Available online at www.sciencedirect.com

SciVerse ScienceDirect

Journal homepage: www.elsevier.com/locate/cortex

Research report

Structural plasticity in the language system related to increased second language proficiency

Maria Stein [a,1], *Andrea Federspiel* [a,1,*], *Thomas Koenig* [a], *Miranka Wirth* [a], *Werner Strik* [a], *Roland Wiest* [b], *Daniel Brandeis* [c,d] and *Thomas Dierks* [a]

[a] *University Hospital of Psychiatry, Department of Psychiatric Neurophysiology, Bern, Switzerland*
[b] *Institute of Diagnostic and Interventional Neuroradiology, Inselspital, University of Bern, Switzerland*
[c] *Department of Child and Adolescent Psychiatry, University of Zurich, Switzerland*
[d] *Department of Child and Adolescent Psychiatry and Psychotherapy, Central Institute of Mental Health, Mannheim, Germany*

ARTICLE INFO

Article history:
Received 15 April 2010
Revised 18 June 2010
Accepted 20 October 2010
Action editor Stefano Cappa
Published online 28 October 2010

Keywords:
Bilingualism
Learning
Language
Structural plasticity
VBM

ABSTRACT

While functional changes linked to second language learning have been subject to extensive investigation, the issue of learning-dependent structural plasticity in the fields of bilingualism and language comprehension has so far received less notice. In the present study we used voxel-based morphometry to monitor structural changes occurring within five months of second language learning. Native English-speaking exchange students learning German in Switzerland were examined once at the beginning of their stay and once about five months later, when their German language skills had significantly increased. We show that structural changes in the left inferior frontal gyrus are correlated with the increase in second language proficiency as measured by a paper-and-pencil language test. Contrary to the increase in proficiency and grey matter, the absolute values of grey matter density and second language proficiency did not correlate (neither on first nor on second measurement). This indicates that the individual amount of learning is reflected in brain structure changes, regardless of absolute proficiency.

1. Introduction

A substantial amount of research on second language comprehension has demonstrated that increasing second language proficiency is linked to functional changes in language-related brain regions (e.g., Yetkin et al., 1996; Chee et al., 2000, 2001; Sakai et al., 2004; Perani and Abutalebi, 2005; Tatsuno and Sakai, 2005; Stein et al., 2006). Concerning semantic and syntactic processing, literature shows that low proficient bilinguals activate additional, mainly prefrontal, brain regions for the processing of the second language (e.g., Chee et al., 2001; Tatsuno and Sakai, 2005; Stein et al., 2006) while they show less left temporal activity (e.g., Perani et al., 1998). These differences disappear, or are at least diminished, with increasing second language proficiency (e.g., Perani et al., 1998; Tatsuno and Sakai, 2005; Stein et al., 2006, 2009).

* *Corresponding author*. University Hospital of Psychiatry, Department of Psychiatric Neurophysiology, Bolligenstrasse 111, CH-3000 Bern 60, Switzerland.
E-mail address: Federspiel@puk.unibe.ch (A. Federspiel).

[1] Equally contributed.

0010-9452/$ – see front matter © 2010 Elsevier Srl. All rights reserved.
doi:10.1016/j.cortex.2010.10.007

Whether and how changes in brain structure accompany an increase in second language proficiency has until now not been sufficiently investigated.

The work of Golestani et al. investigated the ability to learn the differentiation of foreign speech sound and reported increased white matter density anterior of the parieto-occipital sulcus (Golestani et al., 2002, 2007) and in left Heschels gyrus (Golestani et al., 2007) of fast learners when compared to slow learners. Their findings thus demonstrate a connection between brain anatomy and language abilities in the phonological domain. Mechelli et al. (2004) used a global measure of overall second language proficiency and showed that grey matter density in the left gyrus supramarginalis correlated with achieved second language proficiency. In the same region, grey matter density was found to correlate with vocabulary knowledge in monolingual adolescents but not adults (Richardson et al., 2010; Lee et al., 2007). Cross-sectional studies however did not allow to estimate the timescale on which the observed language-related grey matter alterations occurred and to clearly separate effects of genetic predisposition or increased density before learning from experience-induced structural plasticity. It was therefore emphasized that – similar to research in other domains (e.g., Draganski et al., 2004; Boyke et al., 2008) – longitudinal studies are necessary to further clarify the role of experience-dependent structural plasticity in language acquisition, including second language learning (Sakai, 2005; Richardson and Price, 2009).

The present study now uses a longitudinal approach to examine learning-related structural changes induced by five months of second language learning: Ten native English-speaking exchange students learning German in Switzerland were examined once at the beginning of their stay (day 1) and a second time about five months later, when their second language proficiency had significantly increased (day 2). On both days, the subjects' brain structure was measured with magnetic resonance imaging (MRI) on a Siemens 1.5 Tesla Scanner and their German language proficiency was assessed using two written language test (thus omitting). The relation between changes in language proficiency (as assessed by the language tests) and changes in brain structure was then examined using voxel-based morphometry (VBM). VBM yields voxelwise indices of cortical grey matter density, thus allowing objective investigation of subtle changes in grey matter (Ashburner and Friston, 2000; Good et al., 2001); It has been successfully applied before to investigate experience-dependent changes in brain structure (e.g., Draganski et al., 2004). The swiss cantonal ethics committee approved the study and written informed consent was given by each participant.

