Language processing is strongly left lateralized in both sexes
Evidence from functional MRI

Julie A. Frost,[1] Jeffrey R. Binder,[1,2] Jane A. Springer,[1] Thomas A. Hammeke,[1] Patrick S.F. Bellgowan,[1] Stephen M. Rao[1,2] and Robert W. Cox[3]

Departments of [1]Neurology, [2]Cellular Biology and Anatomy and [3]Biophysics Research Institute, Medical College of Wisconsin, Milwaukee, Wisconsin, USA

Correspondence to: Jeffrey R. Binder, MD, Department of Neurology, Medical College of Wisconsin, 9200 West Wisconsin Avenue, Milwaukee, WI 53226, USA
E-mail: jbinder@post.its.mcw.edu

Summary

Functional MRI (fMRI) was used to examine gender effects on brain activation during a language comprehension task. A large number of subjects (50 women and 50 men) was studied to maximize the statistical power to detect subtle differences between the sexes. To estimate the specificity of findings related to sex differences, parallel analyses were performed on two groups of randomly assigned subjects. Men and women showed very similar, strongly left lateralized activation patterns. Voxel-wise tests for group differences in overall activation patterns demonstrated no significant differences between women and men. In further analyses, group differences were examined by region of interest and by hemisphere. No differences were found between the sexes in lateralization of activity in any region of interest or in intrahemispheric cortical activation patterns. These data argue against substantive differences between men and women in the large-scale neural organization of language processes.

Keywords: sex differences; language; lateralization; functional MRI

Abbreviations: fMRI = functional MRI; MANOVA = multivariate analysis of variance

Introduction

Numerous studies report that women, on average, have slightly better verbal skills than men. Although the magnitude of this sex difference is small when all language measures are combined, tests of speech production and verbal fluency show clear differences favouring women (for a review, see Halpern, 1992). Sex differences on language tasks involving vocabulary, verbal analogies and reading comprehension are less consistent and may vary with age (Hyde and Linn, 1988; Clarke *et al.*, 1990). Several studies even showed a slight male advantage on tests involving comprehension of verbal analogies (Hyde and Linn, 1988).

These small sex-related differences in ability have led to a great deal of interest in understanding their neurophysiological basis. One unresolved question is whether sex-related differences in brain function arise from genetic/hormonal sources, from environmental influences or from some interaction of these factors (Maccoby and Jacklin, 1974; Halpern, 1992). Apart from this 'nature versus nurture' question, however, is the more tractable problem of characterizing and quantifying the neurophysiological differences themselves. At least three types of sex-related neurophysiological differences are possible and could exist in various combinations.

One possibility is that sex-related differences exist at a microscopic level, involving differences in connectivity, neuronal density or synaptic efficiency. Such factors could account for ability differences even in the absence of large-scale differences in functional organization. This hypothesis is supported by a recent finding of sex differences in neuronal density in brain regions thought to be involved in language function (Witelson *et al.*, 1995). Indirect support comes from several other sources, including gross morphometric, lesion, behavioural and functional mapping studies that show no evidence for sex differences in large-scale organization or hemispheric lateralization of language functions (Brust *et al.*, 1976; De Renzi *et al.*, 1980; Kertesz and Sheppard, 1981; Kertesz, 1982; Warrington *et al.*, 1986; Kertesz *et al.*, 1987; Oppenheim *et al.*, 1987; Simon and Sussman, 1987; Byne *et al.*, 1988; Damasio *et al.*, 1989; Kertesz and Benke, 1989; Seth-Smith *et al.*, 1989; Allen *et al.*, 1991; Ashton and McFarland, 1991; Habib *et al.*, 1991; Aboitiz *et al.*, 1992; Buckner *et al.*, 1995; Price *et al.*, 1996; Jancke *et al.*, 1997). In two recent PET studies, for example, no significant differences between men and women were shown on func-

tional activation maps produced during various language tasks (Buckner *et al.*, 1995; Price *et al.*, 1996). One of these negative studies employed a word stem completion task which was compared with visual fixation, and a verb generation task which was compared with noun reading (Buckner *et al.*, 1995). In this study verbal fluency processes that should maximally distinguish between women and men were thus engaged (Halpern, 1992).

A second possibility (not exclusive of the first) is that macroscopic differences in brain morphology, intra-hemispheric topography or interhemispheric lateralization contribute to sex differences in verbal abilities. In contrast to the previously cited research, a number of studies reported sex-related differences in regional brain size (Witelson, 1989; Steinmetz *et al.*, 1992; Witelson and Kigar, 1992; Clarke and Zaidel, 1994; Kulynych *et al.*, 1994; Harasty *et al.*, 1997), patterns of aphasia after brain lesion (Lansdell, 1962; Lansdell and Urbach, 1965; McGlone, 1977, 1978; Kimura, 1983; Butler, 1984), and language lateralization determined by non-invasive techniques (Kinsbourne and Cook, 1971; Lake and Bryden, 1976; McGlone, 1980; Lewis and Christiansen, 1989; Shaywitz *et al.*, 1995). In contrast to the two negative PET studies, Shaywitz *et al.* (1995) reported large sex differences in lateralization of activation using functional MRI (fMRI) during a phonological processing task in which subjects determined whether two printed non-words rhyme. When this task was compared with non-linguistic visual control tasks involving consonant letter string matching and line orientation matching, women showed symmetric activation of the frontal lobes during the phonological task, whereas men showed left-lateralized activation. Interpreting the observed activation as indicative of phonological processing, the authors suggested that women and men have differently organized phonological systems, possibly accounting for sex differences observed in some studies of phonological processing (Lukatela *et al.*, 1986).

