



ACADEMIC
PRESS

Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

NeuroImage

NeuroImage 19 (2003) 113–124

www.elsevier.com/locate/ynimg

Learning-induced neural plasticity associated with improved identification performance after training of a difficult second-language phonetic contrast

Daniel E. Callan,^{a,*} Keiichi Tajima,^a Akiko M. Callan,^b Rieko Kubo,^a Shinobu Masaki,^b and Reiko Akahane-Yamada^a

^a Human Information Science Laboratories, ATR International, Kyoto 619-0288, Japan

^b Brain Activity Imaging Center, ATR International, Kyoto 619-0288, Japan

Received 22 July 2002; accepted 10 November 2002

Abstract

Adult native Japanese speakers have difficulty perceiving the English /r-l/ phonetic contrast even after years of exposure. However, after extensive perceptual identification training, long-lasting improvement in identification performance can be attained. This fMRI study investigates localized changes in brain activity associated with 1 month of extensive feedback-based perceptual identification training by native Japanese speakers learning the English /r-l/ phonetic contrast. Before and after training, separate functional brain imaging sessions were conducted for identification of the English /r-l/ contrast (difficult for Japanese speakers), /b-g/ contrast (easy), and /b-v/ contrast (difficult), in which signal-correlated noise served as the reference control condition. Neural plasticity, denoted by exclusive enhancement in brain activity for the /r-l/ contrast, does not involve only reorganization in brain regions concerned with acoustic-phonetic processing (superior and medial temporal areas) but also the recruitment of additional bilateral cortical (supramarginal gyrus, planum temporale, Broca's area, premotor cortex, supplementary motor area) and subcortical regions (cerebellum, basal ganglia, substantia nigra) involved with auditory-articulatory (perceptual-motor) mappings related to verbal speech processing and learning. Contrary to what one may expect, brain activity for perception of a difficult contrast does not come to resemble that of an easy contrast as learning proceeds. Rather, the results support the hypothesis that improved identification performance may be due to the acquisition of auditory-articulatory mappings allowing for perception to be made in reference to potential action.

© 2003 Elsevier Science (USA). All rights reserved.

Introduction

Adult speakers have great difficulty learning to identify certain second-language L2 phonetic contrasts that their native language does not use. A classic example of this is the English /r-l/ contrast for native Japanese speakers. Even after years of exposure to English, it has been generally found that native Japanese adults cannot reliably identify and discriminate minimal pairs of words contrasting in /r/ and /l/ (e.g., *rake* and *lake*) (Miyawaki et al., 1975). How-

ever, extensive perceptual identification training with feedback, using a number of different speakers as stimuli, can improve forced choice identification performance of /r/ and /l/ approximately 20 percentage points (Akahane-Yamada, 1996). The training has been found to generalize to novel stimuli, is long lasting (Lively et al., 1994), and also improves speech production of English /r/ and /l/ even though no formal production training is given (Lively et al., 1994). Native Japanese listeners who are acquiring English as a second language must somehow reorganize and/or recruit neural processes in a manner such that /r-l/ identification performance is improved. Brain regions that undergo reorganization and/or recruitment are conjectured to be reflected by learning-induced neural plasticity.

The primary objective of the current study is to deter-

* Corresponding author. Human Information Science Laboratories, ATR International, 2-2-2 Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-0288, Japan. Fax: +81-774-95-1085.