As mentioned above, functional studies have mainly indicated prefrontal and temporal cortices as regions where brain activation changes occur when second language proficiency in the semantic and syntactic domain increases (e.g., Perani et al., 1998; Tatsuno and Sakai, 2005; Stein et al., 2009). Following the same line of investigation, with the exception that we concentrate on structural changes, we therefore expect to find structural changes in prefrontal and temporal regions as well.

2. Methods

2.1. *Subjects & timing of first and second measurements*

Ten native English-speaking exchange students learning German in Switzerland participated in the study (3 male, 7 female; mean age 17.5 years, range: 16–18.5 years, all right-handed, countries of origin: Australia, Canada, USA). They were recruited via the exchange organization (rotary youth exchange) at the very beginning of their stay in Switzerland. Before coming to Switzerland, 6 of our subjects had no exposure to the German language, 3 subjects had started learning German by themselves (using books) between two and five months before their arrival in Switzerland, a single subject had attended German classes in high school for the last 4 years. After arrival in Switzerland, all of them attended during 3 weeks an intense German language course organized by the exchange organization. In all cases, first measurement (day 1) was after this initial language course. As the subjects with prior exposure to the German language (especially the subject with the history of German classes in high school) were by no means outliers concerning their language skills or their grey matter density on day 1, all subjects were included in the analyses.

Subjects were invited back for second measurement (day 2) about 5 months after first measurement (days between first and second measurement: mean: 166 days, range: 133–224 days).

2.2. *Assessment of German language proficiency*

As described earlier (Stein et al., 2006), German language proficiency was measured using two tests: The first measure (test 1) was a multiple choice test developed by "inlingua" language schools (www.inlingua.com). It consisted of 100 cloze sentences where the subjects had to fill in blanks by picking the correct answer from four possible options. To answer correctly, syntactic as well as semantic knowledge was needed. This test was applied on day 1 and day 2. The second measure (test 2) was a vocabulary test, where subjects were asked to write down the English equivalents of 40 German nouns. This test served as a measure of semantic single word knowledge in the second language. Two versions of this vocabulary test were created. Half of the subjects had version 1 on day 1 and version 2 on day 2, the other half of the subjects had the order reversed.

As both measures were correlated (see also Section 3.1), the percentage of correct answers from both tests was averaged for each subject, yielding one single (more robust) language proficiency score per day. To assess changes in language proficiency, language proficiency scores on day 1 were subtracted from language proficiency scores on day 2 [Δ(*Lang.Test*)].

2.3. *Magnetic resonance (MR) data*

2.3.1. *MR acquisition*

All images were acquired using a 1.5 Tesla whole-body MRI system (Siemens Vision, Erlangen, Germany) equipped with a standard radio-frequency head coil. Anatomical images on day 1 and day 2 were acquired using a 3D T1-weighted

(Magnetization Prepared Rapid Acquisition Gradient Echo - MP-RAGE) sequence for each subject, providing 192 sagittal slices of 1.0 mm thickness, 256 mm × 256 mm field of view (FOV), and a matrix size of 256 × 256. Further scan parameters were 2000 msec repetition time (TR), 4.4 msec echo time (TE), and a flip angle of 15° (FA). Throughout the duration of this study, the MR acquisition protocol was not subject of any changes and neither the magnet was upgraded nor was the gradient system changed.

2.3.2. *VBM*

The structural high-resolution images were used for VBM analysis (Ashburner and Friston, 2000). We aimed to test for specific regional differences in grey matter density between first and second measurement. Therefore each subject's structural images were preprocessed according to the "optimized VBM protocol" (Good et al., 2001) using SPM5 (Wellcome Department of Imaging Neuroscience, London, England; www.fil.ion.ucl.ac.uk).

All structural images were initially checked for artifacts and the center point was placed on the anterior commissure. Each structural image was normalized to the T1 Template of SMP5. The resulting images were then segmented into grey matter, white matter and cerebrospinal fluid (CSF) and finally smoothed with an isotopic Gaussian kernel of 10 mm Full Width at Half Maximum (FWHM) using SPM5 (Wellcome Department of Imaging Neuroscience, London, England; www.fil.ion.ucl.ac.uk). Then, VBM was applied to the data using the "optimized" protocol as described by Good et al. (Good et al., 2001). We then calculated the total volume of grey matter (TGM), white matter (TWM), and cerebrospinal fluid TCSF, as well as the total intracranial volume (TIV). For the VBM analysis of the grey matter density, we used the spatially normalized, smoothed and unmodulated images of the each subject's grey matter and smoothed with 10 mm FWHM kernel (Ashburner and Friston, 2001). In each subject and each voxel, the grey matter density value for day 1 was then subtracted from the grey matter density value for day 2, yielding a grey matter difference index for each subject and each voxel [Δ(*GM*)].

2.4. Statistics

2.4.1. *Voxelwise statistics*

2.4.1.1. Main linear regression analysis (REG). Again, we used statistical parametric mapping (SPM) to analyze the grey matter difference indexes
[Δ(*GM*)]. First, a linear regression in the framework of general linear model was computed using the increase in second language proficiency [Δ(*Lang.Test*)] as a predictor for the voxelwise change in grey matter density [Δ(*GM*)]. This regression will be abbreviated as "*main* REG" from now on and can be expressed by the following formula: $\Delta(GM) = \beta_0 + \beta_1 \times \Delta(Lang.Test) + \varepsilon$.