A third possibility is that there are large-scale sex differences in the neural organization of language that are unrelated to behavioural capacity. In a more extensive description of the Shaywitz *et al.* (1995) fMRI study, Pugh *et al.* (1996) reported that men and women show large differences in activation across a variety of different language tasks and task comparisons. Sex effects were measured during the phonological and consonant letter string matching tasks used by Shaywitz *et al.* (1995) as well as during a lexical–semantic task that required subjects to decide if two words belonged to the same semantic category. For all language tasks and all task combinations, men showed greater leftward asymmetry of activation in the frontal lobe compared with women. That this difference showed no specificity for a particular task suggests that it may represent a general effect of sex on many language processing components. Men in this study showed increased activation in bilateral visual association and temporal lobe regions during the lexical–semantic task compared with the phonological task, whereas women activated these areas equally during the two tasks.

These results suggest that men and women differ both in terms of lateralization of language processes and in the degree of overlap between phonological and lexical–semantic systems. Women and men typically do not show substantive behavioural differences in lexical–semantic processing (Halpern, 1992), and no performance differences between sexes were observed by Pugh *et al.* (1996) on any of the tasks used. These findings thus suggest that men and women carry out identical language processes with the same degree of functional capacity using very differently organized brain systems.

To confirm the findings of Pugh *et al.* (1996), and to address the continuing uncertainty over whether large-scale sex differences in language organization exist, we designed the present fMRI study of sex effects on functional activation during a language task that requires subjects to determine if heard words belong to specified semantic categories. This task was contrasted with a non-linguistic auditory control task. This task combination results in reproducible left-lateralized activation in healthy right-handers (Binder *et al.*, 1997) and produces language lateralization patterns that are strongly correlated with language lateralization results from the Wada test in epilepsy patients (Binder *et al.*, 1996*a*). This task combination, which contrasts an auditory lexical–semantic task with a non-linguistic control, is conceptually similar to the Pugh *et al.* (1996) comparison in which a visual lexical–semantic task was contrasted with a non-linguistic line orientation task. Because this comparison produced significant sex differences in brain activation patterns in the Pugh *et al.* (1996) study, it was expected that such differences would be observed in the present investigation.

By including a larger number of subjects (50 men and 50 women) than in previous functional imaging studies, we hoped to attain greater statistical power to detect subtle differences in activation patterns between women and men. As a means of estimating the specificity of such findings, we also employed a randomization control procedure whereby parallel analyses were performed on two randomly assigned groups of subjects. Activations were compared between groups on a voxel-wise basis to detect focal differences in activation across the entire brain. Group differences were also investigated by region of interest and by hemisphere. In one region of interest analysis an anatomically based method for determining regions identical to that of Pugh *et al.* (1996) was used, in which the number of active voxels above a statistical threshold in each region of interest provided the dependent measures. In a second region of interest analysis functionally defined regions and average activation magnitude values in each region of interest were used as the dependent measures.

Material and methods
Subjects
Subjects were 100 consecutively encountered, healthy, native English speakers who indicated right-handed preferences

Table 1 *Summary of demographic data (mean ± SD)*

	Age	Education	Handedness
Men	23.78 ± 3.79	15.48 ± 2.81	79.50 ± 19.81
Women	22.32 ± 3.99	14.70 ± 2.40	85.72 ± 15.18

(laterality quotient >50) on the Edinburgh Handedness Inventory (Oldfield, 1971). There were 50 men and 50 women, matched on age, education and handedness scores. Means and standard deviations are given in Table 1. Subjects were recruited from classes at local universities and via advertisements in local newspapers. After full explanation of the risks and purposes of this study, all subjects gave written informed consent according to institutional guidelines and were paid a small hourly stipend.

Imaging methods

Scanning was performed on a 1.5 T General Electric Signa scanner (Milwaukee, Wis., USA) using three-axis local gradient and insertable transmit/receive radio frequency coils designed for whole-brain imaging. A gradient-echo echo-planar sequence was used for fMRI with the following parameters: TE (echo time) = 40 ms, TR (repetition time) = 4 s, FOV (field of view) = 24 cm, matrix = 64 × 64, slice thickness = 7 mm. Seventeen to 19 contiguous sagittal slice locations were imaged covering the entire brain, and 100 time series images were obtained at each slice location during the scan.

Tasks

During semantic monitoring, subjects heard spoken English nouns designating animals (e.g. 'rabbit') and were instructed to respond to animals that are both 'found in the United States' and 'used by humans'. During tone monitoring, a non-linguistic control task, subjects heard sequences of three to seven tones in which each constituent tone was either low (500 Hz) or high (750 Hz) in pitch. Subjects were instructed to respond to sequences containing two 'high' tones. During eight activation cycles, semantic monitoring was performed for 24 s with intervening 24 s intervals of tone monitoring. Stimuli were presented at a rate of one every 3 s, and targets occurred on three out of eight trials in each condition. Responses consisted of a button press with the left hand. These tasks were described previously (Binder *et al.*, 1995, 1996*a*, 1997).

Performance on the semantic monitoring task was calculated by comparing each response made by a given subject with those given by a control group of 50 normal subjects on the same stimuli. Items responded to by controls with a probability $P > 0.75$ were categorized as targets, and items responded to with a probability $P < 0.25$ were categorized as distractors. Performance on the tone

monitoring task was calculated as the percentage of trials on which subjects responded correctly.

Image processing and voxel-wise analyses

An automated alignment program was used to minimize possible artefacts due to head motion (Cox, 1996*a*). Images five through to 100 were registered to image four, and only these final 96 images were used in further analyses. For each subject, differences in the MRI signal between semantic monitoring and intervening tone monitoring epochs were calculated on a voxel-wise basis for each activation cycle, using the last four images of each semantic monitoring and tone monitoring epoch. Difference maps showing the mean absolute difference in signal change between semantic monitoring and tone monitoring, and t-maps showing the significance of these differences as a t statistic, were computed for each subject using the eight difference measurements. Individual t-maps and difference maps were transformed into standard 3D stereotaxic space, resampled to a 1 mm grid, and smoothed slightly with a 4 mm root mean square Gaussian filter using the MCW-AFNI software package (Cox, 1996*b*).

In order to determine the specificity of results from the sex comparisons described below, we performed parallel analyses on two pseudorandomly assigned groups comprising equal numbers of women and men. This procedure provides an additional check on whether observed 'sex' differences could be due to other factors or to chance. This technique also demonstrates empirically the number of significant findings and the amount of variability expected by chance in this sample.