E-mail address: dcallan@atr.co.jp (D.E. Callan).

mine the brain regions involved with perceptual learning of a difficult L2 phonetic contrast using fMRI. Brain imaging studies investigating learning-induced plasticity have shown decreases, increases, as well as shifts in the location of activity (as measured by the blood oxygen level-dependent, BOLD, response) after relative to before training (Chollet, 2000; Poldrack, 2000; Van Mier, 2000). Decreases denote activity that is greater before training than after training. Decreases are thought to reflect a sharpening of responses such that a minority of neurons fire more vigorously, whereas the majority of neurons show decreased firing (Poldrack, 2000). Increases denote activity that is present before training but is greater after training. Increases are thought to reflect an expansion of cortical representations, as is seen in auditory cortex of monkeys after frequency discrimination training (Recanzone, 2000). Shifts denote activity that is present after training that is not present before training. Shifts are thought to reflect the use of new neural processes or representations. The type of task, the extent of learning, and the brain region involved all appear to be factors determining whether there is a decrease, increase, or shift in activity (Chollet, 2000; Poldrack, 2000; Van Mier, 2000). Although there are exceptions (Jancke et al., 2001), decreases in activity are generally found when training and exposure to the stimuli are less extensive and when testing is conducted soon after training (Desmond and Fiez, 1998; Grafton et al., 2000; Van Mier, 2000). In contrast, increases and shifts in activity are generally found after extensive training, with a great deal of exposure to the stimuli and a period of consolidation after training from hours to days before brain imaging is conducted (Desmond and Fiez, 1998; Grafton et al., 2000; Van Mier, 2000).

In this study, neural activity underlying speech perception before and after 1 month of extensive training of the English /r-l/ contrast was investigated using fMRI. The English /b-v/ contrast (difficult for native Japanese speakers) and the /b-g/ contrast (easy) were also tested (but not trained). The reference control stimuli consisted of signal-correlated noise, SCN. SCN stimuli preserve the temporal envelope of the original word stimuli but do not contain the spectral information (see Methods for details of how SCN stimuli were constructed). In part, the /b-v/ and /b-g/ conditions serve as controls for changes in general processing strategies as well as for changes in activity related to better signal-to-noise ratio associated with experience in the fMRI environment. Additionally, the common activity before and after training for the /b-g/ contrast relative to the control SCN condition can be used to assess to some extent brain regions involved with processing of a contrast that is easily identified and discriminated. The English /b-g/ contrast is easy for native Japanese speakers because the contrast exists in Japanese and is easily assimilated into similar phonetic categories. Conversely, the English /r-l/ contrast is difficult for native Japanese speakers because both the English phonemes /r/ and /l/ are assimilated into the same Japanese phonetic category /r/ (the same pattern is true for /b-v/). See

Guenther and Bohland (2002) for a self-organizing neural network instantiating auditory representational space of English /r/ and /l/ phonetic categories and the Japanese /r/ phonetic category. Also see McCandliss et al. (2002) for a model of English /r-l/ acquisition by native Japanese speakers based on unsupervised Hebbian learning.

One possible hypothesis maintains that acquisition of a difficult L2 phonetic contrast involves reorganization of the same neural mechanisms as those used during processing of a native phonetic contrast. Studies suggest that processing of complex acoustic features of the human voice are carried out in a region of the brain located in the left and right superior temporal gyrus, STG, superior temporal sulcus, STS, and medial temporal gyrus, MTG, anterior to the primary auditory area, as well as a region in the left posterior STS (Belin et al., 2000; Binder et al., 1996, 2000; Demonet et al., 1992; Hickok and Poeppel, 2000; Scott et al., 2000; Wise et al., 2001; Zatorre et al., 1992). Other brain regions that are thought to be involved with aspects of phonological processing (especially when the task is difficult and involves processing segmental aspects of speech) as well as for speech production include the left planum temporale, PT (Scott et al., 2000), left supramarginal gyrus, SMG (Celsis et al., 1999; Damasio and Damasio, 1980; Hickok and Poeppel, 2000), left inferior frontal cortex, IFC, including Broca's area (Demonet et al., 1992; Hickok and Poeppel, 2000; Zatorre et al., 1992; Price et al., 1996), and left premotor cortex, PMC (Zatorre et al., 1992). Based on the hypothesis stated above, one might expect that as a result of learning a difficult nonnative phonetic contrast, neural processes involved with acoustic-phonetic processing will reorganize in a manner such that brain activity will resemble a phonetic contrast that is easily identified and discriminated. Accordingly, given the extensive training used, one would expect learning-induced plasticity reflected by increases in activity to occur focally in regions of the brain involved with processing of complex speech-related acoustic features. Additionally, one would expect a reduction in the spatial extent of activity as well as a decrease in activity in regions known to be involved with speech processing that are also linked to task difficulty, as training proceeds and performance improves.

