

Second-language learning and changes in the brain

Lee Osterhout^{a,*}, Andrew Poliakov^b, Kayo Inoue^a,
Judith McLaughlin^a, Geoffrey Valentine^a, Ilona Pitkanen^a,
Cheryl Frenck-Mestre^d, Julia Hirschensohn^c

^a*Department of Psychology, University of Washington, Seattle, WA 98195, USA*

^b*Department of Biological Structure, University of Washington, Seattle, WA 98195, USA*

^c*Department of Linguistics, University of Washington, Seattle, WA 98195, USA*

^d*Centre National de la Recherche Scientifique, Aix-Marseille University, Aix-en-Provence, France*

Received 4 April 2007; received in revised form 15 January 2008; accepted 15 January 2008

Abstract

Presumably, second-language (L2) learning is mediated by changes in the brain. Little is known about what changes in the brain, how the brain changes, or when these changes occur during learning. Here, we illustrate by way of example how modern brain-based methods can be used to discern some of the changes that occur during L2 learning. Preliminary results from three studies indicate that classroom-based L2 instruction can result in changes in the brain's electrical activity, in the location of this activity within the brain, and in the structure of the learners' brains. These changes can occur during the earliest stages of L2 acquisition.

© 2008 Elsevier Ltd. All rights reserved.

Keywords: Second language; Plasticity; ERPs; N400; P600; VBM; Language processing

1. Introduction

Experience can change both the function and the structure of the brain (Münste, Altenmüller, & Jäncke, 2002; van Praag, Kempermann, & Gage, 2000). Like other experiences, the experience of learning a second language (L2) is presumably accompanied by changes in the brain. It seems reasonable to presume that how, when, and where these changes occur is relevant (and possibly even essential) to a truly compelling understanding

*Corresponding author. Tel.: +1 206 329 8667.

E-mail address: losterho@u.washington.edu (L. Osterhout).

of L2 acquisition. At present, however, almost nothing is known about what changes in the brain during L2 learning, when these changes occur, and how they reflect L2 learning.

Fortunately, the current era is one of rapid methodological innovation with respect to non-invasive measurement of the human brain. The question of interest is whether these modern methods can detect changes in the brain that occur with L2 acquisition. Here, we describe preliminary results from ongoing experiments showing that these methods might in fact be sensitive to some of the brain changes that occur during L2 acquisition. Our preliminary data suggest these methods might be sensitive to changes over time in the brain's electrophysiological response to L2 stimuli, changes in the neural sources of that electrical activity, and even changes to the structure of the brain itself. Most of the work reviewed below involves longitudinal studies of novice, English-speaking L2 learners progressing through their first years of university-based L2 instruction. It seems likely, therefore, that the changes we report here are relevant to the classroom settings that typify L2 instruction in the United States and many other countries.

2. Changes in the brain's electrophysiological response to L2 stimuli

Aspects of the brain's electrophysiological activity can be recorded non-invasively from the scalp. For example, event-related brain potentials (ERPs) reflect synchronized postsynaptic activity in cortical pyramidal neurons. In our laboratory, we have used ERPs to track learning-related changes in brain function. In particular, we have examined the rate at which L2 knowledge is incorporated into the learner's on-line, real-time language comprehension system. To achieve this goal, we have recorded ERPs while learners read or listen to tokens of the L2 (McLaughlin, Osterhout, & Kim, 2004; Osterhout, McLaughlin, Pitkanen, Frenck-Mestre, & Molinaro, 2006). Our work has primarily involved longitudinal studies that assess changes in the brain response to L2 sentences that occur during the earliest stages of L2 learning. This approach was motivated by prior work showing that certain linguistic manipulations elicit robust effects in the ERP. The crucial finding has been that syntactic and semantic anomalies elicit qualitatively distinct ERP effects, and that these effects are characterized by distinct temporal properties. Semantic anomalies (e.g., *The cat will bake the food ...*) elicit a negative wave that peaks at about 400 ms after the anomalous word appears (the *N400 effect*; Fig. 1a) (Kutas & Hillyard, 1980; Osterhout & Nicol, 1999). By contrast, syntactic anomalies (e.g., *The cat will eating the food ...*) elicit a large positive wave that onsets at about 500 ms after presentation of the anomalous word and persists for at least half a second (the *P600 effect*; Fig. 1b) (Osterhout & Holcomb, 1992, 1993; Osterhout & Mobley, 1995; Osterhout & Nicol, 1999). In some studies, syntactic anomalies have also elicited a negativity over anterior regions of the scalp, with onsets ranging from 100 to 300 ms (Friederici, 1995; Osterhout & Holcomb, 1992).

These results suggest that separable syntactic and semantic processes exist. In an L2 learning context, one implication is that L2 learners must somehow segregate linguistic input into those aspects of the language that relate to sentence form and those that relate to sentence meaning. That is, learners "grammaticalize" some aspects of the L2, but not others. In our work, what we mean by "grammaticalization" is specifically the instantiation of grammatical knowledge into the learner's on-line, real-time language processing system. Our assumption is that, once a feature of the L2 has been grammaticalized, violations of that aspect of the grammar should elicit a P600 effect.