Individual t-maps were averaged across subjects to produce group t-maps for men, women and the two random groups. The average t-maps were thresholded at $P < 0.0001$ for qualitative comparisons of group activation patterns, as described by Binder *et al.* (1997). Between-group, voxel-wise t-tests were also performed on the individual difference maps, contrasting men with women and the two random groups with each other. Between-group t-maps representing t statistics at each voxel for these comparisons were thresholded at a nominal $P < 0.0001$ to eliminate false positive voxels in non-brain regions. A Bonferroni-corrected significance threshold would have been more conservative, but this more lenient threshold was used to increase the likelihood of detecting subtle sex differences. Voxel clusters surviving this threshold that were smaller than 200 μl (approximately two original voxels) were excluded. Effect sizes for the between-groups t-tests were estimated at each voxel using a calculation of the effect size $d = 2t/\mathrm{sqrt(d.f.)}$ where t is the t-test value and d.f. is the degrees of freedom used in the t-test calculation (Cohen, 1988).

Region of interest analyses

Voxel-wise comparisons allow an unbiased assessment of sex effects at each coordinate position in the brain. This

technique may be relatively insensitive to such effects, however, because variance at a voxel level is likely to result mainly from random, local gyral/sulcal variations. Comparisons at a regional level may improve sensitivity to group differences in large-scale neuronal organization by minimizing sensitivity to local anatomical variability.

The first region of interest analysis performed replicated the methods of Pugh *et al.* (1996). Stereotaxic regions of interest were identified in exactly the same locations as reported by Pugh *et al.* (1996) using the coordinate system of Talairach and Tournoux (1988). Rectangular regions of interest were created in the axial plane for the lateral orbital gyrus (volume coordinates Ab, Bb, Bc, Cb, Cc at $z = -8$), prefrontal dorsolateral region (Bc, Bd at $z = 8$ and Bc at $z = 20$), inferior frontal gyrus (Cc, Cd, Dc, Dd at $z = 8$, and Cc, Cd, Dc, Dd at $z = 20$), superior temporal gyrus (Dc, Dd at $z = -8$, E3c, E3d, Fc, Fd at $z = 8$ and Fc, Fd, Gc, Gd at $z = 20$), middle temporal gyrus (E1c, E1d, E2c, E2d, E3d, Fd at $z = -8$, Gc, Gd at $z = 8$, and Hc at $z = 20$), lateral extrastriate region (Hc, Ib at $z = -8$ and $z = 8$, Ib at $z = 20$) and medial extrastriate region (Gb, Ha, Hb, Ia at $z = -8$, Ha, Ia at $z = 20$). Each volume was 8 mm thick and parallel to the line connecting anterior and posterior commissures, as in the Pugh *et al.* (1996) study. The individual *t*-maps were thresholded at $P < 0.05$, and the number of voxels surviving this threshold in each region of interest was calculated. Two two-factor multivariate analyses of variance (MANOVAs) (sex by hemisphere and random group by hemisphere) were performed to compare men with women and the two random groups with each other at each region of interest.

An alternative region of interest method using magnitude differences and functionally derived regions was included to explore the effects of using these different methods. Compared with counting the number of activated voxels above a statistical threshold, magnitude difference measures are less sensitive to head motion, brain pulsatility and other noise sources. Functionally derived regions of interest are less likely to include non-brain or non-activated brain regions, or to combine spatially contiguous but functionally distinct areas in the same region of interest. Functionally derived regions of interest were identified in an average activation map created by merging individual t-maps from an original sample of 80 subjects (40 men and 40 women). This average map was thresholded at a nominal $P < 10^{-9}$ to create several non-contiguous voxel clusters representing regions of relatively strong activation. Binary region of interest mask images were created for each of these voxel clusters, including clusters in left prefrontal, left angular, left temporal and left retrosplenial cortex, a left thalamocapsular region and a right cerebellar region. Each of these binary mask images was reflected across the midline to create regions of interest for homologous regions in the opposite hemisphere (Fig. 1). The region of interest mask images were then used to select, in each subject, voxels within the region of interest for subsequent analyses. Because we selected large regions from

a group activation map, the region of interest volumes were sufficiently large to ensure that the majority of activated voxels were included in each subject.

For each subject, the average absolute difference value for each region of interest was calculated from the difference maps by averaging the difference values for all voxels in the region of interest. A MANOVA on the average difference values for each of the dependent region of interest measures included two factors: sex (male versus female) and hemisphere (left versus right). An identical MANOVA was performed on the random groups with the following two factors: random group (group 1 versus group 2) and hemisphere.

Intrahemispheric difference analyses

So that groups could be compared in terms of intrahemispheric activation patterns, relative intrahemispheric differences in activation were calculated for each subject using the functionally derived regions of interest. Differences in activation magnitude were calculated between the prefrontal and temporal regions of interest, the prefrontal and angular regions of interest, and the temporal and angular regions of interest. MANOVAs were performed to compare men with women, and the two random groups with each other, on each of the six intrahemispheric difference values (three difference values for each hemisphere).

Results
Performance

Women and men did not differ significantly in performance on either the semantic monitoring or tone monitoring task, and both groups performed well above chance levels. On the semantic monitoring task, men averaged 90.8% correct (SD = 6.3) and women averaged 90.4% correct (SD = 6.2). On the tone monitoring task, men averaged 97.6% correct (SD = 2.8) and women averaged 97.1% correct (SD = 3.9).