2.4.1.2. Correction for multiple comparisons. The correction for multiple comparisons (in order to control for Type I errors) was then performed using the method of cluster-size thresholding (Forman et al., 1995), which was computed by a self-written Matlab program. We started from the uncorrected voxel-level probability threshold of $p_{uncorrected} < .001$, and we performed 1000 Monte Carlo simulations to achieve a corrected probability threshold of $p_{corrected} < .05$.

2.4.2. *Clusterwise statistics*

For each subject, the voxelwise grey matter difference indices [Δ(*GM*)] were extracted from the significant clusters and averaged for each cluster, yielding for each cluster a clusterwise grey matter difference index $[\overline{\Delta}(GM_{cluster})]$. This grey matter difference index was then used for further analyses: to explore the correlations between this index $[\overline{\Delta}(GM_{cluster})]$ and the increase in second language proficiency [Δ(*Lang.Test*)], we performed a correlation analysis using the program "regstats" of MATLAB (MATLAB version 7, release 14; The MathWorks Inc., Natick, USA).

2.4.3. *Check for additional potential predictors*

In a next step, we used nonparametric Spearman rank correlation coefficient (r_s) in order to check for dependencies between the increase in second language proficiency [Δ(*Lang.Test*)] and (a) knowledge of the second language on day 1 [language test score on day 1 (L.T.day1)], (b) days between first and second measurement (Days), (c) gender (gender), (d) age of the subject on day 1 (*age*) (e) general maturation effects [assessed through the overall TGM-difference (maturation)]. Those variables showing a correlation with our main predictor [the increase in the second language test (Δ(*Lang.Test*))] were then included in an additional multivariate linear regression (*extended* REG) to test if the observed grey matter changes could not only be described by an increase in second language proficiency [Δ(*Lang.Test*)] alone, but also by possible confounding predictors.

Because of the limited sample size, we preferred the above described analysis procedure over a within-subjects regression including multiple covariates (i.e., ANCOVA in SPM).

3. Results

3.1. Behavioral data

Second language proficiency, as indicated in a combined measure of two tests described above, increased in every single subject (see Table 1). In the mean, this increase was highly significant: The mean percentage of correct answers in the language tests increased from 27% (±5%) on day 1 to 47% (±7%) on day 2 [$t_{(df=9)} = 10.6$; $p < .001$].

Even though the vocabulary test mainly tested semantic knowledge while a combination of semantic and syntactic knowledge was needed to solve the cloze-sentence-test, both measures were highly correlated (day 1: $r = .8$, $p = .002$; day 2: $r = .7$, $p = .008$).

3.2. VBM data

Overall indices like TGM, TWM, TCSF and TIV were extracted. Comparison of the overall indices showed that TIV increased significantly from day 1 to day 2 [$t_{(df=9)} = 2.3$; $p = .0046$]. This increase was mainly due to a significant increase in TGM [$t_{(df=9)} = 4.6$; $p = .0013$], while TWM [$t_{(df=9)} = 1.9$; $p = .09$] and TCSF [$t_{(df=9)} = .47$; $p = .65$] remained rather stable.

Table 1 – Single subject data.

Subject	Language test score day 1 [% cor. ans.]	Language test score day 2 [% cor. ans.]	Language test score difference (day 2 – day 1) [% cor. ans.]	Total grey matter volume difference (day 2 – day 1) [µl]	Time between day 1- and day 2-measurement [days]	Age [years]	Gender
S01	22.63	37.13	14.5	123.07615	224	16.6	m
S02	38	57.88	19.88	53.15978	161	16.8	m
S03	28.5	37	8.5	113.59018	133	18.3	f
S04	22.5	47.63	25.13	92.90347	210	16.2	m
S05	21.63	49.25	27.63	81.0527	175	18.2	f
S06	29.63	46.25	16.63	85.00624	147	18.6	f
S07	31.13	52.75	21.63	101.52301	133	18.1	f
S08	22.38	38.38	16	124.99286	147	18.4	f
S09	25.38	47.5	22.13	-28.5802	133	18.3	f
S10	29.38	54.88	25.5	1.4478	168	16	f

For every subject, the following information is given: language test scores (i.e., combined scores from both language tests), total grey matter volume (whole brain), time passing between first and second measurement, age, gender. Abbreviations: cor.ans. = correct answers; m = male; f = female.

3.3. *Relationship between second language proficiency and VBM data*

3.3.1. *Voxelwise analysis*

The relationship between the increase in second language proficiency [language test score on day 2 minus language test score on day 1; Δ(*Lang.Test*)] and the grey matter density change [measured in each voxel as: grey matter density on day 2 minus grey matter density on day 1; Δ(*GM*))] was evaluated: A linear regression analysis within the framework of the general linear model was computed using the increase in second language proficiency as a predictor for the voxelwise increase in grey matter density (*main* REG).

