Further fMRI validation of the visual half field technique as an indicator of language laterality: A large-group analysis

Lise Van der Haegen, Qing Cai, Ruth Seurinck, Marc Brysbaert

Department of Experimental Psychology and Ghent Institute for Functional and Metabolic Imaging,
Ghent University, Belgium

Address for Correspondence:

Lise Van der Haegen
Department of Experimental Psychology
Henri Dunantlaan 2
B-9000 Ghent (Belgium)
E-mail: Lise.VanderHaegen@UGent.be
Tel: 32-9-264-94-29 - Fax: 32-9-264-64-96
Abstract

The best established lateralized cerebral function is speech production, with the majority of the population having left hemisphere dominance. An important question is how to best assess the laterality of this function. Neuroimaging techniques such as functional Magnetic Resonance Imaging (fMRI) are increasingly used in clinical settings to replace the invasive Wada-test. We evaluated the usefulness of behavioral visual half field (VHF) tasks for screening a large sample of healthy left-handers. Laterality Indices (LIs) calculated on the basis of the latencies in a word and picture naming VHF task were compared to the brain activity measured in a silent word generation task in fMRI (pars opercularis/BA44 and pars triangularis/BA45). Results confirmed the usefulness of the VHF-tasks as a screening device. None of the left-handed participants with clear right visual field (RVF) advantages in the picture and word naming task showed right hemisphere dominance in the scanner. In contrast, 16/20 participants with a left visual field (LVF) advantage in both word and picture naming turned out to have atypical right brain dominance. Results were less clear for participants who failed to show clear VHF asymmetries (below 20 ms RVF advantage and below 60 ms LVF advantage) or who had inconsistent asymmetries in picture and word naming. These results indicate that the behavioral tasks can mainly provide useful information about the direction of speech dominance when both VHF differences clearly point in the same direction.

Keywords: hemispheric asymmetry, interhemispheric transfer, language, laterality index, visual half field task
1. Introduction

Since Dax (1865) and Broca (1865) localized speech production in the left cerebral hemisphere, a large number of researchers have investigated the functional asymmetry of the two brain halves, which look so similar at the anatomical level. Today, language processing is still one of the most frequently investigated lateralized functions (Hugdahl & Westerhausen, 2010). It is now well-established that speech production is not always controlled by the left hemisphere, not even in healthy right-handers. Some 1-5\% of the right-handers are right language dominant or have bilateral language control (Knecht et al., 2000; Pujol, Deus, Losilla & Cadevila, 1999).

Surprisingly, the majority of left-handers are left dominant as well (against Broca’s initial assumption). Only 20-25\% of the left-handers are thought to be right language dominant or to have bilateral control. Because of the limited covariation between handedness and cerebral dominance, it is important for researchers to have a reliable measure of language lateralization. Simply comparing a group of left-handers with a group of right-handers is unlikely to provide clear findings.

Localization of language functions has been studied most intensively in patients undergoing brain surgery (e.g., to remove the seizure-causing tissue in epilepsy). For these patients, it is important to know where the language areas are, so that they can be spared (Möddel, Lineweaver, Schuele, Reinholz, & Loddenkemper, 2009). Traditionally, the Wada test was used (Wada & Rasmussen, 1960). This test consists of the injection of sodium amobarbital in the left or right internal carotid artery while the subject performs a language task such as counting aloud. The Lateralization Index (LI) is then calculated by comparing the performance after left and right injection (Binder et al., 1996). The Wada test is clearly invasive and does not provide intrahemispheric information. In recent years, a range of neuroimaging paradigms have been used as an alternative. These are non-invasive techniques such as functional Magnetic Resonance Imaging (fMRI) and magnetoencephalography (MEG). The validity of LIs based on these techniques was established by comparing them with results from the Wada test both in healthy and epileptic populations (e.g., Binder et al., 1996; Hirata et al., 2010; Jansen et al., 2006; Pirmoradi, Béland, Nguyen, Bacon & Lassonde, 2010).
Although fMRI and MEG are much less invasive than the Wada-test, they have some drawbacks for every-day laterality research. First, they are expensive and time-consuming. Second, a considerable percentage of participants do not qualify for this type of research, because they suffer from claustrophobia or because their body contains irremovable ferromagnetic matter. Finally, there may be concerns about repeated, intensive testing (e.g., to try out various manipulations or to establish psychophysical functions). Although some of the disadvantages can be overcome, behavioral LI measures would be an interesting alternative for fast, less expensive language lateralization of large groups and for repeated, intensive testing.

Surprisingly, despite decades of research only a few studies have looked directly at the validity of behavioral laterality measures by comparing them with brain imaging data (e.g., Bethmann, Templemann, De Bleser, Scheich & Brechmann, 2006; Gonzalez & Goodale, 2009; Hunter & Brysbaert, 2008a; Krach, Chen & Hartje, 2006). In addition, Hunter and Brysbaert (2008a) criticized most of these studies, because they were not well designed.