Activation patterns

The averaged *t*-maps for women and men (Fig. 2, bottom 2 rows) were similar to each other and to those for the random groups (Fig. 2, top 2 rows). For all group averages, activations were strongly left-lateralized, with large activations in prefrontal, temporal, angular, retrosplenial and thalamocapsular regions (Binder *et al.*, 1997). Right-sided activations occurred primarily in the cerebellum. Qualitative comparison of the activation maps for the two random groups demonstrated the extent of variation that can occur from chance alone. For example, group one showed small activation foci in the left cerebellum, left thalamocapsular region and right frontal lobe that were not visible in the group two activation map. Small differences between the average maps
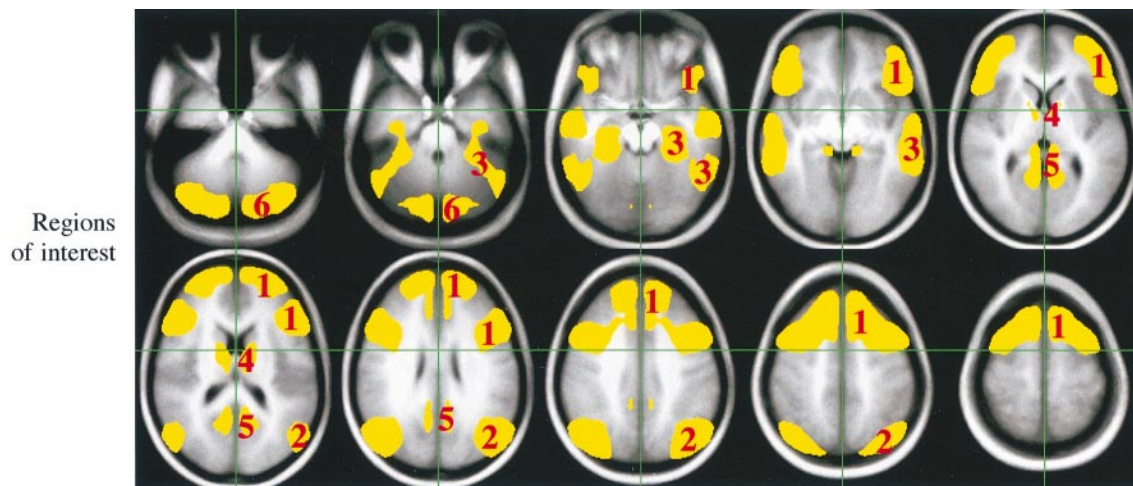
Fig. 1 Regions of interest identified in an average activation map from 80 subjects. The left side of the brain is on the reader's right in this figure and in Fig. 2. Regions are numbered for the left hemisphere (and apply to homologous regions in the right hemisphere) as follows: 1 = prefrontal, 2 = angular gyrus, 3 = temporal, 4 = thalamocapsular, 5 = retrosplenial, 6 = cerebellar. Talairach z coordinates for slices are −36, −26, −16, −6, 4, 14, 24, 34, 44 and 54.
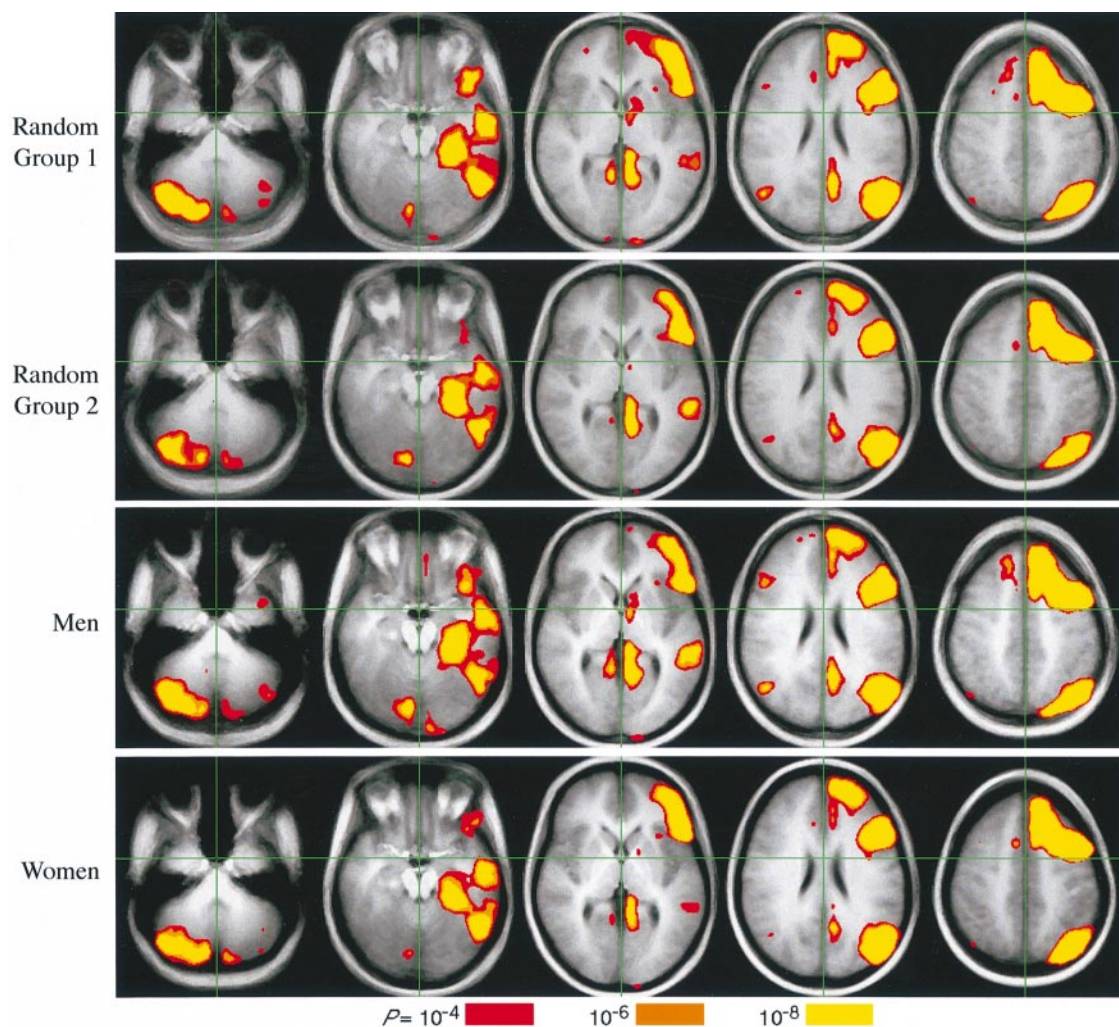
Fig. 2 Top two rows: averaged semantic monitoring–tone monitoring (SM–TM) t-maps for two random groups with equal numbers of women and men. Bottom two rows: semantic monitoring–tone monitoring t-maps for men and women. Between-groups t-tests revealed no significant differences in overall activation patterns. Talairach z coordinates for slices are −36, −16, 4, 24 and 44.

for men and women did not exceed those observed in the random comparison.