The correction for multiple comparisons (in order to control for Type I errors) yielded an extent threshold of 1242 voxels (i.e., a volume of 1242 mm^3) to protect against false positives at a probability level of $p_{corrected} < .05$. At a corrected voxel-level probability threshold of $p_{corrected} < .05$, the *main* REG linear regression analysis revealed two big clusters with a significant relationship between the increase in second language proficiency and the grey matter density change (see Fig. 1): These clusters were located in the left inferior frontal gyrus (lIFG; volume of 2249 mm^3) and in the left anterior temporal lobe (lATL; volume of 2891 mm^3).

Fig. 1 shows the two clusters with significant correlations, indicating that subjects with a greater increase in second language proficiency also showed a greater increase in grey matter density in those brain regions: One cluster (2249 mm^3) was located in the lIFG [mostly located in Brodmann area (BA) 46, but spanning also parts of BA 45, 47, 13 and 10]. The second cluster (2891 mm^3) was situated in the lATL (mostly located in BA 28, but spanning also parts of BA 20, 36, and 21, thus including medial as well as lateral parts of the anterior temporal cortex).

3.3.2. *Clusterwise analysis*

The mean grey matter difference indices for the clusters lIFG and 1ATL were computed as described above (see Section 2.4.2), yielding for each cluster a mean grey matter difference index for each subject [$\overline{\Delta}(GM_{lIFG})$ and $\overline{\Delta}(GM_{lATL})$]. Subsequent correlation analyses describing the relationship between the increase in second language proficiency [Δ(*Lang.Test*)] and the grey matter density change in the two significant clusters revealed for the left IFG [$\overline{\Delta}(GM_{lIFG})$] a significant correlation with an explained variance of 63% [$t_{(df=8)} = 4.01$; $p = .0039$; $r^2 = .63$, see Fig. 2] and for the left ATL [$\overline{\Delta}(GM_{lATL})$] a significant correlation with an explained variance of 50% [$t_{(df=8)} = 2.84$; $p = .021$; $r^2 = .503$, see Fig. 3].

3.4. *Check for possible alternative predictors*

As described in Section 2.4.3, we used a nonparametric correlation (Spearman rank, r_s) to check for dependencies between our first predictor [the increase in second language proficiency, Δ(*Lang.Test*)] and possible alternative predictors. Of these possible alternative predictors the following did not show a statistical relationship to increase in second language proficiency: (a) knowledge of the second language on day 1 [language test score on day 1 (L.T.day1); $r_s = -.18$; $p = .63$], (b) days between first and second measurement (*Days*; $r_s = .27$; $p = .45$), (c) gender ($r_s = -.11$; $p = .75$) and (d) age (*age*; $r_s = -.38$;

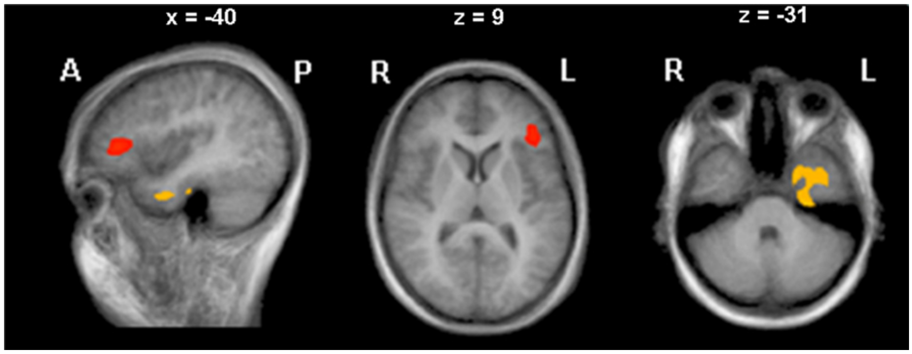

Fig. 1 – Clusters with significant correlation ($p_{corrected}$ < .05) between increase in second language proficiency and cortical grey matter density change plotted on the averaged brain of the 10 subjects; red: cluster in the lIFG (2249 mm^3); yellow: cluster in the lATL (2891 mm^3). A: anterior; P: posterior, L: left; R: right.

$p = .28$). Only the overall TGM difference, which we used as an indicator of general maturation effects showed a significant correlation with the increase in second language proficiency was observed (*maturation*; $r_s = -.68$; $p = .035$).

Therefore, we performed an additional multivariate linear regression (*extended* REG, as described in Section 2.4.3). With this *extended* REG, we tested if the grey matter changes in our clusters [$\overline{\Delta}(GM_{lIFG})$ and $\overline{\Delta}(GM_{lATL})$] could not only be described by an increase in second language proficiency [$\Delta(Lang.Test)$] alone, but also by a general maturation effects [assessed through the overall TGM-difference (maturation)]. This *extended* REG can be expressed through the following formula: $\overline{\Delta}(GM) = \beta_0 + \beta_1 \times \Delta(Lang.Test) + \beta_2 \times (maturation) + \varepsilon$

In the left IFG, this *extended* REG revealed: (i) predictor increase in second language proficiency [$\Delta(Lang.Test)$]: $t_{(df=7)} = 4.62$; $p = .024$; and (ii) predictor general maturation effects (maturation): $t_{(df=7)} = 1.63$; p = n.s.. The explained variance by this *extended* REG was found to be $r^2 = .69$ (obtained by the adjusted r^2), thus exceeding the explained variance obtained with the *main* REG ($r^2 = .63$, see above) by only 6%.