Hunter and Brysbaert (2008a) examined the visual half field (VHF) task with bilateral stimulus presentation. In this task participants fixate the middle of a screen where a fixation cross appears and after 500 ms is replaced by tachistoscopically presented parafoveal stimuli in the left visual field (LVF) and the right visual field (RVF), together with a central arrow pointing to one of them. Participants are asked to name the stimulus to which the arrow points. LIs are calculated by subtracting the mean reaction time (RT) to stimuli in RVF from the mean RT to stimuli in LVF. This method is based on the contralateral projection of visual information in humans (e.g., Bradshaw & Nettleton, 1983; Bryden, 1982). Due to the partial crossing of the human visual pathways, stimuli from LVF/RVF are sent to the right hemisphere (RH)/left hemisphere (LH) respectively. As a result, participants with left brain dominance are expected to name stimuli faster in RVF, whereas participants with right dominance are expected to name stimuli faster in LVF.\(^1\)

Hunter and Brysbaert (2008a) used two different types of stimuli: words and pictures. They obtained high correlations between the LIs of these tasks and fMRI brain activity measured

---

\(^1\) The VHF technique is usually not criticized when the stimuli are presented outside central vision (i.e., parafoveally, about 1.5° of visual angle away from the centre), although evidence motivated by the split fovea theory indicates that laterality effects can also be found in central vision (see Ellis & Brysbaert, 2010 for a review).
in a silent word generation task (for more details about the tasks, see below). The correlation was \( r = 0.63 \) for word naming and \( r = 0.77 \) for picture naming. Furthermore, Hunter and Brysbaert (2008b) claimed that the combined results of word and picture naming allowed them to predict brain dominance as measured with fMRI with 100% accuracy.

A main problem with the Hunter and Brysbaert (2008a) study, however, is that it was based on a very small sample. Only 26 left-handers took part in the behavioral VHF-tasks, and only 10 of them were selected for the fMRI task. Six of these showed clear RVF advantages in both VHF tasks and, as expected, turned out to be left dominant in the scanner. Two participants had clear LVF advantages and were confirmed as right dominant. Two final participants showed an LVF advantage in the word task, but no clear advantage in the picture naming task, and they were classified as bilateral in the fMRI task.

All in all, although Hunter and Brysbaert’s results look promising, for two reasons it would be good to have a retest on a larger group. First, it would be good to see a confirmation of the RH dominance in participants with clear LVF advantages, given that this group is rare and that there were only two hits in Hunter and Brysbaert. Second, it would be interesting to know what can be concluded of the many participants who do not show a clear VHF advantage. Do these participants have a reduced LI, or is their VHF measure simply less informative?

In the current study, we present data of 250 left-handed students who participated in the behavioral tasks (word and picture naming) and of whom 50 were scanned. We were particularly interested in those participants that deviated from the mainstream pattern (i.e., those that did not show a RVF advantage indicative of LH dominance). As a result, the majority of our findings involve participants either showing a clear LVF-advantage or a reduced VHF-asymmetry.

We adopted the method of Hunter and Brysbaert (although we had to use Dutch stimuli). We used the naming VHF task, because this task comes closest to the brain activation measured by the silent word generation task and because naming is the most lateralized function (see Ellis & Brysbaert, 2010 vs. Jordan & Paterson, 2009, for a discussion of this).  

---

2 The relationship between the laterality of the frontal language regions and the occipito-temporal language regions is an interesting topic as well (see Cai, Lavidor, Brysbaert, Paulignan & Nazir, 2008; Cai, Paulignan, Brysbaert, Ibarrola & Nazir, 2010 vs. Pinel & Dehaene, 2010), but one that is orthogonal to the subject of the present paper. If language comprehension or the relationship between production and comprehension is the focus of research, other
We also introduced some improvements. A first change was that we monitored eye-position with an eye-tracker, to directly test the extent to which VHF data may be invalidated by imperfect adherence to the central fixation instruction (Bourne, 2006). It has been claimed that participants in VHF tasks do not always fixate at the instructed location and that this may invalidate the findings (Jordan, Patching & Milner, 1998), although in a previous study we failed to find evidence for this possibility, at least when the stimuli were presented in such a way that one letter always fell on the fixation position (i.e., the so-called optimal viewing position paradigm; Van der Haegen, Drieghe & Brysbaert, 2010). Eye movement monitoring also allowed us to check to what extent participants make eye movements in a VHF-task. Although stimulus presentation time was tachistoscopic in Hunter and Brysbaert (2008a), the authors did not explicitly test the presence of fast eye movements, for which they have been criticized (Jordan & Paterson, 2009).

The second improvement of our study was that we calculated the fMRI LIs in a different way. Recent studies have pointed out that different definitions of region of interest (ROIs), statistical thresholds of brain activity, boundaries of LIs to classify dominance, and baseline conditions may result in different LIs (Abbott, Waites, Lillywhite and Jackson, 2010; Chlebus, Mikl, Brazdil, Pazourkova, Krupa & Rektor, 2007; Jansen et al., 2006; see Seghier, 2008 for a review). Wilke and Schmithorst (2006) proposed a combination of a bootstrapping procedure and histogram analysis to calculate robust LIs in neuroimaging data (see the Method section for a more detailed description of this technique). Hunter and Brysbaert used a normalized subtraction of the number of activated voxels in each hemisphere as LI, which may not always have resulted in the best estimate for each participant. In order to have more robust LIs, we used the approach of Wilke and Schmithorst to calculate a global mean LI for activation in the pars opercularis (approximately BA 44) and the pars triangularis (approximately BA 45). These two regions are the most active areas in the silent word generation task we adopted. They are known to be involved in many linguistic functions, including semantic, phonological and syntactic processes (Amunts et al., 2004; Heim, Eickhoff & Amunts, 2008), with semantic processing located more anterior than syntax and phonology. In addition, we looked at the LIs in the pars orbitalis (approximately BA 47), the insula, and the ventral premotor cortex (BA6), as these regions are tasks are more appropriate (see for example van Ettinger-Veenstra et al., 2010 and Pirmoradi et al., 2010 for a discussion of appropriate tasks to assess LIs in posterior brain areas such as Wernicke).
more and more considered as part of the language production network as well. BA 47 is cytoarchitectonically more similar to BA 45 than BA 44, because it is part of the same granular layer of the cortex (Hagoort, 2006, 2009). BA 47 has also been found to be involved in semantic processing (De Carli et al., 2007) and the processing of fine-grained temporal sequences (Vuust, Roepstorff, Wallentin, Mouridsen, & Østergaard, 2006). The insula is involved in speech motor control (Ackermann & Riecker, 2010). Finally, BA 6 has been found active when overt speech is programmed (Shuster & Lemieux, 2005).