Voxel-wise comparisons

In the between-groups *t*-tests, there were no voxel clusters that passed both the significance threshold ($P < 0.0001$) and the size threshold (volume >200 μl). Effect sizes for the between-groups *t*-tests were small (0–0.4) and were the same magnitude for the sex and random contrasts. Less than 1% of voxels had absolute effect sizes greater than 0.2, and this proportion was the same for the sex and random group comparisons. These results thus reveal no significant differences between women and men in a voxel-wise comparison of activation levels.

Anatomical region of interest comparisons

In the analysis replicating Pugh *et al.* (1996), both men and women had significantly more activated voxels in the left hemisphere than the right hemisphere in all regions of interest ($P < 0.00001$ for all regions of interest) except the medial extrastriate, which showed no significant lateralization. Men had more activated voxels than women in the medial extrastriate regions of interest bilaterally [$F(1,196) = 7.443$, $P < 0.005$]. No other region of interest showed this sex difference. No sex by hemisphere interaction effects were significant, indicating that the degree of lateralization of activation was the same for women and men in all regions of interest.

In the parallel random group analysis, no significant group differences were found. For both groups, greater activation was observed in the left hemisphere for all regions of interest except the medial extrastriate ($P < 0.00001$ for all significant effects). No group by hemisphere interactions were significant.

Functional region of interest comparisons

The functionally derived region of interest analysis revealed that both women and men had greater activation in the left than right hemisphere for all regions of interest ($P < 0.00001$ for all comparisons) except the cerebellum, which was more strongly activated on the right side ($P < 0.00001$). Right lateralization of the cerebellar activation probably accounts for the lack of medial extrastriate lateralization in the other region of interest analysis, as inspection of the data showed that this 'medial extrastriate' region of interest included activated voxels located in the right superior cerebellum. Men showed greater activation than women bilaterally in the retrosplenial [$F(1,196) = 12.367$, $P < 0.001$] and thalamocapsular [$F(1,196) = 5.433$, $P < 0.05$] regions of interest. Notably, there was no significant interaction between sex and hemisphere for any region of interest, indicating that the degree of lateralization of activation was the same for men and women in all regions of interest.
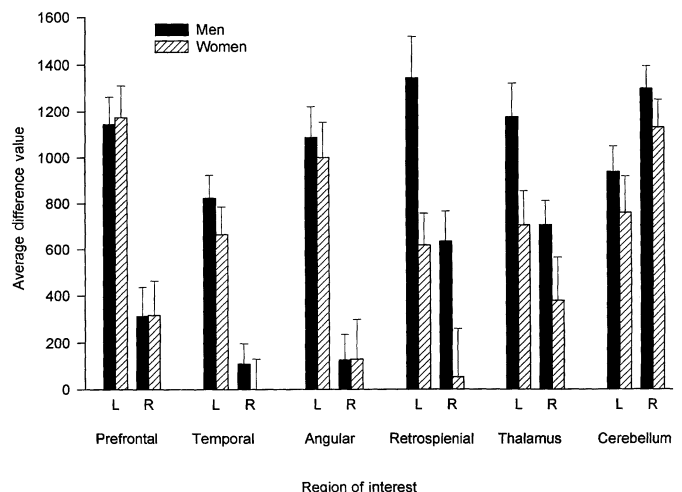
Fig. 3 Average difference values for women and men by hemisphere and region of interest. Error bars indicate standard error of the mean. L = left, R = right.

Figure 3 demonstrates the overall similarity between women and men on average difference values by region of interest, with men showing larger activation values bilaterally in retrosplenial and thalamocapsular regions. No significant main effects for random group assignment were found.

Intrahemispheric comparisons

Analysis of intrahemispheric differences in activation between prefrontal, temporal and angular regions of interest revealed no differences between women and men or between the two random groups.

Discussion

Compared with previous studies, ours had the advantage of a very large sample size, which provided greater statistical power to detect small sex differences. In this study we also employed a validated, reproducible measure of language-related brain activation (Binder *et al.*, 1996a, 1997). Because the language and control tasks employed differ in many ways, including stimulus complexity, phonetic perceptual demands, lexical content and semantic content, the resulting signals are likely to represent the combined activation of several language-related component processors, including speech perceptual, lexical and semantic systems. This activation was strongly lateralized to the left hemisphere in both women and men. No differences were found between the sexes in a voxel-by-voxel analysis, and there were no differences between women and men in lateralization of activity in any region of interest. Men and women also did not differ in terms of intrahemispheric cortical activation patterns. The sexes thus showed very similar, strongly left-lateralized activation patterns, arguing against substantive sex differences in the large-scale neural organization of language functions. While failing to confirm the finding

that women have more bilateral representation of language processing systems than men (Pugh *et al*., 1996), our results are not entirely unexpected. Men and women do not, on average, show significant differences in performance on most language tasks (Hyde and Linn, 1988; Clarke *et al*., 1990), and no sex differences in performance were observed on the activation tasks used in this study. These data thus do not support the hypothesis that women and men carry out identical language processes with the same degree of functional capacity using differently organized neural systems.

In agreement with these results, two other functional imaging groups reported identical activation patterns in men and women during language activation tasks (Buckner *et al*., 1995; Price *et al*., 1996). The tasks used by Price *et al*. (1996) involved phonological and semantic aspects of reading, thus engaging many of the same processing components studied by Shaywitz *et al*. (1995) and Pugh *et al*. (1996). Large, statistically significant effects of task, task order and type of baseline task were found, but sex effects were small and insignificant. Buckner *et al*. (1995) employed word-stem completion and verb generation tasks, which are speech production measures like those on which women and men show significant performance differences in large group studies (Halpern, 1992). Thus, no significant sex differences in large-scale activation patterns were found even on tasks for which there is some evidence of sex-related differences in processing capacity. These results make it even less likely that sex differences in activation are present during language tasks on which men and women perform equivalently.