In the left ATL, the *extended* REG revealed: (i) predictor increase in second language proficiency [$\Delta(Lang.Test)$]: $t_{(df=7)} = 2.79$; $p = .027$; and (ii) predictor general maturation effects (maturation): $t_{(df=7)} = .78$; p = n.s. The explained variance by this *extended* REG was found to be $r^2 = .41$ (obtained by the adjusted r^2). Here, the *extended* REG explained even a lower percentage of the variance than the *main* REG ($r^2 = .50$, see above). In both clusters, the additional predictor general maturation effects did not contribute significantly to the explained variance. It is thus unlikely that the observed effects are better explained by general maturation than by increasing second language proficiency.

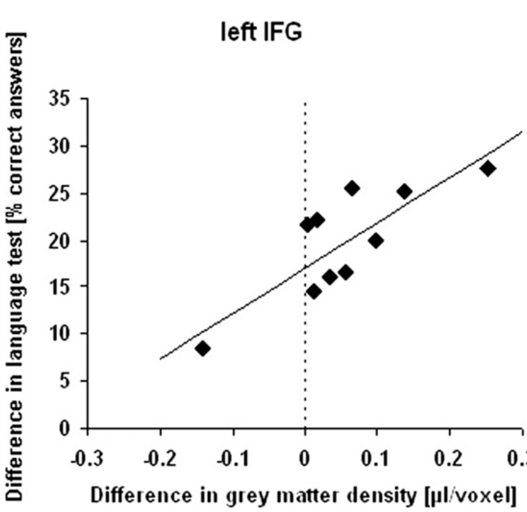

Fig. 2 – Grey matter difference in the left IFG cluster (mean of grey matter density over voxels belonging to the cluster on day 2 minus mean over same voxels on day 1) plotted against the difference in the language test (test core on day 2 minus test score on day 1). The linear regression line is overlaid ($r = .79$, $r^2 = .63$, $p = .0039$).

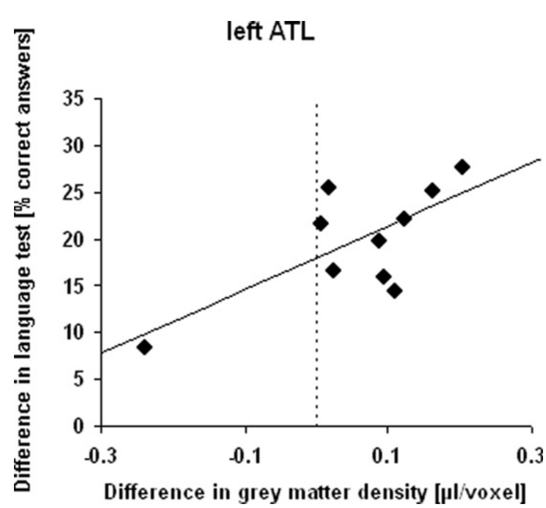

Fig. 3 – Grey matter difference in the left ATL cluster (mean of grey matter density over voxels belonging to the cluster on day 2 minus mean over same voxels on day 1) plotted against the difference in the language test (test core on day 2 minus test score on day 1). The linear regression line is overlaid ($r = .71$, $r^2 = .5$, $p = .02$).

4. Discussion

In the present longitudinal study examining exchange students before and after 5 months of second language learning, we observed a correlation between increase in second language proficiency and increase in grey matter density in the left IFG and in the lATL. Notably, the two clusters are thus located in brain regions that have repeatedly and robustly been found to be involved in semantic as well as syntactic language processing (see Vigneau et al., 2006 for a meta-analysis of 129 scientific reports). Furthermore, research on bilingual subjects has shown that neuronal activity is modulated through second language proficiency in left inferior frontal (Perani et al., 1998; Chee et al., 2001; Perani and Abutalebi, 2005; Tatsuno and Sakai, 2005; Stein et al., 2006, 2009) as well as in left anterior temporal brain regions (Perani et al., 1998; Perani and Abutalebi, 2005).

More specifically, it has been proposed that BA 45 and BA 47 in the left IFG are responsible for selection and integration of semantic information as needed for example in sentence comprehension (Sakai, 2005) and that BA 47 is involved in effortful semantic retrieval (Fiez, 1997). Furthermore, imaging studies have shown that activation in the left IFG increases with semantic task difficulty, but not with task difficulty in general (Raichle et al., 1994; Demb et al., 1995). Concerning second language acquisition, activation changes in the left IFG were reported with increasing second language proficiency in semantic (Perani et al., 1998; Chee et al., 2001; Stein et al., 2006) as well as syntactic tasks (Sakai et al., 2004). It is thus very reasonable that structural changes occur in this region during five months of intense language training.

For the left ATL, it has been shown that high proficient bilinguals show higher activation during language comprehension tasks when compared to low proficient bilinguals (e.g., Perani et al., 1998). According to a recent meta-analysis of language studies dealing with the mother tongue, lATL is involved in semantic and syntactic tasks and is part of a semantic network that constructs overall meaning by integrating knowledge from external (vision, audition) and internal (long-term memory, emotion) messages (Vigneau et al., 2006). While this functional description holds for the more lateral parts of our left ATL cluster, the medial parts are probably even closer linked to the memory system: the medial temporal lobe system holds newly acquired information in memory before a consolidation process including changes in the neocortex begins (Miyashita, 2004).