2. Method

2.1 Behavioral VHF tasks

2.1.1 Participants

A total of 250 students from Belgian universities and higher education schools participated in this experiment (68 males, 182 females; age ranging from 17 to 30 with mean age = 19.9 years). They were recruited via advertisements on a research website, e-mail, or word of mouth. All participants were left-handed native Dutch speakers and had normal or corrected-to-normal vision. Sixty-four students earned credit points for a psychology course by participating, the others were paid. We recruited more widely than the traditional undergraduate psychology students because we needed a large number of left-handers. This group shows a higher incidence of atypical brain laterality (Knecht et al., 2000; McKeever, Seitz, Krutsch & Van Eys, 1995; Pujol et al., 1999). By not including right-handers we excluded confounds related to handedness. We also wanted to have extra variability in our sample.

Handedness was assessed with a Dutch version of the Edinburgh Handedness Inventory (Oldfield, 1971). This was combined with a questionnaire about eyedness, earedness and footedness (Porac & Coren, 1981). Participants were asked to choose a number between -3 and -1 to indicate their degree of left side preference, and a number between +1 and +3 to indicate their degree of right side preference (Brysbaert, 1994). Additionally, they performed the Miles (1930) test of eye dominance. In this test participants are asked to look at a distant target through
a small opening formed by putting together the thumbs and index fingers of both hands. Binocular viewing through the opening is alternated with monocular viewing with each eye. The eye that sees the target when it is open is selected as the dominant eye. The Miles test was administered to determine the participant’s eyedness by means of an unconscious sighting task, which controls for contamination of handedness. For example, participants may indicate a right eye preference for sighting down a rifle, simply because they prefer to have their right hand on the trigger (Porac & Coren, 1976). The questionnaire and the Miles test were administered prior to participation. Only students that reported to write and draw with their left hand were accepted. We did not include participants based on a cut-off value for left-handedness in order to obtain large variability. The Appendix shows the mean ratings reported in the questionnaires for the 50 students who also participated in the fMRI study. The data of all 250 participants are available as electronic supplementary materials.

2.1.2 Stimuli

*Picture naming.* The line drawings used in the VHF picture naming task were adopted from Hunter and Brysbaert (2008a). Five pictures were randomly presented: a boat ([boot] in Dutch), a book [boek], a house [huis], a lamp [lamp] and a star [ster]. The figure of a tree in Hunter and Brysbaert (2008a) was replaced by a symmetrical figure of a star, because of the large phonological overlap between the Dutch words *boom* (tree) and *boot* (boat). All names were monosyllabic and all stimuli were symmetrical not to favor a VHF. The five pictures are displayed in Figure 1.
**Word naming.** A list of 96 Dutch three-letter words and a list of 96 four-letter words were selected for the VHF word naming task. Half of them served as targets, half as filler words to create matched word pairs. Targets and fillers of each pair had an equal number of letters, belonged to the same word class (substantive or adjective), and were pairwise controlled for summated type bigram frequency, log frequency per million and number of neighbors in the CELEX database \((p > .40;\) Baayen, Piepenbrock, & Van Rijn, 1993). Words were selected with the Wordgen software (Duyck, Desmet, Verbeke, & Brysbaert, 2004). Similar to Hunter and Brysbaert (2008a) the three-letter and the four-letter words were matched on initial phoneme as this is the best predictor of naming latencies (Balota, Cortese, Sergent-Marshall, Spieler & Yap, 2004). Finally, targets and fillers that formed a bilateral pair never started with the same letter, so that errors could easily be detected and participants would not be able to start their response on the basis of the wrong stimulus. The full list of word pairs can be found in the electronic supplementary materials.
2.1.3 Design

There was only one repeated-measure variable in the naming tasks, namely VHF. In the picture naming task, each of the five line drawings was presented four times in combination with each of the other four pictures: once as a target in the LVF, once as a target in the RVF, once as a filler in the LVF and once as a filler in the RVF. As a result, each picture was presented 16 times in the 40 possible stimulus pairs (5*4*2). A randomized sequence of these 40 trials was repeated four times during the experiment.

In the word naming task, each of the 96 target words was presented twice with its matched filler word: once in LVF and once in RVF. Hence, the participants named 192 trials in total. The trials were divided in two blocks, in such a way that all targets were named once before the second presentation block began. Two lists were created and distributed over the participants, with a counterbalanced order of the VHF in which the trials were presented.

The VHF task has been criticized because possible confounding variables have not always been controlled for. Therefore, we paid attention to the variables that have been mentioned. First, bilateral presentation was used, so that participants were not subject to attentional biases due to the sudden appearance of a stimulus in LVF or RVF. Second, participants only had to name the stimulus the arrow pointed to. When participants have to process two stimuli and are free to choose which half field is processed first, the VHF differences are confounded by individual attention strategies (Voyer & Boles, 2007). The central arrow pointing to the target stimulus further ensured that participants were motivated to pay attention to the fixation location rather than look around (Schmuller & Goodman, 1980). Stimulus duration was limited to 200 ms, which was short enough to prevent eye movements in a paradigm with bilaterally presented stimuli (Hulme, 1979; Walker & McSorley, 2006) but long enough to make the stimulus perceptible. The stimuli in the parafovea were also large enough to make sure that participants could see them reasonably well. Each participant responded to all stimuli both in LVF and RVF, so that the individual LIs were not influenced by the stimuli presented in RVF and LVF. Finally, we presented enough stimuli to make sure that the VHF estimates had reasonably small confidence intervals (Brysbaert & d’Ydewalle, 1990).
**2.1.4 Apparatus**

