koutstaal W, Maril A, et al. **Task-specific repetition priming in left inferior prefrontal cortex.** *Cereb Cortex* 2000;10:1176–84
18. Meister IG, Weidemann J, Foltys H, et al. **The neural correlate of very-long-term picture priming.** *Eur J Neurosci* 2005;21:1101–06
19. Buckner RL, Koutstaal W, Schacter DL, et al. **Functional MRI evidence for a role of frontal and inferior temporal cortex in amodal components of priming.** *Brain* 2000;123:620–40
20. Oldfield RC. **The assessment and analysis of handedness: the Edinburgh inventory.** *Neuropsychologia* 1971;9:97–113
21. Ogawa S, Lee TM. **Magnetic resonance imaging of blood vessels at high fields: in vivo and in vitro measurements and image simulation.** *Magn Reson Med* 1990;16:9–18
22. Ogawa S, Lee TM, Nayak AS, et al. **Oxygenation-sensitive contrast in magnetic resonance image of rodent brain at high magnetic fields.** *Magn Reson Med* 1990;14:68–78
23. Thulborn KR, Waterton JC, Matthews PM, et al. **Oxygenation dependence of the transverse relaxation time of water protons in whole blood at high field.** *Biochim Biophys Acta* 1982;714:265–70
24. Talairach J, Tournoux P. *Co-planar Stereotactic Atlas of the Human Brain.* New York: Thieme; 1988
25. Friston KJ, Ashburner J, Poline JB, et al. **Spatial registration and normalisation of images.** *Hum Brain Mapp* 1995a;2:165–89
26. Poline JB, Worsley KJ, Holmes AP, et al. **Estimating smoothness in statistical parametric maps: variability of P values.** *J Comput Assist Tomogr* 1995;19:788–96
27. Friston KJ, Holmes AP, Worsley KJ, et al. **Statistical parametric maps in functional imaging: a general linear approach.** *Hum Brain Mapp* 1995b;2:189–210
28. Tzourio-Mazoyer N, Landeau B, Papathanassiou D, et al. **Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain.** *Neuroimage* 2002;15:273–89
29. Rombouts SAR, Barkhof F, Hoogenraad FGC, et al. **Within-subject reproducibility of visual activation patterns with functional magnetic resonance imaging using multislice echo planar imaging.** *Magn Reson Imaging* 1998;16:105–13
30. Indefrey R, Levelt WJ. **The neural correlates of language production.** In: Gazzaniga MS, ed. *The New Cognitive Neurosciences.* Cambridge, Mass: MIT Press; 2000:845–65
31. Ojemann JG, Ojemann GA, Lettich E. **Cortical stimulation mapping of language cortex by using a verb generation task: effects of learning and comparison to mapping based on object naming.** *J Neurosurg* 2002;97:33–38
32. Etard O, Mellet E, Papathanassiou D, et al. **Picture naming without Broca's and Wernicke's area.** *Neuroreport* 2000;11:617–22
33. Kan IP, Thompson-Schill SL. **Effect of name agreement on prefrontal activity during overt and covert picture naming.** *Cogn Affect Behav Neurosci* 2004;4:43–57
34. Raichle ME, Fiez JA, Videen TO, et al. **Practice-related changes in human brain functional anatomy during nonmotor learning.** *Cerebral Cortex* 1994;4:8–25
35. Thompson-Schill SL, D'Esposito M, Kan IP. **Effects of repetition and competition on activity in left prefrontal cortex during word generation.** *Neuron* 1999;23:513–22
36. Herholz K, Reulen HJ, von Stockhausen HM, et al. **Preoperative activation and intraoperative stimulation of language-related areas in patients with glioma.** *Neurosurgery* 1997;41:1253–60