Two alternative explanations for the observed structural changes may be considered.

First, one could argue that the observed structural changes may be due to an effect of enriched environment. Living for 5 months in a foreign country is a challenging situation with high attentional and computational demands, a situation possibly comparable to an "enriched environment" that has been shown to lead to structural changes in animal studies (van Praag et al., 2000). This new environment (including exposure to the new language) might have been a modulator variable of grey matter density as well as for second language proficiency, thus leading to the observed correlation. A second argument might be that maturational effects are responsible for the observed structural changes as well as for the increase in second language proficiency. In both cases, one would expect that the days between first and second measurement (thus the time period of possible maturation and enriched environment exposure) would explain the grey matter changes. This alternative explanation was tested by computing an additional linear regression analysis within the framework of the generalized linear model (GLM), this time using the days between the two measurements as a predictor for the voxelwise increase in grey matter density. Again, a corrected probability level of $p_{corrected} < .05$ was used. This yielded a cluster in the left ATL overlapping with the left ATL cluster found with the *main* REG using increase in second language proficiency as predictor. It must thus be concluded that the results observed in the left ATL can at least in part be explained by those additional effects. The left IFG-results however can – according to this result – not be explained through effects of enriched environment or general maturation.

In the case of maturational effects, one would also expect the overall grey matter differences to be a significant predictor of the observed correlation. As the *extended* REG (see Results section) shows, this was however not the case. For the frontal lobes, literature suggests furthermore that in normal maturation grey matter volume in the frontal lobes peaks between the age of 11 and 12 and then constantly declines (Giedd et al., 1999).

We thus conclude, that enriched environment exposure might explain at least in part the grey matter changes observed in the left ATL. This conclusion is in accordance with the above mentioned view that the left ATL is part of a network constructing overall meaning from the integration of external messages with stored emotional and long-term memory associations (Vigneau et al., 2006). It is very well conceivable that such a network should be influenced by enriched environment exposure at least as much as by increasing language proficiency.

Contrary to that, neither mere exposure to a new environment nor pure maturational effects can explain the observed correlation between second language improvement and grey matter changes in the left IFG. The grey matter changes in the left IFG thus seem to reflect the individual amount of second language learning. However, more fine-grained methods will have to clarify the underlying cellular changes. The discussion about cellular events responsible for MRI signal changes so far includes changes in neuronal or glial cell size, changes in number or strength of dendritic spines, neuropil or glial changes (Trachtenberg et al., 2002; Chklovskii et al., 2004; Volterra and Meldolesi, 2005; Draganski and May, 2008).

The finding of structural changes in the parietal lobes reported by Golestani et al. (2002, 2007) in relation to speech sound learning or by Mechelli et al. (2004) in relation to global second language proficiency was not replicated. The discrepancy between the effects reported here and those reported by Golestani et al. is probably due to the fact that this study did not encompass oral speech comprehension or production, but focused on structural changes that are related to an increase in second language proficiency as assessed with a written language test. The fact, that the present study did not replicate the effects reported by Golestani et al. is thus not challenging the established link between the learning of non-native speech sounds and white matter volume.

The difference in comparison to Mechelli et al. (2004) might either be due to differences in the behavior measure used (Mechelli et al. used a language test battery that covered a wider range of linguistic functions). Or it might be due to a difference in the analysis approach: While Mechelli et al. defined their regions through inter-individual differences (indicating: the higher the absolute grey matter density, the higher the absolute language proficiency) the regions reported here where defined through a correlation of intra-individual differences (indicating the higher the intra-individual increase in grey matter density, the higher the intra-individual increase in the language test). A third possible explanation, concerns the timescale of the observed changes: While Mechelli's subjects looked back on many years of second language learning, the subjects in the present study were mostly within their first year of second language learning. The discrepancy between the two studies might thus indicate that structural changes in the parietal lobes occur later during second language learning.

One limitation of the present study certainly is its rather small sample size. It is nowadays accepted that between-group comparisons using VBM are optimally performed with large samples. However, there are also highly valuable studies performing between-group comparisons with smaller sample sizes (i.e., Draganski et al., 2004, 12 subjects per group). Furthermore, the present study is based on a within-group comparison, where intra-individual variability in the brain structure provides a more sensitive measurement of individual growth patterns (Hua et al., 2009). In this longitudinal approach each subject serves as his or her own control and any putative change is established directly over repeated measures. We thus feel confident that our results have an implication despite the smaller sample size.

One might argue that the present study only analyzed linear relationships and left non-linear relationships unnoticed. However, the small sample size limits the exploration of additional different relationships, which should be investigated in future studies with larger samples.

Considering the single subject data, one subject (S03) seems to depart from the average. Even if this subject is not a statistical outlier (neither with Grubb's test, nor with Dixon's Q), one might argue that it has an impact on the correlation. We therefore recalculated the whole brain analysis without this subject applying the same statistical criteria as described in Section 2.4.1. This recalculation yielded only one significant cluster close to the ATL cluster reported above, but slightly superior (cluster center at talairach coordinates −45/−16/−17, cluster size 197 mm^3, mostly located in BA 20, see Supplementary Fig. 4 in the online material). In no other brain region, a significant correlation was observed. One cluster overlapping with the above reported IFG cluster was among the clusters showing a correlation (at an uncorrected threshold); however it did not survive the cluster size correction. This recalculation with slightly different results supports the view that the present study needs replication with larger samples in the future.