Despite these findings, many investigators have presented indirect evidence for sex differences in the large-scale neural representation of language functions. In the following section we briefly examine some of this evidence and consider some of the conceptual/methodological issues associated with the various experimental techniques.

Sex differences in macroscopic brain language organization

The corpus callosum and the superior temporal region are both believed to play a role in hemispheric specialization and language function. Many studies of sex differences in the morphology of these structures have been reported, but there are discrepancies between the findings. For example, women were reported to have larger subregions of the corpus callosum in several studies (Witelson, 1989; Steinmetz *et al*., 1992; Clarke and Zaidel, 1994). However, many investigators have not replicated these findings (Kertesz *et al*., 1987; Oppenheim *et al*., 1987; Byne *et al*., 1988; Allen *et al*., 1991; Habib *et al*., 1991; Aboitiz *et al*., 1992). The inconsistency among studies may be due to methodological variations related to how brain regions were measured, and whether these measurements were normalized for brain size (Jancke *et al*., 1997; Leonard, 1997). One report showed that, regardless of sex, larger brains are associated with

smaller corpus callosa (Jancke *et al*., 1997). Thus, sex differences in corpus callosum morphology may be due to sex differences in brain size, as women tend to have smaller brains.

Some recent studies of sex differences in brain morphology have focused on the planum temporale, a region believed to be important for auditory (Binder *et al*., 1996*b*) and associative language processes (for a review, see Kolb and Whishaw, 1990). When measurements were adjusted for total brain size, one study revealed larger plana temporale bilaterally in women (Harasty *et al*., 1997), but others showed no differences between women and men (Aboitiz *et al*., 1992; Witelson and Kigar, 1992). Several investigators reported sex differences in leftward asymmetry of the plana temporale, again with inconsistent results. Men showed greater asymmetry than women in two studies (Witelson and Kigar, 1992; Kulynych *et al*., 1994), but no sex differences in asymmetry were found in another (Aboitiz *et al*., 1992). Future studies may explain the disparate findings reported by different investigators, although at present there is little consistent evidence for sex-related differences in regional brain gross morphology. More importantly, there is currently no evidence directly linking sex-related size differences to differences in language ability.

Deficit-lesion correlation methods have been used extensively to examine sex differences in language organization, but the findings from these studies have also been equivocal. McGlone (1977, 1978) found adverse effects on verbal IQ only after left hemisphere damage in men, but in women verbal IQ was affected after either left or right hemisphere injury, suggesting a more diffuse and bilateral representation of language in women. Kimura (1983), however, reached nearly opposite conclusions, finding evidence that language functions in women are more focally organized in the left frontal lobe than in men. Other investigators have generally not replicated either of these findings (De Renzi *et al*., 1980; Kertesz and Sheppard, 1981; Basso *et al*., 1982; Kertesz, 1982; Warrington *et al*., 1986; Kertesz and Benke, 1989). Kertesz and Sheppard (1981) provided evidence that sex differences reported in some deficit-lesion studies may be due to differences in the location and extent of naturally occurring lesions rather than to differences in underlying cerebral organization. Differences in lesion size and location, which could result from confounding factors such as differences in stroke mechanism or degree of atherosclerotic disease, were not controlled for in several other studies reporting sex differences in aphasia incidence or recovery (McGlone, 1977; Kimura, 1983; Butler, 1984; Pizzamiglio and Mammucari, 1985). Overall, aphasia studies thus provide relatively little evidence for underlying sex differences in the large-scale neural organization of language.

Several behavioural techniques have been used to infer brain language organization in normal subjects. The reports of greater laterality effects in men than women from studies of speech perception using dichotic listening techniques are inconsistent (Lake and Bryden, 1976; McGlone, 1980; Munro

and Govier, 1993). Perceptual advantages in dichotic listening can be biased by attentional factors, however, and may not directly reflect underlying functional asymmetries (Mondor and Bryden, 1991; Mondor, 1994). Sex differences in language lateralization have been studied extensively using divided visual field techniques, although again with very inconsistent results. Fairweather (1982), for example, reviewed 188 such studies, and in 87 of them no evidence of sex effects was found. Of those in which differences were found, some suggested greater lateralization effects in men, while others found stronger lateralization in women (e.g. Healy *et al.*, 1985). The dual-task paradigm, originally described by Kinsbourne and Cook (1971), assumes that simultaneous performance of a language and motor task will lead to a lateralized motor performance decrement if the language task is preferentially performed by one hemisphere. Reports of sex differences on these tasks have also been inconsistent (Simon and Sussman, 1987; Lewis and Christiansen, 1989; Seth-Smith *et al.*, 1989; Ashton and McFarland, 1991). This technique may be more reflective of manual than language dominance, as left-handed subjects consistently showed right hemisphere dominance with this paradigm in one report (Simon and Sussman, 1987). Although some of the inconsistency arising from these techniques may be explained by differences in methodology and task requirements, no such account has yet emerged despite the vast amount of available data. Across the many studies that have been reported, no consistent sex differences have been found using speech production, phonological, lexical or semantic tasks.

Taken as a whole, this literature does not provide strong evidence for sex differences in the large-scale neural organization of language functions. If present, these differences are likely to be small in comparison with the degree of similarity in language system organization between men and women. This conclusion is in accord with the findings of the present study and with the general similarity between women and men on most measures of language processing ability.

Sex differences in activation magnitude

Although the present findings do not support the notion that women and men differ substantively in large-scale language system organization, several subtle sex differences were observed. The stronger activation observed in men in retrosplenial, thalamocapsular and medial extrastriate regions of interest has not been reported previously. These differences occurred bilaterally, and there were no sex differences in the degree of lateralization of activity in these regions of interest. Although these differences were not large, they were not observed in the random comparisons.