Alternatively, given the known improvement in speech production that accompanies perceptual learning of the English /r-l/ contrast (Bradlow et al., 1997), it could be hypothesized that /r-l/ phonetic identification performance is facilitated by recruitment of neural mechanisms involved with auditory-articulatory (perceptual-motor) mappings related to verbal speech processing and learning as well as by reorganization of neural processes used for acoustic-phonetic aspects of speech perception. Facilitation could arise from establishment of auditory-articulatory mappings in which a context-dependent prediction of phonetic identity is made based on ongoing auditory input and internally generated phonetic production (akin to analysis by synthesis). According to this hypothesis, reorganization of perceptual

maps in brain regions involved primarily with complex speech-related acoustic features is induced by global neural networks composed of both cortical and subcortical brain regions concerned with the formation and processing of perceptual–motor mappings. One would expect learning-induced plasticity to be reflected by an enhancement (defined by both increases and shifts) in activity in brain regions involved with acoustic aspects of phonetic processing (see above) as well as in regions involved with perceptual–motor mappings related to verbal speech processing and learning. These regions include the cerebellum (Desmond and Fiez, 1998; Desmond et al., 1997; Ivry and Fiez, 2000; Silveri et al., 1998), basal ganglia (Muller et al., 2000; Trepanier et al., 1998), Insula (Kent and Tjaden, 1997; Paulesu et al., 1993), STG (Kent and Tjaden, 1997; Paulesu et al., 1993), SMG (Kent and Tjaden, 1997; Paulesu et al., 1993), supplementary motor area, SMA (Kent and Tjaden, 1997; Paulesu et al., 1993; Silveri et al., 1998), dorsolateral prefrontal cortex, DLPFC (Kent and Tjaden, 1997; Gaillard et al., 2000; Paulesu et al., 1993; Silveri et al., 1998; Smith and Jonides, 1999), PMC (Kent and Tjaden, 1997; Smith and Jonides, 1999), and IFC, including Broca's area and ventral inferior frontal cortex, VIFC (Desmond and Fiez, 1998; Desmond et al., 1997; Gaillard et al., 2000; Hickok and Poeppel, 2000; Kent and Tjaden, 1997; Paulesu et al., Silveri et al., 1998; Smith and Jonides, 1999).

Methods

Subjects

Nine right-handed monolingual Japanese speakers participated in this study. Subjects were between 21 and 30 years old. Two subjects were female and seven were male. All of the subjects received 6 years of English education in junior and senior high school, in which reading and writing were emphasized rather than conversational skills. None of the subjects had lived in an English-speaking community. Subjects were paid for their participation and gave written informed consent for experimental procedures, approved by the ATR Human Subject Review Committee.

Stimuli and procedure

The speech stimuli used for the fMRI experiment consisted of 24 minimal pairs of monosyllabic English words differing between /r/ and /l/ word initially. Analogous sets were also made for /b–v/ (e.g., *boat* and *vote*) and for /b–g/ (e.g., *bun* and *gun*). The speech stimuli were spoken by a male native English speaker and were recorded digitally in an anechoic chamber with a sampling rate of 22,050 Hz. SCN stimuli were constructed from each of the monosyllabic speech stimuli. Each SCN stimulus was created by first generating a random sequence of positive and negative ones containing the same number of samples as the monosyllabic

speech stimulus at hand and then multiplying this monosyllabic speech stimulus by the random sequence of positive and negative ones. The SCN stimuli have temporal envelopes similar to those of the original stimuli but do not contain the same spectral information. The SCN stimuli constructed from the /r–l/ monosyllabic words were used as control stimuli in the /r–l/ session. This same method of control was also used for the /b–v/ and /b–g/ sessions.