Fixation locations were monitored with an Eyelink 1000 eye tracking device (SR Research, Ontario, Canada) and naming response latencies were registered with a voice key. Head movements were restricted with a chin rest and a brace at forehead height, without discomforting the participants too much when giving vocal responses. Appropriate calibration and validation were carried out with a 9-point grid. The drift between the computed fixation location based on the calibration and the current fixation was checked after each trial by displaying a single fixation target. If the drift was too large, calibration was rerun. Participants’ fixation location was measured every millisecond (sampling rate of 1000 Hz). Eye movements were recorded from the moment the trial started. For the fixation analyses only fixation locations during stimulus presentation were taken into account. Viewing was binocular throughout the experiment, but the eye-tracker only recorded the dominant eye (as assessed by the above described Miles test).

**2.1.5 Procedure**

Figure 2 illustrates the procedure of the VHF tasks. Participants were instructed to fixate the center of the screen placed at a reading distance of 60 cm from the moment the fixation cross appeared. At a viewing distance of 57 cm, an image size of 1 cm corresponds to 1 degree of visual angle. Participants were told that the fixation cross after 500 ms would be replaced by a tachistoscopically presented arrow and two stimuli (one in LVF and one in RVF). The arrow pointed in the direction of the target stimulus, which had to be named as fast and accurately as possible. The target and filler stimulus were presented for 200 ms and followed by a mask. The mask consisted of randomly oriented lines in the picture naming task, and four ASCII codes 35 (#) in the word naming task. At the onset of the mask, the arrow was replaced by the fixation cross, which remained visible until the voice key registered a response, or until 5000 ms elapsed.
The two stimuli were presented at an equal distance from the screen center. Pictures subtended a visual angle between 1.91° en 10.98°. Words were presented in Courier New font, size 15, between 1.6° (four letter words) or 2.07° (three letter words) and a fixed outer edge of 3.39°.

Each participant received the same practice trials before the beginning of the experimental blocks. These consisted of 8 randomly chosen picture pairs or 16 word pairs that did not return in the experimental phase. All participants first performed the picture naming task, as the word task was experienced as more difficult. Because we wanted the procedure to be standardized for all participants, we did not counterbalance the order of the VHF tasks but presented them in increasing order of difficulty. A limitation of this decision is that we cannot generalize the results across presentation order of the VHF tasks.

Completing the questionnaire and the informed consent form, giving instructions, setting up the eye-tracking device and presenting all practice and experimental trials took about 60 minutes.
2.2 fMRI word generation task

2.2.1 Participants

We selected a subgroup of 50 participants from the behavioral VHF tasks. All selected participants who were willing to undergo the fMRI test fulfilled all conditions to be scanned. They signed an informed consent form according to the guidelines of the Ethics Committee of the Ghent University Hospital. Because we especially wanted to find participants with atypical language dominance, all participants with a LVF advantage of at least 10 ms on both the picture and word naming task were asked to take part in the fMRI study. Twenty participants did so; participants 58, 61, 65 and 66 were also invited but refused to take part. We further included the following comparison patterns: no clear advantage on either task (N = 1), a clear RVF advantage on both tasks (N = 14), a clear RVF advantage for picture naming, but no clear advantage for word naming (N = 1), a clear LVF/RVF advantage for word naming, but not for picture naming (N = 7 and 4 respectively), and a LVF/RVF advantage for picture naming, but a RVF/LVF advantage for word naming (N = 2 and 1 respectively). Figure 3 shows the distribution of the 250 VHF differences. The height of the black bars reflects the number of participants that were selected to take part in the fMRI study.
2.2.2 Task design

A silent word generation task was used to determine language dominance. It is the task used by Hunter and Brysbaert (2008a) and others interested in hemispheric dominance (e.g., Abbott et al., 2010; Badzakova-Trajkov, Häberling, Roberts & Corballis, 2010; Knecht, Henningsen, Deppe, Huber, Ebner & Ringelstein, 1996). Participants were asked to mentally
think of as many words as possible beginning with a letter presented in the middle of the screen for 15s. Ten different letters were presented in randomized order. The baseline condition consisted of ten 15s blocks with silent repetition of the non-word *baba*. Experimental and baseline blocks were alternated with 20 rest periods of again 15s, during which a horizontal line was displayed at the screen centre. Subjects were familiarized with the method prior to scanning.

2.2.3 *Image acquisition*

Images were acquired on a 3-Tesla Siemens Trio MRI scanner (Siemens Medical Systems, Erlangen, Germany) with an 8-channel radiofrequency head coil. First, a high-resolution anatomical image was collected using a T1-weighted 3D MPRAGE sequence (TR = 1550 ms, TE = 2.39 ms, image matrix = 256 × 256, FOV = 220 mm, flip angle = 9°, voxel size = 0.9 × 0.9 × 0.9 mm³). Functional images were then obtained using a T2*-weighted gradient-echo EPI sequence. Forty axial slices covering the whole brain were acquired (TR = 2630 ms, TE = 35 ms, image matrix = 64 * 64, FOV = 224 mm, flip angle = 80°, slice thickness = 3.0 cm, distance factor = 17%, voxel size = 3.5 * 3.5 * 3 mm³).