Interpretation of these magnitude differences is not straightforward, however, due to the fact that the methods used here measure relative differences in activation rather than absolute activation levels. Thus, the stronger activation

bilaterally in these regions of interest in men could be explained equally well by postulating that men have relatively greater activation during the linguistic task, or that women have greater activation during the non-linguistic control task. Given this uncertainty, it would be premature to infer that these relative differences in activation magnitude have any relationship to differences in neural organization of language functions or differences in verbal ability. Future research could perhaps resolve this ambiguity by examining sex differences in activation produced by the control task in comparison with a more neutral baseline.

Conclusions

The study of sex-related differences in brain language organization has important practical value. Various male-female subject ratios are used in neuroimaging studies, and many studies report data only for male subjects. If men and women differ significantly in brain organization for language processing, these results would not be directly comparable or generalizable across the sexes. In agreement with PET imaging data on verbal fluency and reading tasks, the results from this study indicate that the large-scale organization of language function in the brain is very similar in men and women. Combined with the PET results, these data suggest that it may be appropriate in many circumstances to generalize PET and fMRI language activation results across sex groups and across experiments using different sex ratios.

Although the preponderance of current evidence suggests much greater similarity than difference between the sexes in the large-scale neural organization of language, more functional imaging data are needed to account for the conflicting results. Future studies should involve large subject samples to better detect small effects and to estimate more reliably the size of these effects. For fMRI studies, particular attention will need to be given to measuring and minimizing head motion, an ubiquitous noise source that can mask true activation signals and result in bilateral false positive activations (Hajnal *et al.*, 1994).

Although the same general brain regions appear to subserve language functions in men and women, it is possible that sex-related differences exist at a microscopic level, involving differences in connectivity, neuronal density or synaptic efficiency (Witelson *et al.*, 1995). Such factors could account for ability differences even in the absence of large-scale differences in functional organization, and may not be detectable using macroscopic functional imaging methods. If present, these microscopic differences could be (i) genetically determined, (ii) the result of hormonal or other metabolic factors, (iii) the result of neural plasticity induced by environmental or experiential factors, or (iv) any combination of these.

Acknowledgements

References

Aboitiz F, Scheibel AB, Zaidel E. Morphometry of the Sylvian fissure and the corpus callosum, with emphasis on sex differences. Brain 1992; 115: 1521–41.

Allen LS, Richey MF, Chai YM, Gorski RA. Sex differences in the corpus callosum of the living human being. J Neurosci 1991; 11: 933–42.

Ashton R, McFarland K. A simple dual-task study of laterality, sex-differences and handedness. Cortex 1991; 27: 105–9.

Basso A, Capitani E, Moraschini S. Sex differences in recovery from aphasia. Cortex 1982; 18: 469–75.

Binder JR, Rao SM, Hammeke TA, Frost JA, Bandettini PA, Jesmanowicz A, et al. Lateralized human brain language systems demonstrated by task subtraction functional magnetic resonance imaging. Arch Neurol 1995; 52: 593–601.

Binder JR, Swanson SJ, Hammeke TA, Morris GL, Mueller WM, Fischer M, et al. Determination of language dominance using functional MRI: a comparison with the Wada test. Neurology 1996a; 46: 978–84.

Binder JR, Frost JA, Hammeke TA, Rao SM, Cox RW. Function of the left planum temporale in auditory and linguistic processing. Brain 1996b; 119: 1239–47.

Binder JR, Frost JA, Hammeke TA, Cox RW, Rao SM, Prieto T. Human brain language areas identified by functional magnetic resonance imaging. J Neurosci 1997; 17: 353–62.

Brust JC, Shafer SQ, Richter RW, Bruun B. Aphasia in acute stroke. Stroke 1976; 7: 167–74.

Buckner RL, Raichle ME, Petersen SE. Dissociation of human prefrontal cortical areas across different speech production tasks and gender groups. J Neurophysiol 1995; 74: 2163–73.

Butler S. Sex differences in human cerebral function. Prog Brain Res 1984; 61: 443–55.

Byne W, Bleier R, Houston L. Variations in human corpus callosum do not predict gender: a study using magnetic resonance imaging. Behav Neurosci 1988; 102: 222–7.

Clarke JM, Zaidel E. Anatomical-behavioral relationships: corpus callosum morphometry and hemispheric specialization. Behav Brain Res 1994; 64: 185–202.

Clarke E, Gardner MK, Brown G, Howell RJ. Changes in analogical reasoning in adulthood. Exp Aging Res 1990; 16: 95–9.

Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. Hillsdale (NJ): Lawrence Erlbaum; 1988.

Cox RW. Algorithms for image registration, motion detection, and motion correction. In: Savoy R, editor. FMRI2DAY: a two day workshop on functional MRI. Boston: MGH-NMR Center; 1996a. p. 25–43.

Cox RW. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. Comput Biomed Res 1996b; 29: 162–73.

Damasio H, Tranel D, Spradling J, Alliger R. Aphasia in men and women. In: Galaburda AM, editor. From reading to neurons. Cambridge (MA): MIT Press; 1989. p. 307–25.

De Renzi E, Faglioni P, Ferrari P. The influence of sex and age on the incidence and type of aphasia. Cortex 1980; 16: 627–30.

Fairweather H. Sex differences: little reason for females to play midfield. In: Beaumont JG, editor. Divided visual field studies of cerebral organisation. London: Academic Press; 1982. p. 147–94.

Habib M, Gayraud D, Oliva A, Regis J, Salamon G, Khalil R. Effects of handedness and sex on the morphology of the corpus callosum: a study with brain magnetic resonance imaging. Brain Cogn 1991; 16: 41–61.

Hajnal JV, Myers R, Oatridge A, Schwieso JE, Young IR, Bydder GM. Artifacts due to stimulus correlated motion in functional imaging of the brain. Magn Reson Med 1994; 31: 283–91

Halpern DF. Sex differences in cognitive abilities. 2nd ed. Hillsdale (NJ): Lawrence Erlbaum; 1992.