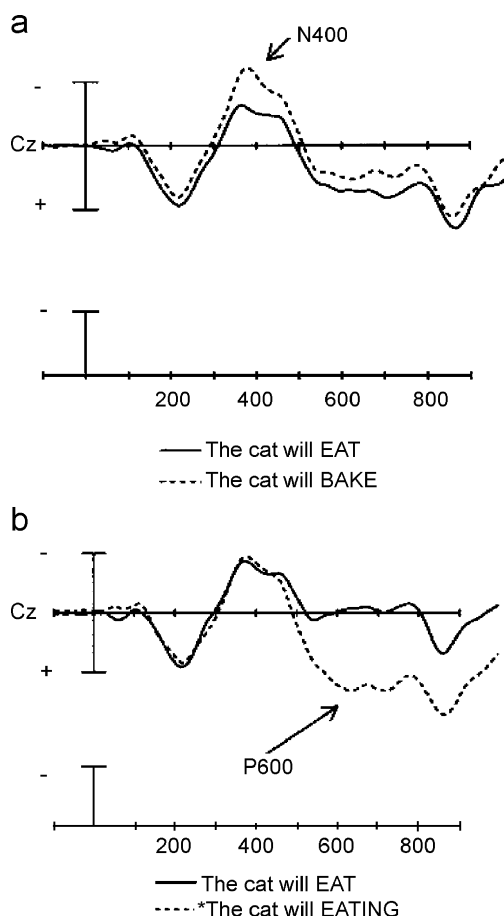


Fig. 1. ERPs, recorded at electrode site Cz. Onset of critical word is indicated by vertical calibration line, which represent 5 μV. Negative voltage is plotted up. Each hashmark indicates 100 ms. (a) ERP response to semantically well-formed (solid line) and anomalous (dashed line) critical words. (b) ERP response to syntactically well-formed (solid line) and anomalous (dashed line) critical words (adapted from Osterhout & Nicol, 1999).

To investigate grammaticalization during L2 learning, we have focused on the acquisition of grammatical features and their associated morphosyntactic rules. These features (and how they are involved in morphosyntax) vary across languages. For example, English and French both have sentential agreement (i.e., agreement of the verb with the subject in verbal person and number; e.g., *I like* vs. *He likes*), but only French has noun phrase agreement, that is, agreement between the noun and its determiner/adjective in number and gender (e.g., *le garçon* vs. *les garçons* ‘the-masc-sg boy, the-pl boys’, excluding the restricted English case of *this/those/these*).

What factors might inhibit or facilitate grammaticalization of these features and their morphosyntactic rules? One frequent claim is that only grammatical features that are present in the L1 can be acquired during L2 acquisition (Hawkins & Franceschina, 2004). Other researchers argue that novel L2 features can be learned, albeit more slowly than

those that are present in the L1 (White, 2003). Thus, there is no consensus about whether, or when during acquisition, L2 learners acquire L2 features and morphosyntactic rules that are not present in their L1.

Another factor that seems likely to play a role in L2 grammatical morpheme learning is the covariation between morphology and phonology. For example, French has an opaque orthography due to many suffixes being phonologically silent. Thus, the plural suffix—*s*, which marks the plural orthographically across all elements in the NP (*le-s jeune-s fille-s* ‘the young girls’) is silent on the noun, determiner, and adjective in almost all instances. A similar situation arises in the verb phrase (VP), where variations in verbal person are marked orthographically on the verb but are silent in most oral forms. Thus, the different inflections for a regular verb such as *marcher* (to walk) in present tense sound identical across four different persons/spellings. The effect of the ‘missing’ phonological cue is notorious on spelling. Errors such as *les chien* or *Ils mange* are frequent (Negro & Chanquoy, 2000), and are often made by native-French-speaking adults as well as by children. Such errors are much rarer when phonology is available as a cue (Largy & Fayol, 2001).

Given this evidence, a reasonable prediction is that L2 learners will acquire an L2 feature or morphosyntactic rule more quickly when the relevant inflectional morphology is phonologically realized. However, this possibility has received little direct attention in the recent L2 literature. It also seems likely that L1–L2 similarity and phonological–morphological covariation might have interactive effects during L2 learning. For example, L1–L2 similarity combined with phonological realization of the relevant grammatical morphemes might lead to very fast learning, whereas L1–L2 dissimilarity combined with no phonological realization might lead to very slow learning.