Apart from the small sample size, a limitation lies in the rather rough language test (testing semantic and syntactic knowledge) which doesn't allow for a detailed analysis of the underlying language skills: The present study describes a relationship between grey matter changes in the IFG and increasing second language proficiency. The data analyzed in the present study however does not allow stating precisely which of the abilities related to higher second language proficiency is represented in this relationship: The present study cannot differentiate between the effects of language switching (Rodriguez-Fornells et al., 2002), controlled semantic retrieval (Miyashita, 2004; Segalowitz and Hulstijn, 2005) or language conflict management (see Abutalebi and Green, 2007 for a review). Furthermore, the current discourse on bilingualism includes evidence supporting the view that bilingualism has a generalized (i.e., not language-specific) effect on executive functioning (Bialystok, 2009). Future studies aiming to disentangle these effects should therefore not only operate with more precise language tests (allowing to differentiate between the above mentioned abilities associated with bilingualism) but also include measures of domain independent executive functioning (i.e., Bialystok and Martin, 2004).

Prior studies in other fields have indicated that "the denser the cortex, the better the performance" (Maguire et al., 2000; Draganski et al., 2004), others have linked lower density to better performance (Golestani et al., 2002). The present study stresses the importance of intra-individual changes: The intra-individual increases in density and proficiency correlated, while absolute density value was not linked to absolute second language proficiency. This indicates that the individual amount of learning is reflected in brain structure changes, regardless of absolute proficiency.

In summary, the present results show that within 5 months, second language learning is linked to grey matter increase in the left IFG. In this region, functional activation is known to decrease with increasing language proficiency. The structural changes in the IFG seen here macroscopically using MRI might reflect the building of a more specialized network, which in turn allows the processing of two languages without recruiting additional resources.

Acknowledgements

This research was supported by the Swiss National Competence Center on Research on Neuroplasticity and Repair. We also wish to thank: Rotary youth exchange, especially B. Fischer and W. Wyser, for great help in recruiting the subjects; all the exchange students for their readiness to participate; S. Früh for help in data acquisition.

Supplementary material

Supplementary data related to this article can be found online at doi:10.1016/j.cortex.2010.10.007.

REFERENCES

Abutalebi J and Green DW. Bilingual language production: The neurocognition of language representation and control. *Journal of Neurolinguistics*, 20: 242–275, 2007.

Ashburner J and Friston KJ. Voxel-based morphometry – The methods. *NeuroImage*, 11: 805–821, 2000.

Ashburner J and Friston KJ. Why voxel-based morphometry should be used. *NeuroImage*, 14: 1238–1243, 2001.

Bialystok E. Bilingualism: The good, the bad and the indifferent. *Bilingualism: Language and Cognition*, 12: 3–11, 2009.

Bialystok E and Martin MM. Attention and inhibition in bilingual children: Evidence from the dimensional change card sort task. *Developmental Science*, 7: 325–339, 2004.

Boyke J, Driemeyer J, Gaser C, Buchel C, and May A. Training-induced brain structure changes in the elderly. *Journal of Neuroscience*, 28: 7031–7035, 2008.

Chee MW, Hon N, Lee HL, and Soon CS. Relative language proficiency modulates bold signal change when bilinguals perform semantic judgments. Blood oxygen level dependent. *NeuroImage*, 13: 1155–1163, 2001.

Chee MW, Weekes B, Lee KM, Soon CS, Schreiber A, Hoon JJ, et al. Overlap and dissociation of semantic processing of Chinese characters, English words, and pictures: Evidence from fMRI. *NeuroImage*, 12: 392–403, 2000.

Chklovskii DB, Mel BW, and Svoboda K. Cortical rewiring and information storage. *Nature*, 431: 782–788, 2004.

Demb JB, Desmond JE, Wagner AD, Vaidya CJ, Glover GH, and Gabrieli JD. Semantic encoding and retrieval in the left inferior prefrontal cortex: A functional MRI study of task difficulty and process specificity. *Journal of Neuroscience*, 15: 5870–5878, 1995.

Draganski B, Gaser C, Busch V, Schuierer G, Bogdahn U, and May A. Neuroplasticity: Changes in grey matter induced by training. *Nature*, 427: 311–312, 2004.

Draganski B and May A. Training-induced structural changes in the adult human brain. *Behavioural Brain Research*, 192: 137–142, 2008.

Fiez JA. Phonology, semantics, and the role of the left inferior prefrontal cortex. *Human Brain Mapping*, 5: 79–83, 1997.

Forman SD, Cohen JD, Fitzgerald M, Eddy WF, Mintun MA, and Noll DC. Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): Use of a cluster-size threshold. *Magnetic Resonance in Medicine*, 33: 636–647, 1995.

Giedd JN, Blumenthal J, Jeffries NO, Castellanos FX, Liu H, Zijdenbos A, et al. Brain development during childhood and adolescence: A longitudinal MRI study. *Nature Neuroscience*, 2: 861–863, 1999.