Harasty J, Double KL, Halliday GM, Kril JJ, McRitchie DA. Language-associated cortical regions are proportionally larger in the female brain. Arch Neurol 1997; 54: 171–6.

Healy JM, Waldstein S, Goodglass H. Sex differences in the lateralization of language discrimination vs language production. Neuropsychologia 1985; 23: 777–89.

Hyde JS, Linn MC. Gender differences in verbal ability: a meta-analysis. Psychol Bull 1988; 104: 53–69.

Jancke L, Staiger JF, Schlaug G, Huang Y, Steinmetz H. The relationship between corpus callosum size and forebrain volume. Cereb Cortex 1997; 7: 48–56.

Kertesz A. Sex distribution in aphasia. Behav Brain Sci 1982; 5: 310.

Kertesz A, Benke T. Sex equality in intrahemispheric language organization. Brain Lang 1989; 37: 401–8.

Kertesz A, Sheppard A. The epidemiology of aphasic and cognitive impairment in stroke: age, sex, aphasia type and laterality differences. Brain 1981; 104: 117–28.

Kertesz A, Polk M, Howell J. Black SE. Cerebral dominance, sex, and callosal size in MRI. Neurology 1987; 37: 1385–88.

Kimura D. Sex differences in cerebral organization for speech and praxic functions. Can J Psychol 1983; 37: 19–35.

Kinsbourne M, Cook J. Generalized and lateralized effects of concurrent verbalization on a unimanual skill. Quart J Exp Psychol 1971; 23: 341–5.

Kolb B, Whishaw IQ. Fundamentals of human neuropsychology. 3rd ed. New York: W.H. Freeman; 1990.

Kulynych JJ, Vladar K, Jones DW, Weinberger DR. Gender differences in the normal lateralization of the supratemporal cortex:

MRI surface-rendering morphometry of HeschL's gyrus and the planum temporale. Cereb Cortex 1994; 4: 107–8.

Lake DA, Bryden MP. Handedness and sex differences in hemispheric asymmetry. Brain Lang 1976; 3: 266–82.

Lansdell H. A sex difference in effect of temporal-lobe neurosurgery on design preference. Nature 1962; 194: 852–4.

Lansdell H, Urbach N. Sex differences in personality measures related to size and side of temporal lobe ablations. Proc Amer Psychol Assoc 1965; 73: 113–4.

Leonard CM. Corpus callosum: sex or size? Cereb Cortex 1997; 7:2.

Lewis RS, Christiansen L. Intrahemispheric sex differences in the functional representation of language and praxic functions in normal individuals. Brain Cogn 1989; 9: 238–43.

Lukatela G, Carello C, Savic M, Turvey MT. Hemispheric asymmetries in phonological processing. Neuropsychologia 1986; 24: 341–50.

Maccoby EE, Jacklin CN. The psychology of sex differences. Stanford (CA): Stanford University Press; 1974.

McGlone J. Sex differences in the cerebral organization of verbal functions in patients with unilateral brain lesions. Brain 1977; 100: 775–93.

McGlone J. Sex differences in functional brain asymmetry. Cortex 1978; 14: 122–8.

McGlone, J. Sex differences in human brain asymmetry: a critical survey. Behav Brain Sci 1980; 3: 215–27.

Mondor TA. Interaction between handedness and the attentional bias during tests of dichotic listening performance. J Clin Exp Neuropsychol 1994; 16: 377–85.

Mondor TA, Bryden MP. The influence of attention on the dichotic REA. Neuropsychologia 1991; 29: 1179–90.

Munro P, Govier E. Dynamic gender-related differences in dichotic listening performance. Neuropsychologia 1993; 31: 347–53.

Oldfield RC. The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia 1971; 9: 97–113.

Oppenheim JS, Lee BC, Nasa R, Gazzaniga MS. No sex-related differences in human corpus callosum based on magnetic resonance imagery. Ann Neurol 1987; 21: 604–6.

Pizzamiglio L, Mammucari A, Razzano C. Evidence for sex differences in brain organization in recovery in aphasia. Brain Lang 1985; 25: 213–23.

Price CJ, Moore CJ, Friston KJ. Getting sex into perspective. Neuroimage 1996; 3 (3 Pt 2): S586.

Pugh KR, Shaywitz BA, Shaywitz SE, Constable RT, Skudlarski P, Fulbright RK, et al. Cerebral organization of component processes in reading. Brain 1996; 119: 1221–38.

Seth-Smith M, Ashton R, McFarland K. A dual-task study of sex differences in language reception and production. Cortex 1989; 25: 425–31.

Shaywitz BA, Shaywitz SE, Pugh KR, Constable RT, Skudlarski P, Fulbright RK, et al. Sex differences in the functional organization of the brain for language [see comments]. Nature 1995; 373: 607–9. Comment in Nature 1995; 373: S61–2.

Simon TJ, Sussman HM. The dual task paradigm: speech dominance or manual dominance? Neuropsychologia 1987; 25: 559–69.

Steinmetz H, Jancke L, Kleinschmidt A, Schlaug G, Volkmann J, Huang Y. Sex but no hand difference in the isthmus of the corpus callosum [see comments]. Neurology 1992; 42:749–52. Comment in: Neurology 1992; 42: 709–13.

Talairach J, Tournoux P. Co-planar stereotaxic atlas of the human brain: 3-dimensional proportional system: an approach to cerebral imaging. Stuttgart: Thieme; 1988.

Warrington EK, James M, Maciejewski C. The WAIS as a lateralizing and localizing diagnostic instrument: a study of 656 patients with unilateral cerebral lesions. Neuropsychologia 1986; 24: 223–39.

Witelson SF. Hand and sex differences in the isthmus and genu of the human corpus callosum. Brain 1989; 112: 799–835.

Witelson SF, Kigar DL. Sylvian fissure morphology and asymmetry in men and women: bilateral differences in relation to handedness in men. J Comp Neurol 1992; 323: 326–40.

Witelson SF, Glezer II, Kigar DL. Women have greater density of neurons in posterior temporal cortex. J Neurosci 1995; 15: 3418–28.