We investigated these predictions using a longitudinal experimental design involving 14 English-speaking novice French learners progressing through their first year of French instruction at the University of Washington.¹ Our stimuli were as follows:

(1) Sept plus cinq?livre font douze.
‘seven plus five/book make twelve’

semantic condition

(2) Tu adores?*adorez le français.
‘you-2-sg adore-2-sg/adore-2-pl the French’

verbal person agreement condition/phonologically realized

(3) Tu manges des hamburgers?*hamburger pour diner.
‘you-2-sg eat-2-sg some-pl hamburgers-pl/hamburger-sg for dinner’

number agreement condition/phonologically unrealized

In (1), the noun *livre* is semantically anomalous. In (2), the verb *adorez* is conjugated incorrectly, given the preceding sentence fragment. In (3), the noun *hamburger* disagrees with the syntactic number of the plural article. Our stimuli were selected from the material in the textbook assigned during the first month of instruction. The anomalous items in the verbal person condition involved a grammatical rule that was present in the L1 and an orally realized contrast between inflectional morphemes. The anomalous items in the

¹A similar preliminary report of this experiment appeared in Osterhout et al. (2006).

number agreement condition involved a rule that was not present in the L1 and a phonologically unrealized contrast between inflectional morphemes. Therefore, our prediction was that L2 learners would respond to the anomaly in (2) with less L2 exposure compared to the anomaly in (3).

For each condition, subjects read 30 exemplars each of anomalous and well-formed control sentences. Sentences were counterbalanced across lists, such that each subject saw only one version of a particular sentence frame. The 180 sentences in each list were pseudorandomly ordered prior to presentation. Sentences were presented word-by-word on a computer screen, with each word being presented for 350 ms and with a 300 ms blank-screen interval between words. The final word in each sentence was followed by a 1450-ms blank screen interval, after which the subject was prompted to make a “sentence acceptability” judgment about the preceding sentence. Continuous EEG was recorded from 13 scalp sites and averaged off-line.

As expected, the native French speakers showed an N400 effect to the semantically anomalous words (1) and large P600 effects to the two types of syntactic anomalies (2)–(3). The learners, as is often the case, showed striking individual differences, both in a behavioral “sentence acceptability judgment” task and in the pattern of ERPs elicited by the anomalous stimuli. We segregated the learners into upper (“fast learners”, $n = 7$) and lower (“slow learners”, $n = 7$) halves, based on their performance in the sentence-acceptability judgment task (mean d -primes for the session 3 sentence acceptability judgments, averaged over the three conditions, were 2.7 and 1.5 for the fast and slow learners, respectively). ERPs were then averaged separately for each group. Results for the “fast learner” group will be described here. At each testing session, including the initial session that occurred after just 1 month of instruction, semantically anomalous words elicited a robust N400 effect (midline electrodes: $F(1, 6) = 24.61, p < .003$), and this effect changed minimally with increasing instruction (d -primes for sentence-acceptability task were 2.0, 3.0, and 3.1 for sessions 1, 2, and 3). Results for the verbal person agreement condition are shown in Fig. 2. After just 1 month of instruction, the learners’ brains discriminated between the syntactically well-formed and ill-formed sentences (midline electrodes: $F(1, 6) = 5.58, p = .05$). However, rather than eliciting the P600 effect (as we saw in native French speakers), the syntactically anomalous words elicited an N400-like

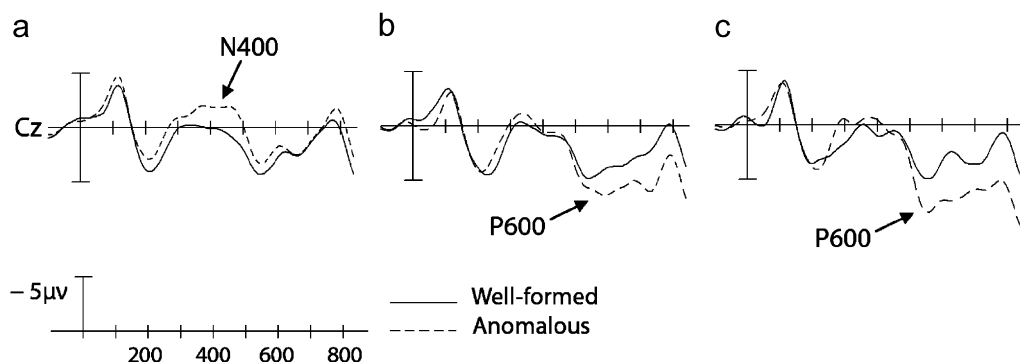


Fig. 2. ERPs to the French (L2) verbal person agreement condition, for fast learners. (a) Session 1. (b) Session 2. (c) Session 3.

effect. (This effect did not differ in distribution from the N400 effect elicited by the semantically anomalous words.) By 4 months, the N400 effect was replaced by a P600-like positivity (midline electrodes: $F(1, 6) = 8.73$, $p < .03$; d-primes were 2.0, 3.5, and 3.5 for sessions 1, 2, and 3). Results for the nominal number agreement condition can be summarized easily: learners performed very poorly in the sentence acceptability judgment task for these materials (d-prime = 0.5, 1.5, and 1.6 for sessions 1, 2, and 3), and there were no robust differences in the ERP responses to the agreeing and disagreeing stimuli.

These results are consistent with the predictions we were testing. First, L1–L2 similarity combined with phonological realization of the relevant grammatical morphemes produced very fast L2 syntactic learning, whereas L1–L2 dissimilarity combined with no phonological realization produced very slow learning. This occurred even though our learners were drilled repeatedly on both rules from nearly the first day in class. However, the two rules we tested represent the ends of a putative continuum of morphosyntactic difficulty; without additional data it is impossible to know whether L1–L2 similarity or phonological realization of grammatical morphemes had a larger impact on the learning rate.