2.2.4 *Data analysis*

Data analysis was performed using SPM5 software (Wellcome Trust Centre for Neuroimaging, London, UK). The first four acquired images were eliminated for each participant, due to the stabilizing of the magnetic field. Data preprocessing consisted of (1) slice time correction because slices were acquired in an interleaved way; (2) realignment using rigid body transformations to correct for movement artifacts; (3) coregistration of the anatomical image to mean functional image; (4) normalization to the Montreal Neurological Institute (MNI) T1 template; and finally (5) spatial smoothing with an isotropic Gaussian Kernel (8 mm full width at half-maximum). Data-analysis was done by using the general linear model (GLM) for modeling the experimental (target letter) and control (non-word *baba*) condition with a boxcar function, convolved with a canonical haemodynamic (BOLD) response function.
3. Results

3.1 LI calculations

Prior to LI calculation for the VHF studies, some elementary data cleaning was done. The following trials were excluded: naming corrections or errors (Picture: 3.5%, Word: 10.0%), voice key failures (Picture: 0.7%, Word: 1.2%), data loss due to eye-tracker recalibration (0.2% in each task), RTs less than 200 ms or greater than 1500 ms (Picture: 1.4%, Word: 1.6%), and for the remaining trials latencies above/below 2.5 SDs from the participant's mean RT (Picture: 2.5%, Word: 2.6%).

LI s were calculated in two different ways. Analysis1 included all trials, whereas Analysis2 took into account the quality of central fixation. This analysis excluded all trials on which the subject initiated a saccade in the 200 ms stimulus presentation period (Picture: 1.8%, Word: 17.2%) or fixated more than .5° to the left or the right of the screen center (Picture: 16.4%, Word: 14.0%). Eye movements of 11/250 participants could not be registered because their pupil was not clearly visible (indicated by * in the Appendix and in the supplementary materials). The data of these participants were all treated as unreliable fixations outside the critical region of 1° in the above described data trimming.

All in all, the behavioral LIs in Analysis 1 were based on 91.9% of the trials for picture naming and on 84.4% of the trials for word naming. In Analysis 2, they were based on 73.5% (Picture) and 55.0% (Word) of the data. LIs of less than -10 ms were considered as evidence for a LVF/RH advantage, LIs of more than 10 ms were considered as evidence for a RVF/LH advantage. Values in between were classified as unclear. On the basis of these criteria, 24 participants showed a clear LVF advantage on both tasks, 143 showed two clear RVF advantages, and the remaining showed either a mixed pattern (N = 81) or no clear VHF advantage (N = 2). Overall, the picture and word VHFs correlated significantly, but not very highly with each other ($r = .45$, $p < .001$, both in Analysis1 and Analysis2).

Because the intercorrelation of measures depends on the reliability of the individual measures, we calculated the split-half reliability by correlating the first and second half of each test and attenuating the correlation for length with the Spearman-Brown formula ($r_{attenuated} = (2 * r)$
Reliability of LI index in the picture naming VHF task was .79 in Analysis1 and .73 in Analysis2; for the word VHF tasks the values were .91 and .85 respectively.

Finally, we compared the LIs of Analysis1 and Analysis2 in the 239 participants whose eye movements could be measured, to investigate to what extent the findings in a VHF-task were invalidated by inadequate fixation control. A first informative finding was the very high correlation between the LIs of both analyses: $r = .98, p < .001$ for picture naming and $r = .94, p < .001$ for word naming. When we looked at our initial classification of the participants (clear LVF or RVF advantage, no clear advantage), we saw that 16 participants or 6.7% got a different classification in the picture naming task and 23 or 9.6% in the word naming task if we used the data of Analysis2 rather than those of Analysis1. Note that Analysis2 was based on very strict selection criteria: A trial was invalid from the moment the eye-tracking device registered an eye movement, regardless of the position or duration of the resulting fixation. When we made the criterion of an eye movement less strict and defined it as no eye movement on the parafoveal stimulus instead of no eye movement at all, the VHF classification changed for only 2.6% of the participants in the picture naming task and 5.7% in the word naming task. In other words, for 95% of the participants eye movement control did not have added value. The small extra value of fixation control was confirmed when we looked at the correlations between the VHF differences and the fMRI LIs [$r = .66, p < .001$ and $r = .67, p < .001$ in Analysis1 (N=50) vs. $r = .65, p < .001$ and $r = .64, p < .001$ in Analysis2 (N=49) for the picture and word VHF respectively].

As indicated above we used a different approach than Hunter and Brysbaert (2008a) for the fMRI LIs. Instead of taking a certain statistical threshold to calculate the normalized difference of number of activated voxels in each hemisphere, we used the LI Toolbox 1.02 provided by Wilke and Lidzba (2007). For each region on which the LI scores were based, 20 equally sized steps from 0 to the maximum t-value were taken as thresholds. At each level, 100 bootstrap resamples with a resample ratio of $k = 0.25$ were taken in the left and right investigated area. Then, all 10 000 possible LI combinations were calculated but only the central 50% of data were kept in order to exclude statistical outliers. In the last step, a weighted mean LI for each individual was calculated with higher thresholds receiving a higher weight. A more detailed description of this procedure can be found in Wilke and Schmithorst (2006).
Individual fMRI LIs were calculated from the activation in the areas formed by the pars opercularis (approximately BA44) and the pars triangularis (approximately BA45) together, in the pars opercularis and pars triangularis separately, in the pars orbitalis (approximately BA 47), in the insula, and in the precentral cortex (according to the AAL template; Tzourio-Mazoyer et al., 2002). All six resulting LIs per participant can be found in the supplementary materials. The analyses below are based on the areas classically seen as Broca's area, namely the combination of BA 44 and 45, because these showed the highest correlations with the VHF data. This measure is included in the Appendix as well. FMRI LI values range from -1 (when there are only voxels active in the right hemisphere) to +1 (when there is only signal in the left hemisphere) as they reflect the normalized difference of activated voxels in the left and the right hemispheres. When running the study we used the following criteria to classify participants: Participants with LI < -.60 were classified as RH dominant (N = 20), those with LI >+.60 as LH dominant (N = 25), and those with LI between −.60 and +.60 as bilateral (N = 5).