Golestani N, Molko N, Dehaene S, LeBihan D, and Pallier C. Brain structure predicts the learning of foreign speech sounds. *Cerebral Cortex*, 17: 575–582, 2007.

Golestani N, Paus T, and Zatorre RJ. Anatomical correlates of learning novel speech sounds. *Neuron*, 35: 997–1010, 2002.

Good CD, Johnsrude IS, Ashburner J, Henson RN, Friston KJ, and Frackowiak RS. A voxel-based morphometric study of ageing in 465 normal adult human brains. *NeuroImage*, 14: 21–36, 2001.

Hua X, Leow AD, Levitt JG, Caplan R, Thompson PM, and Toga AW. Detecting brain growth patterns in normal children using tensor-based morphometry. *Human Brain Mapping*, 30: 209–219, 2009.

Lee H, Devlin JT, Shakeshaft C, Stewart LH, Brennan A, Glensman J, et al. Anatomical traces of vocabulary acquisition in the adolescent brain. *Journal of Neuroscience*, 27: 1184–1189, 2007.

Maguire EA, Gadian DG, Johnsrude IS, Good CD, Ashburner J, Frackowiak RS, et al. Navigation-related structural change in the hippocampi of taxi drivers. *Proceedings of the National Academy of Sciences of the United States of America*, 97: 4398–4403, 2000.

Mechelli A, Crinion JT, Noppeney U, O'Doherty J, Ashburner J, Frackowiak RS, et al. Neurolinguistics: Structural plasticity in the bilingual brain. *Nature*, 431: 757, 2004.

Miyashita Y. Cognitive memory: Cellular and network machineries and their top–down control. *Cognition and Behavior*, 306: 435–440, 2004.

Perani D and Abutalebi J. The neural basis of first and second language processing. *Current Opinion in Neurobiology*, 15: 202–206, 2005.

Perani D, Paulesu E, Galles NS, Dupoux E, Dehaene S, Bettinardi V, et al. The bilingual brain. Proficiency and age of acquisition of the second language. *Brain*, 121(Pt 10): 1841–1852, 1998.

Raichle ME, Fiez JA, Videen TO, MacLeod AM, Pardo JV, Fox PT, et al. Practice-related changes in human brain functional anatomy during nonmotor learning. *Cerebral Cortex*, 4: 8–26, 1994.

Richardson FM and Price CJ. Structural MRI studies of language function in the undamaged brain. *Brain Structure & Function*, 213: 511–523, 2009.

Richardson FM, Thomas MS, Filippi R, Harth H, and Price CJ. Contrasting effects of vocabulary knowledge on temporal and parietal brain structure across lifespan. *Journal of Cognitive Neuroscience*, 22: 943–954, 2010.

Rodriguez-Fornells A, Rotte M, Heinze HJ, Nosselt T, and Munte TF. Brain potential and functional MRI evidence for how to handle two languages with one brain. *Nature*, 415: 1026–1029, 2002.

Sakai KL. Language acquisition and brain development. *Science*, 310: 815–819, 2005.

Sakai KL, Miura K, Narafu N, and Muraishi Y. Correlated functional changes of the prefrontal cortex in twins induced by classroom education of second language. *Cerebral Cortex*, 14: 1233–1239, 2004.

Segalowitz N and Hulstijn J. Automaticity in bilingualism and secondlanguage learning. In Kroll J and De Groot A (Eds), *Handbook of Bilingualism: Psycholinguistic Approaches*. New York: Oxford University Press, 2005: 371–388.

Stein M, Dierks T, Brandeis D, Wirth M, Strik W, and Koenig T. Plasticity in the adult language system: A longitudinal electrophysiological study on second language learning. *NeuroImage*, 33: 774–783, 2006.

Stein M, Federspiel A, Koenig T, Wirth M, Lehmann C, Wiest R, et al. Reduced frontal activation with increasing 2nd language proficiency. *Neuropsychologia*, 47: 2712–2720, 2009.

Tatsuno Y and Sakai KL. Language-related activations in the left prefrontal regions are differentially modulated by age, proficiency, and task demands. *Journal of Neuroscience*, 25: 1637–1644, 2005.

Trachtenberg JT, Chen BE, Knott GW, Feng G, Sanes JR, Welker E, et al. Long-term in vivo imaging of experience-dependent synaptic plasticity in adult cortex. *Nature*, 420: 788–794, 2002.

van Praag H, Kempermann G, and Gage FH. Neural consequences of environmental enrichment. *Nature Reviews. Neuroscience*, 1: 191–198, 2000.

Vigneau M, Beaucousin V, Herve PY, Duffau H, Crivello F, Houde O, et al. Meta-analyzing left hemisphere language areas: Phonology, semantics, and sentence processing. *NeuroImage*, 30: 1414–1432, 2006.

Volterra A and Meldolesi J. Astrocytes, from brain glue to communication elements: The revolution continues. *Nature Reviews. Neuroscience*, 6: 626–640, 2005.

Yetkin O, Zerrin Yetkin F, Haughton VM, and Cox RW. Use of functional mr to map language in multilingual volunteers. *AJNR. American Journal of Neuroradiology*, 17: 473–477, 1996.