We also observed a discontinuous pattern over time in the response to the verbal person anomalies: early in learning, these anomalies elicited an N400 effect in learners, whereas later in learning these same anomalies elicited a P600 effect. Our hypothesis is that our learners were progressing through discrete stages of syntactic learning: they began by memorizing particular combinations of words and morphemes, and only later induced general syntactic rules (Myles, Hooper, & Mitchell, 1998; Wray, 2001; see Tomasello, 2000, for evidence that children go through similar stages during L1 acquisition). To be specific, an L2 learner might initially memorize the fact that certain subjects are followed by certain forms of the verb, without decomposing the verb into root + inflection or applying a general morphosyntactic agreement rule. In this stage of learning, the learner associates meanings with the undecomposed chunk of language, and either memorizes the two words as a chunk or learns about word sequence probabilities (e.g., that *Tu* ‘you-2-sg’ is followed by *marches* ‘walk-2-sg’, whereas *Ils* ‘they-3-pl’/*Nous* ‘we-1-pl’ is followed by *marchent* ‘walk-3-pl’/*marchons* ‘walk-1-pl’). Violations of the verbal person rule (e.g., *tu *adorez*) result in novel word combinations, and hence elicit an N400 effect. After more instruction, learners induce a general verbal person rule (*tu -s*, *nous -ons*, *vous -ez*, etc.); violations of the rule elicit a P600 effect. If our interpretation is correct, then our adult L2 learners grammaticalized this aspect of the L2 after just a few months (~80 h) of L2 instruction.

3. Changes in brain sources

Ideally, one would like to discern the neural sources of these language-sensitive ERP effects. By doing so, one might be able to describe changes in the brain across both time and space. Unfortunately, the source of a given ERP effect (the “inverse solution”) cannot be known with certainty. This follows from the fact that a large number of source configurations could produce an identical pattern of activity across the scalp (Nunez & Srinivasan, 2006). Nonetheless, source estimates are possible given certain limiting assumptions. The traditional approach has been to search for point dipole sources (Hämäläinen & Sarvas, 1989; Henderson, Butler, & Glass., 1975). In general, this entails assuming a small number of dipole sources and iterating through all possible combinations

of dipole location, orientation, and strength, looking for the best match between the source model and the observed scalp distribution. This method brings with it numerous limitations and caveats (Halgren et al., 2002).

More recently developed alternative methods provide a tomographic analysis analogous to that provided by neuroimaging methods, but with vastly superior temporal resolution (and reduced spatial resolution; Michel et al., 1999). These methods, known as *distributed source models*, provide a linear inverse solution (unlike the traditional nonlinear dipole approach) to estimate the current distribution throughout the entire three-dimensional cortex. But in order to reach a unique solution, these models require assumptions that might not be valid. If the assumptions are not valid, then various errors can occur: simultaneously active areas can be missed, spurious sources can be generated, and the results can be excessively blurred (for extensive discussion, see Michel et al., 1999).

In our lab, we have been using one type of distributed source model, Low Resolution Electromagnetic Tomography (LORETA; Pascual-Marqui, Michel, & Lehmann, 1994), to estimate the sources associated with processing sentences in the L1 and L2. The primary assumption of LORETA is that dramatic changes in current do not occur across contiguous areas of cortex (i.e., in adjacent voxels). In practice, this assumption often produces an excessively blurred solution. Because in our study the L1–L2 contrast is within subject, any distortions that result from violations of the smoothness assumption should be relatively constant across languages. Therefore, differences in the source distribution associated with processing sentences in two languages might be meaningful (or at least roughly indicative) even if the exact location of the sources is somewhat distorted.

In an ongoing study, we are recording ERPs from native English-speaking university students who are enrolled in L2 (French) instruction. Presently, we have recorded ERPs from 22 students who were approximately halfway through their second year of French instruction. The stimuli were sentences in the “semantic” and “verbal person agreement” conditions in the L2 ERP experiment described above, plus approximate English translations of the French sentences. Counterbalancing was used such that each subject saw only one version of each sentence across well-formed vs. ill-formed contrasts and L1 vs. L2 contrasts. The methods were identical to those described, except that subjects were tested once with L1 (English) stimuli and once with L2 (French) stimuli; the order of languages was counterbalanced across subjects. Subjects were tested in one session lasting between 2 and 3 h, with a break between the L1 and L2 lists. EEG was recorded from 61 channels rather than 13; the greater spatial sampling is crucial for better source estimates. As expected, the semantically and syntactically anomalous words elicited N400 and P600 effects, for both English (L1) and French (L2) (midline electrodes: semantic, $F(1, 19) = 9.02$, $p < .01$; syntactic L1, $F(1, 19) = 47.54$). Interestingly, the P600 effect was much larger for the L1 than the L2 ($F(1, 19) = 5.28$, $p < .04$), whereas the N400 effect was similar across the two languages ($F(1, 19) = 1.91$, $p > 1.8$).