3.2 Comparison of LIs based on VHF and fMRI
Figure 4 shows the correlation between LIs based on the VHF tasks (Analysis 2; see description in 3.1) and LIs based on the fMRI task (BA44+45). The VHF data vary along the y-axis (two data points per participant), the fMRI data along the x-axis. The data of Participant8 are not included, as there were no eye-monitoring data for this right dominant participant. Panel A shows the data for all 49 participants; Panel B shows the results for the participants with consistent VHF asymmetries in word and picture naming (N = 34). The upper right and lower left quadrant include data indicative of respectively LH/RH dominance both in VHF and fMRI.

It is clear from Figure 4 that fMRI makes a much sharper distinction between LH and RH dominance than the VHF task. There is a distinct gap between both groups, with only two or three participants falling in-between. In contrast, the transitions from LVF advantage to RVF advantage in the VHF tasks are much more continuous, with an unpredictable relationship to the fMRI outcome in the region from -60 ms to +25 ms. All participants with a RVF advantage of more than 25 ms were classified as LH dominant in the scanner, and all but one participant with a LVF advantage of more than 60 ms were classified as RH dominant. Participants with VHF asymmetries between -60 and +25, however, could go either way in the scanner. A comparison between Panel A and Panel B shows that this was particularly true for participants who showed opposite VHF advantages for word and picture naming.

Overall, the fMRI LIs correlated positively with both the picture ($r = .65, p < .001$) and word ($r = .64, p < .001$) naming LIs. As expected on the basis of Figure 4, the correlations were even higher when only the consistent participants were taken into account (Panel B; picture: $r = .76, p < .001$; word: $r = .74, p < .001$). It is clear that a stricter threshold than +/- 10 ms should be taken for clear-cut classifications, but the current data show that the VHF tasks are a useful screening tool for laterality research.

A stepwise multiple regression analysis with four predictor variables (word VHF and picture VHF according to Analysis1 and Analysis2) returned significant effects for word VHF ($t(46) = 3.14, p < .01$) and picture VHF ($t(46) = 2.75, p < .01$) according to Analysis1, and no further contribution of word VHF and picture VHF according to Analysis 2. Apparently, the addition of eye fixation control was not an asset for better prediction of brain dominance as determined with fMRI (see the Appendix for the raw data).
3.3 Correlations with the questionnaire data of lateral preferences

We also correlated the VHF and fMRI LIs with the laterality indices based on the questionnaire data. None of the questionnaire preferences correlated significantly with the behavioral data or with the LI from BA44 and 45 (ps > .11), although footedness seemed stronger for the right dominant group as assessed with fMRI (mean = -1.8) compared to the left dominant group (mean = -0.8). Other authors also reported higher correlations with footedness than with other variables (Day & Macneilage, 1996; Searleman, 1980), although Brysbaert (1994) reported a higher correlation between language laterality and earedness. When interpreting the present null-effects it is important to keep in mind that only left-handers were tested, which seriously reduced the range of laterality indices in the questionnaire data.

4. Discussion

We examined the usefulness of VHF tasks to assess language laterality in a large sample of left-handed, healthy participants (N = 250). All participants took part in two VHF tasks (word naming and picture naming) and the participants we thought most likely to have atypical language dominance were invited to take part in an fMRI validation study, together with a control group of 14 participants with a consistent RVF advantage on both tasks. We additionally examined the influence of saccades and imprecise eye fixation positions in VHF tasks. The following were the main findings.

First, it is clear that both VHF tasks can be used to screen participants for atypical language laterality. Chances of finding such a laterality pattern are much higher for participants with LVF advantages than for participants with RVF advantages. Although it is possible that we would have found a participant with clear RVF advantage in the VHF tasks and RH dominance in the scanner if we had scanned all 250 participants, our data strongly suggest that such occurrences would be very rare (Figure 4; see also Hunter & Brysbaert, 2008a). In contrast, of the 20 participants with consistent LVF advantages we scanned, 16 turned out to have atypical
dominance (i.e., a hit rate of 80%). The fact that the classification was better for participants with consistent VHF advantages than for participants with inconsistent advantages indicates that the combination of the word and picture naming VHF task was worthwhile. We thus recommend using the picture and word VHF as a combined laterality indicator.

The main limitation of the VHF tasks is what to do with participants not showing a clear VHF asymmetry and participants with inconsistent VHF asymmetries. As for the participants with reduced VHF asymmetries, it is not the case that they also have reduced laterality in the scanner (based on the LIs of BA 44 and 45, which had the highest correlations with the VHF tasks). Rather they seem to divide into a group with LH dominance and a group with RH dominance. Further testing is also needed to have more information about participants with opposite VHF asymmetries in the word and picture naming task, because financial constraints prevented us from fully testing them. We were able to test only two participants with a LVF advantage in word naming and a RVF advantage in picture naming (P22, P43), one of whom turned out to be bilateral in the scanner (P22: word -49 ms, picture +37 ms, fMRI -.23), and one LH dominant (P43: word -16 ms, picture +88 ms, fMRI +.89). Of the two participants with a LVF advantage in picture naming and a RVF advantage in word naming, one turned out to be RH dominant (P19: word +10 ms, picture -38 ms, fMRI -.65) and one LH dominant (P46: word +38 ms, picture -42 ms, fMRI +.93).

On the one hand, these deviating patterns may point to differences in laterality patterns between brain regions responsible for word reading and speech production in a subset of participants (as argued by Pinel & Dehaene, 2010). On the other hand, they could also be due to the fact that VHF tasks and fMRI tasks use different dependent variables (RTs vs. BOLD signal) or to the fact that although the tasks are similar they are nevertheless different paradigms (stimulus naming vs. silent word generation).

The fuzzy boundary between LH and RH dominant participants in VHF measures means that researchers can use various criteria to select their participants, depending on the constraints under which they are working. If they have easy access to a large pool of lefthanders, but difficult access to an fMRI scanner, they are advised to include only those participants who show an LVF advantage both in word and picture naming. Alternatively, if access to large groups of participants is a problem whereas scanning costs are not prohibitive, all LIs smaller
than +20 ms become interesting, because this is where we found our RH dominant participants (Figure 4).

The addition of eye movement control to the VHF tasks did not have additional value in the present study. The corrected VHF asymmetries did not differ much from the uncorrected ones and, more importantly, did not correlate more with the fMRI validation data. This agrees with Van der Haegen et al.’s (2010) conclusion that deviations in the fixation position are noise, rather than systematic biases that invalidate the conclusions, as argued by Jordan and colleagues (Jordan et al., 1998; Jordan & Paterson, 2009). In this respect, it is important to keep in mind that our study already contained a fixation incentive in the form of the central arrow pointing to the target stimulus to be named. Results may be different if participants have no incentive at all to properly look at the fixation position when the trial starts. Another way to put control on the participants’ fixation behavior is to add a secondary task, such as naming briefly presented digits at random intervals (Van der Haegen, Brysbaert & Davis, 2009). The fact that strict eye movement control is not needed for valid laterality research, is interesting because it takes away much of the burden for the participants and also makes the testing more mobile. Without the need for an eye-tracking device, larger (left-handed) samples can be tested for screening under more comfortable circumstances. In addition, it must not be forgotten that eye-tracking results in the exclusion of potentially interesting participants. For instance, we had to decline participants with strong glasses and even then we had problems to monitor the eyes of 11 out of 250 participants (including one who turned out to be RH dominant; P8).

Finally, for the interpretation of our findings it is important to keep in mind that only left-handers were tested. It will be interesting to see how a similar group of right-handers perform on our battery of tasks and measures. Given that less than 5% of them are expected to be right language dominant, we would expect very few participants to show a clear LVF advantage both in word and picture naming. However, of these we would expect an equally high percentage (80%) to be right dominant in the word generation task. The most interesting subgroups arguably would be those with reduced VHF asymmetries and inconsistent VHF asymmetries: Would they all be left-dominant given the prevalence of this type within the right-handers or would they divide in two subgroups like the lefthanders?
References


Brysbaert, M. (1994). Lateral preferences and visual field asymmetries: Appearances may have been overstated. *Cortex, 30*, 413-429.