We then used LORETA to estimate the current distribution associated with *normal* sentence processing (rather than the processing of anomalous sentences, which has been the strategy in previously published work) within two critical time windows: the window associated with the N400 component (during which the brain is most robustly sensitive to conceptual aspects of the stimulus) and the window associated with the P600 effect (during which the brain is most robustly sensitive to syntax). LORETA analyses on the critical words in the well-formed verbal person agreement condition are shown in Fig. 3. Two solutions are shown for each word, for representative time points in the N400 (300–500 ms)

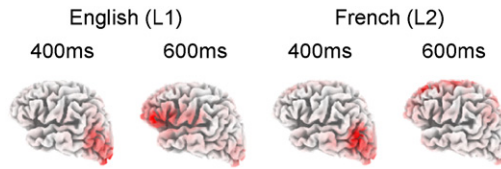


Fig. 3. LORETA solutions at 400 and 600 ms to critical words in the well-formed versions of sentences in the verbal person agreement condition (left panel = English (L1) sentences; right panel = French (L2) sentences).

Table 1

Brodmann's Area (BA) locations of the five voxel clusters (in descending order of current density) with the most current density, as estimated by LORETA, at 400 and 600 ms after presentation of critical words in the well-formed verbal person agreement condition

	English (L1)		French (L2)	
	400 ms	600 ms	400 ms	600 ms
1	BA17, L	BA45, L	BA19, R	BA8, L
2	BA17, L	BA44, L	BA39, L	BA9, L
3	BA18, L	BA44, R	BA18, R	BA6, R
4	BA39, L	BA10, L	BA19, L	BA37, R
5	BA39, R	BA17, L	BA17, L	BA6, L

L: left hemisphere and R: right hemisphere.

and P600 (500–800 ms) windows. For English (L1) stimuli, the LORETA solutions indicated a posterior distribution for the N400 window (including temporoparietal and extrastriate regions), and an anterior distribution for the P600 window (especially the left inferior frontal gyrus). The LORETA solutions for the same critical words in the French (L2) sentences are shown on the right half of Fig. 3; clusters of current density are listed in Table 1. The L2 LORETA solutions in the N400 window were similar to those for the L1 stimuli, with a posterior current distribution for both nouns and verbs. However, the L2 LORETA solution in the P600 window were quite different from the L1 solutions; current density was greatest in the medial dorsal frontal lobe, rather than in the inferior frontal gyrus. Any interpretation would be premature; nonetheless, it is interesting to note that posterior lesions involving the left temporoparietal region often produce a deficit in processing word meanings, whereas lesions to the left inferior frontal lobe often produce agrammatic symptoms. Perhaps L2 learners quickly incorporate aspects of L2 word meaning into their on-line processing system, and engage similar neural circuits that access and integrate these meanings. L2 syntactic rules might take longer to incorporate in this way. One question we hope to answer in this longitudinal design is how much L2 experience is needed before the LORETA solutions become more like the L1 responses (e.g., the point at which L2 verbs engage the “anterior processing stream”). In additional experiments involving first- and third-year French students, our preliminary results indicate that the L2 solutions might become progressively more like the L1 solutions, across the first 3 years of L2 instruction.

4. Changes in brain structure

Anatomical neuroimaging methods can be used to study the effects of learning on brain structures *in vivo*. These methods include morphometric and volumetric techniques that measure certain parameters of well-defined brain structures (for example, hippocampal volume or cortical thickness). More recently, a number of unbiased, whole-brain techniques have been developed. One such technique is voxel-based morphometry (VBM; Mechelli, Price, Friston, & Ashburner, 2005). In its simplest application, VBM involves a voxel-by-voxel comparison of the relative concentration of gray matter (GM) and white matter (WM). This is achieved using whole-brain MRI scans.

VBM was originally devised to examine structural abnormalities in patients. Recent work using neurologically normal subjects, however, has shown that VBM can reveal the effects of learning and practice on brain structure. For example, Maguire et al. (2000) used VBM to test whether structural changes could be detected in the brains of London taxi drivers as a result of extensive experience with spatial navigation. The VBM data indicated that the posterior hippocampi of taxi drivers were significantly larger than those of healthy controls. Furthermore, hippocampal volume correlated with the amount of time spent as a taxi driver. These results suggest that the posterior hippocampi may expand regionally in individuals who have extensive experience with spatial navigation. However, an alternative explanation could attribute this result to selection rather than plasticity (e.g., people with larger posterior hippocampi might gravitate to spatially demanding occupations). Longitudinally designed VBM studies can help reduce such uncertainties. For example, Draganski et al. (2004) demonstrated that acquisition of new skills may indeed change GM density. Brain scans were acquired from healthy subjects before they learned juggling and 3 months later when they had become skilled performers. The comparison of the scans acquired before and after practice revealed an expansion in GM in bilateral mid-temporal areas and left posterior intra-parietal sulcus. These findings were specific to the training stimulus, as a group of controls showed no changes in GM over the same period. More recently, Draganski et al. (2006) described GM density changes in a group of medical students studying for their final examination. Although these findings are intriguing, the mechanism of changes detected by VBM is still poorly understood, particularly in healthy subjects. Methods other than whole-brain MRI will be required to investigate whether the gray-matter increases induced by the acquisition of new skills are related to changes in neuropil, neuronal size, and/or dendritic or axonal arborization.

Just a few studies have used VBM to investigate possible effects of language learning on brain structure. Mechelli et al. (2004) demonstrated that systematic measurable structural changes are present in the brains of individuals who have learned multiple languages, L2 learners' brains, and that the degree of structural variation depends on the age of acquisition of the second language and the proficiency attained (as well as on the number of languages acquired). Golestani et al. (2006) described anatomical correlates of learning novel speech sounds studied using sulcul anatomy and VBM of white as well as GM. Although the results from these VBM-and-language studies are quite interesting, they have employed cross-sectional designs. Cross-sectional VBM designs introduce certain problems of interpretation (for discussion see Ashburner & Friston, 2001; Bookstein, 2001; Mechelli et al., 2004), and also uncertainty as to whether VBM differences are due to "learning" or to "pre-selection."

In our lab, we have completed a longitudinal VBM pilot study of L2 learning. Our four subjects were university students enrolled in an intensive first-year Spanish course. The course lasted 9 weeks and met daily for $3\frac{1}{2}$ h of instruction. In our longitudinal design, each learner participated in two MRI sessions of six MRI scans each. Session 1 occurred after 20 or 21 days of instruction. Session 2 occurred between 1 day before and 3 days after the end of instruction. For each subject, approximately 5 weeks of L2 instruction intervened between Sessions 1 and 2. The data from the four subjects (48 structural T1 scans total) were subsequently preprocessed and analyzed. We used the SPM5 software package and the VBM5 toolkit (C. Gaser) that implements enhancements and modifications needed for more advanced VBM analysis techniques.

Due to the small sample size ($n = 4$), we did not expect to find (nor did we find) statistically significant changes in GM density using an anatomically open hypothesis (i.e., statistical analyses not constrained to any anatomical region). Instead, we used a region of interest approach based on the results of the Mechelli et al. (2004) cross-sectional VBM study of bilingual and multilingual speakers. Mechelli et al. report that the amount of experience with a second language correlates with GM changes in a left-hemisphere inferior parietal region (MNI coordinates: $x = -48$, $y = -62$, $z = 44$). In our VBM analysis, we used a small volume correction with a spherical 10 mm region of search centered at the MNI coordinates reported by Mechelli et al. This analysis detected a cluster of significant ($p < .05$) GM density increase. The $K E$ value reported for the cluster-level statistics of VBM was 373. This result suggests that structural changes resulting from learning a second language can be observed after a relatively short but intense instructional period, even in a small sample of learners. Of course, these results are preliminary and await replication in a larger-scale study.

Each method of investigation comes attached with cautions and limitations, and this is certainly true of VBM and related methods. For example, significant effects might be attributable to scanner or MR sequence differences rather than to the subjects themselves. Other potential issues relate to the difficulty of spatially normalizing brains, the robustness of standard parametric tests and the interpretation of the results. Some apparently volumetric differences may be due to displacement without volume change. Any factor which causes changes in voxel intensities in the original images (e.g., vascular changes, changes in hydration status leading to relaxation time changes, etc.) may result in significant VBM differences.

One other limitation of VBM deserves particular mention: tissue segmentation (into GM and WM) is easier and more reliable in some parts of the brain than in other parts. Segmentation results will inevitably be affected by the presence of other tissues with similar contrast (for example, dura matter, large blood vessels, etc.) and imaging artifacts (movement related artifacts including blood flow and eye movement artifacts, susceptibility artifacts near paranasal sinuses and ear canals and many others).

To illustrate this point, we assessed the variability of GM density maps across several scans within a session using data from our pilot VBM study described above. Variability for one subject/session was estimated by computing the standard deviation of GM density maps estimated from each of the six scans of a single session. These were normalized and averaged to produce a map shown in Fig. 4. This color-coded map is overlaid onto a GM probability density map. Yellow and red areas represent increased variability in GM density estimates. Fig. 4 clearly indicates that variability is much higher in some parts of the brain (e.g., some regions in the temporal and parietal lobes) than in others;