Appendix

Data from the 50 participants who participated in both the behavioral VHF tasks and the fMRI silent word generation task. The various columns include respectively: mean scores for handedness, earedness, eyedness, footedness, and an overall sidedness score (as reported in the questionnaires); the mean RT difference between LVF and RVF on the picture and the word naming task (both when uncorrected and corrected for imperfect fixation positions); and the fMRI Laterality Index for BA44+45. We also include the initial assessments we gave to the participants on the basis of the data. A VHF advantage of at least 10 ms was considered as evidence for a reliable VHF difference; advantages below this criterion were classified as “unclear”. fMRI LIs between -.60 and +.60 were considered as evidence for bilaterality. Participants are ordered and numbered according to fMRI LI for the first 50 participants, and according to the picture VHF difference for the remaining 200 participants (who did not take part in the fMRI task). Participants of whom no eye-tracking could be collected (e.g. when the eyelid or eyelashes are too close to the pupil, the eye-tracker receives insufficient contrast to localize the pupil position) are indicated by an asterisk.
<table>
<thead>
<tr>
<th>Subject</th>
<th>MEAN SCORES QUESTIONNAIRE</th>
<th>PICTURE VHF</th>
<th>WORD VHF</th>
<th>SILENT WORD GENERATION</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>-2.2</td>
<td>-1.8</td>
<td>-2.0</td>
<td>-2.5</td>
</tr>
<tr>
<td>2</td>
<td>-2.3</td>
<td>-3.0</td>
<td>-3.0</td>
<td>-3.0</td>
</tr>
<tr>
<td>3</td>
<td>-3.2</td>
<td>-2.8</td>
<td>-2.8</td>
<td>-1.5</td>
</tr>
<tr>
<td>4</td>
<td>-2.0</td>
<td>-1.5</td>
<td>-1.5</td>
<td>-2.0</td>
</tr>
<tr>
<td>5</td>
<td>-1.3</td>
<td>-2.3</td>
<td>-2.3</td>
<td>-1.9</td>
</tr>
<tr>
<td>6</td>
<td>-3.0</td>
<td>-3.0</td>
<td>-3.0</td>
<td>-3.0</td>
</tr>
<tr>
<td>7</td>
<td>-3.0</td>
<td>-3.0</td>
<td>-3.0</td>
<td>-3.0</td>
</tr>
<tr>
<td>8*</td>
<td>-3.0</td>
<td>-3.0</td>
<td>-3.0</td>
<td>-2.3</td>
</tr>
<tr>
<td>9</td>
<td>-2.8</td>
<td>-2.5</td>
<td>-2.5</td>
<td>-2.6</td>
</tr>
<tr>
<td>10</td>
<td>-2.4</td>
<td>-3.0</td>
<td>-3.0</td>
<td>-3.0</td>
</tr>
<tr>
<td>11</td>
<td>-3.0</td>
<td>-1.5</td>
<td>-1.5</td>
<td>-2.5</td>
</tr>
<tr>
<td>12</td>
<td>-2.9</td>
<td>-1.5</td>
<td>-1.8</td>
<td>-2.3</td>
</tr>
<tr>
<td>13</td>
<td>-1.9</td>
<td>-1.8</td>
<td>0.8</td>
<td>-1.3</td>
</tr>
<tr>
<td>14</td>
<td>-2.1</td>
<td>-6.8</td>
<td>0.8</td>
<td>-0.8</td>
</tr>
<tr>
<td>15</td>
<td>-3.0</td>
<td>-3.0</td>
<td>-3.0</td>
<td>-3.0</td>
</tr>
<tr>
<td>16</td>
<td>-3.0</td>
<td>-2.0</td>
<td>-2.0</td>
<td>-1.0</td>
</tr>
<tr>
<td>17</td>
<td>-2.5</td>
<td>-3.0</td>
<td>-1.5</td>
<td>-2.0</td>
</tr>
<tr>
<td>18</td>
<td>-2.9</td>
<td>-2.8</td>
<td>-2.8</td>
<td>-2.8</td>
</tr>
<tr>
<td>19</td>
<td>-0.5</td>
<td>0.3</td>
<td>2.0</td>
<td>0.4</td>
</tr>
<tr>
<td>20</td>
<td>-2.8</td>
<td>-2.0</td>
<td>-2.5</td>
<td>-2.5</td>
</tr>
<tr>
<td>21</td>
<td>-2.6</td>
<td>-2.3</td>
<td>-3.0</td>
<td>-2.0</td>
</tr>
<tr>
<td>22</td>
<td>-1.5</td>
<td>-1.0</td>
<td>-0.3</td>
<td>0.5</td>
</tr>
<tr>
<td>23</td>
<td>-1.8</td>
<td>-3.0</td>
<td>-1.5</td>
<td>-1.6</td>
</tr>
<tr>
<td>24</td>
<td>-1.8</td>
<td>-6.5</td>
<td>-3.0</td>
<td>-2.5</td>
</tr>
<tr>
<td>25</td>
<td>-2.7</td>
<td>-2.3</td>
<td>-3.0</td>
<td>-1.3</td>
</tr>
<tr>
<td>26</td>
<td>-2.7</td>
<td>-2.3</td>
<td>-1.0</td>
<td>-3.0</td>
</tr>
<tr>
<td>27</td>
<td>-2.9</td>
<td>-2.0</td>
<td>-3.0</td>
<td>-1.5</td>
</tr>
<tr>
<td>28</td>
<td>0.5</td>
<td>1.5</td>
<td>0.8</td>
<td>0.8</td>
</tr>
<tr>
<td>29</td>
<td>-3.0</td>
<td>-3.0</td>
<td>-3.0</td>
<td>-3.0</td>
</tr>
<tr>
<td>30</td>
<td>-2.8</td>
<td>-3.0</td>
<td>-2.8</td>
<td>-2.8</td>
</tr>
<tr>
<td>31</td>
<td>-2.1</td>
<td>-2.8</td>
<td>3.0</td>
<td>-0.4</td>
</tr>
<tr>
<td>32</td>
<td>-1.1</td>
<td>0.8</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>33</td>
<td>-2.4</td>
<td>0.8</td>
<td>-2.8</td>
<td>-1.8</td>
</tr>
<tr>
<td>34</td>
<td>-2.8</td>
<td>-1.8</td>
<td>-1.8</td>
<td>-2.3</td>
</tr>
<tr>
<td>35</td>
<td>-1.8</td>
<td>-1.3</td>
<td>-2.0</td>
<td>-1.9</td>
</tr>
<tr>
<td>36</td>
<td>-3.0</td>
<td>1.3</td>
<td>3.0</td>
<td>1.0</td>
</tr>
<tr>
<td>37</td>
<td>-2.3</td>
<td>-2.0</td>
<td>-2.8</td>
<td>-1.3</td>
</tr>
<tr>
<td>38</td>
<td>-2.4</td>
<td>-3.0</td>
<td>-2.8</td>
<td>-2.7</td>
</tr>
<tr>
<td>39</td>
<td>0.6</td>
<td>-1.8</td>
<td>-2.5</td>
<td>0.3</td>
</tr>
<tr>
<td>40</td>
<td>-2.6</td>
<td>-1.3</td>
<td>2.0</td>
<td>-1.3</td>
</tr>
<tr>
<td>41</td>
<td>-2.6</td>
<td>-2.0</td>
<td>-3.0</td>
<td>-2.6</td>
</tr>
<tr>
<td>42</td>
<td>-2.2</td>
<td>0.5</td>
<td>-2.0</td>
<td>-0.5</td>
</tr>
<tr>
<td>43</td>
<td>-3.0</td>
<td>0.0</td>
<td>3.0</td>
<td>-0.6</td>
</tr>
<tr>
<td>44</td>
<td>-2.6</td>
<td>-1.5</td>
<td>-1.3</td>
<td>-2.0</td>
</tr>
<tr>
<td>45</td>
<td>-2.9</td>
<td>-1.5</td>
<td>-2.3</td>
<td>-2.4</td>
</tr>
<tr>
<td>46</td>
<td>-3.0</td>
<td>-2.5</td>
<td>-3.0</td>
<td>-1.3</td>
</tr>
<tr>
<td>47</td>
<td>-2.8</td>
<td>-2.0</td>
<td>-2.0</td>
<td>-2.5</td>
</tr>
<tr>
<td>48</td>
<td>-1.5</td>
<td>-2.0</td>
<td>1.5</td>
<td>-0.5</td>
</tr>
<tr>
<td>49</td>
<td>-2.2</td>
<td>-2.3</td>
<td>-1.8</td>
<td>0.5</td>
</tr>
<tr>
<td>50</td>
<td>-2.9</td>
<td>-2.3</td>
<td>-2.0</td>
<td>-2.5</td>
</tr>
</tbody>
</table>

R C V T H O F E L I D M A N G T E C C e e e e e e