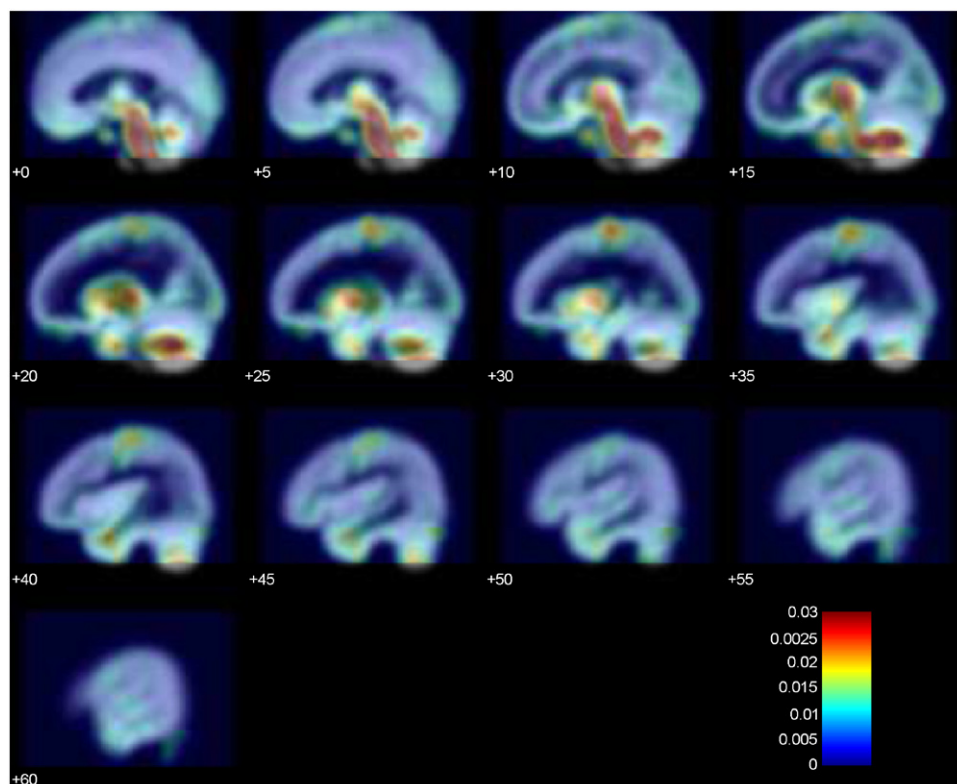


Fig. 4. Variability in VBM signal. A color-coded map depicting heterogeneity in variability is overlaid onto a GM probability density map. Yellow and red areas represent increased variability in GM density estimates.

correspondingly, changes in GM density will be more observable where the variability is relatively low.

5. Conclusions

Our results suggest that the brain of an adult second-language learner is a highly dynamic place, even during the earliest stages of L2 learning. Our results also suggest that modern methods are capable of revealing at least some of these changes. In particular, the methods used here are seemingly sensitive to changes in the brain's electrical activity, changes in the location of this activity within the brain, and changes in the structure of the learners' brains. Of course, our results are preliminary and await completion of the ongoing work, as well as subsequent replication.

At first glance, our results might seem surprising. A large literature has shown that the ability to learn a language seems to degrade with age; our learners were not young children but instead were young adults. The causes of age effects on L2 proficiency are controversial, but two frequently cited theoretical explanations stand out. One explanation involves the putative existence of a critical period for L2 learning (Johnson & Newport, 1989). According to this explanation, language learning is constrained by maturational

factors (specifically, brain maturation) that circumscribe a critical period (which by most accounts ends around puberty) for native-like attainment. A second explanation attributes the age-related declines to the effects of increasing experience with a first language (Kuhl, 2004). Neural network simulations provide a convenient framework for understanding effects of L1 usage on L2 learning. In these simulations, early learning results in the entrenchment of optimal network patterns, after which new learning requires considerable training. Consistent with this view, it has been demonstrated that early experience with certain aspects of a first language (e.g., phonemes) seems to degrade the ability to learn aspects of a second language later in life (Kuhl, 2004).

Both of these explanations implicate the same underlying cause for age-related effects on language learning, namely, a reduction in neural plasticity that degrades the ability to learn and retain new linguistic information. However, despite the popularity of this view, there is little direct evidence to support it. The hypothesis also seems to be inconsistent with accumulating evidence that (at least in many respects) the brain remains remarkably plastic throughout much of life (van Praag et al., 2000). The findings reported here are clearly too preliminary and insufficiently detailed to inform this debate in a substantive way. But these findings do allow us to hope that the methods in hand might, eventually, tell us something useful about the degree of neural plasticity in older learners and the role it plays in enabling and/or hindering L2 acquisition.

Acknowledgments

Preparation of this article was supported by grants R01DC01947 and P30DC04661 from the National Institute on Deafness and Other Communication Disorders.

References

- Ashburner, J., & Friston, K. J. (2001). Why voxel-based morphometry should be used. *NeuroImage*, *14*, 1238–1243.
- Bookstein, F. L. (2001). Voxel-based morphometry should not be used with imperfectly registered images. *NeuroImage*, *14*, 1454–1462.
- Draganski, B., Gaser, C., Busch, V., Schuierer, G., Bogdahn, U., & May, A. (2004). Neuroplasticity: Changes in grey matter induced by training. *Nature*, *427*, 311–312.
- Draganski, B., Gaser, C., Kempermann, G., Kuhn, H. G., Winkler, J., Buchel, C., et al. (2006). Temporal and spatial dynamics of brain structure changes during extensive learning. *Journal of Neuroscience*, *26*, 6314–6317.
- Friederici, A. D. (1995). The time course of syntactic activation during language processing: A model based on neuropsychological and neurophysiological data. *Brain and Language*, *50*, 259–284.
- Golestani, N., Molko, N., Dehaene, S., Le Bihan, D., & Pallier, C. (2007). Brain structure predicts the learning of foreign speech sounds. *Cerebral Cortex*, *17*, 575–582.
- Halgren, E., Dhond, R. P., Christensen, N., Van Petten, C., Marinkovic, K., Lewine, J., et al. (2002). N400-like magnetoencephalography responses modulated by semantic context, word frequency, and lexical class in sentences. *NeuroImage*, *17*, 1101–1116.
- Hämäläinen, M. S., & Sarvas, J. (1989). Realistic conductivity geometry model of the human head for interpretation of neuromagnetic data. *IEEE Transactions on Biomedical Engineering*, *36*, 165–171.
- Hawkins, R., & Franceschina, F. (2004). Explaining the acquisition and non-acquisition of determiner-noun gender concord in French and Spanish. In P. Prevost, & J. Paradis (Eds.), *The acquisition of French in different contexts*. Amsterdam: Benjamins.
- Henderson, C. J., Butler, S. R., & Glass, A. (1975). The localization of equivalent dipoles of EEG sources by the application of electrical field theory. *Electroencephalography and Clinical Neurophysiology*, *39*, 117–130.
- Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, *21*, 60–99.

- Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews: Neuroscience*, 5, 831–843.
- Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic anomaly. *Science*, 207, 203–205.
- Largy, P., & Fayol, M. (2001). Oral cues improve subject-verb agreement in written French. *International Journal of Psychology*, 36, 121–131.
- Maguire, E. A., Gadian, D. G., Johnsrude, I. S., Good, C. D., Ashburner, J., Frackowiak, R. S. J., et al. (2000). Navigation-related structural change in the hippocampi of taxi drivers. *Proceedings of the National Academy of Sciences*, 97, 4398–4403.
- McLaughlin, J., Osterhout, L., & Kim, A. (2004). Neural correlates of second-language word learning: Minimal instruction produces rapid change. *Nature Neuroscience*, 7, 703–704.
- Mechelli, A., Price, C. J., Friston, K. J., & Ashburner, J. (2005). Voxel-based morphometry of the human brain: Methods and applications. *Current Medical Imaging Review*, 1, 105–113.
- Mechelli, A., Crinion, J. T., Noppeney, U., O'Doherty, J., Ashburner, J., Frackowiak, R. S., et al. (2004). Neurolinguistics: Structural plasticity in the bilingual brain. *Nature*, 431, 757.
- Michel, C. M., de Peralta, R. G., Lantz, G., Andino, S. G., Spinelli, L., Blanke, O., et al. (1999). Spatiotemporal EEG analysis and distributed source estimation in presurgical epilepsy evaluation. *Journal of Clinical Neurophysiology*, 16, 239–266.
- Münte, T. F., Altenmüller, E., & Jäncke, L. (2002). The musician's brain as a model of neuroplasticity. *Nature Reviews Neuroscience*, 3, 473–478.
- Myles, F., Hooper, J., & Mitchell, R. (1998). Rote or rule? Exploring the role of formulaic language in classroom foreign language learning. *Language Learning*, 48, 323–363.
- Negro, I., & Chanquoy, L. (2000). Subject-verb agreement with present and imperfect tenses: A developmental study from 2nd to 7th grade. *European Journal of Psychology of Education*, 15, 113–134.
- Nunez, P. L., & Srinivasan, R. (2006). *Electric fields of the brain*. Oxford: Oxford University Press.
- Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language*, 31, 785–806.
- Osterhout, L., & Holcomb, P. J. (1993). Event-related potentials and syntactic anomaly: Evidence of anomaly detection during the perception of continuous speech. *Language and Cognitive Processes*, 8, 413–438.
- Osterhout, L., & Mobley, L. A. (1995). Event-related brain potentials elicited by failure to agree. *Journal of Memory and Language*, 34, 739–773.
- Osterhout, L., & Nicol, J. (1999). On the distinctiveness, independence, and time course of the brain responses to syntactic and semantic anomalies. *Language and Cognitive Processes*, 14, 283–317.
- Osterhout, L., McLaughlin, J., Pitkanen, I., Frenck-Mestre, C., & Molinaro, N. (2006). Novice learners, longitudinal designs, and event-related potentials: A paradigm for exploring the neurocognition of second-language processing. *Language Learning*.
- Pascual-Marqui, R. D., Michel, C. M., & Lehmann, D. (1994). Low resolution electromagnetic tomography: A new method for localizing electrical activity in the brain. *International Journal of Psychophysiology*, 18, 49–65.
- Tomasello, M. (2000). The item-based nature of children's early syntactic development. *Trends in Cognitive Sciences*, 4, 156–160.
- Van Praag, H., Kempermann, G., & Gage, F. H. (2000). Neural consequences of environmental enrichment. *Nature Reviews Neuroscience*, 1, 191–199.
- White, L. (2003). *Second Language Acquisition and Universal Grammar*. Cambridge: Cambridge University Press.
- Wray, A. (2001). *Formulaic language*. Cambridge: Cambridge University Press.