Subjects were trained to identify the English /r–l/ contrast by undergoing perceptual training with feedback (Bradlow et al., 1997). Stimuli used during training were 68 pairs of English words minimally contrasting in /r/ and /l/ in different word positions, spoken by five native American English speakers. On each trial, subjects saw one minimal pair on the computer screen and heard one of the two words presented through headphones. Subjects responded by clicking the word that they thought they heard. The program then gave immediate feedback on whether the response was correct; if incorrect, the trial was repeated with the same stimulus until the subject responded correctly. The format of the training program was that of a game in which a pleasant audio signal rewarded correct responses, whereas an unpleasant audio signal was presented for incorrect responses. The total number of correct responses was also recorded in units of 3 and depicted as coins on the computer monitor. The training lasted approximately 1 month and consisted of 45 sessions, each session containing 272 trials. In each session, the 68 minimal pairs spoken by one of the native speakers were presented twice each in a randomized order. No training was given for the /b–v/ or /b–g/ contrasts.

Before and after training, subjects underwent behavioral testing outside of the scanner, in which the same two-alternative forced-choice task was used. However, subjects did not receive feedback about their responses during these tests. Each test consisted of one session with about 160 English words containing /r/ and /l/, none of which appeared during the training phase.

Identification responses were also collected during the fMRI scanning before and after training for three phonetic contrasts, /r–l/, /b–v/, and /b–g/. Since these identification tests were conducted during scanning, the signal-to-noise ratio for these tests was much lower than that in the above mentioned tests before and after training. The same stimuli were used for tests before and after training during fMRI recording.

The fMRI procedure consisted of a block design in which 12 monosyllabic words were presented (approximately 85–90 dB SPL) at a rate of 1 every 2 (synchronized with fMRI scanning using Neurobehavioral System's Presentation software) via MR-compatible headphones (Hitachi Advanced Systems' ceramic transducer headphones; frequency range 30–40,000 Hz, approximately 20 dB SPL passive attenuation) followed by 12 control SCN stimuli. This pattern was repeated four times until all 48 monosyllabic words were presented. The speeded task helped to ensure that response times would not be different before and

after training. This is important in order to control for changes in brain activity after relative to before training that is a result of faster response time rather than learning-induced plasticity (Poldrack, 2000). Separate experimental sessions were conducted for the three contrasts /b–g/, /b–v/, and /r–l/. The task for the subject was to determine the initial phoneme of each word by button press in the word blocks and to randomly press one of the two buttons in the control blocks. Although this response paradigm attempts to control for brain activity related to the act of pressing a button, it is understood that random versus directed button press in the control versus experimental blocks is likely to involve somewhat different neural processes (Van Oostende et al., 1997). However, given that response times do not differ, comparison of after and before conditions will alleviate this potential confound. Functional brain imaging was conducted before and after approximately 1 month of perceptual identification training.

fMRI data collection and preprocessing

For structural and functional brain imaging, Shimadzu–Marconi's Magnex Eclipse 1.5T PD250 was used at the ATR Brain Activity Imaging Center. Functional T2*-weighted images were acquired using a gradient echo-planar imaging sequence (echo time, 55 ms; repetition time, 6000 ms; flip angle, 90°). Depending on head size, a total of 50–52 contiguous axial slices were acquired with a 3 × 3 × 3 mm voxel resolution covering the cortex and cerebellum. Anatomical T1-weighted MRI scans were also acquired for each individual with a 1 × 1 × 1 mm voxel resolution.

Images were preprocessed using programs within SPM99b. Differences in acquisition time between slices were accounted for, movement artifact was removed, and images were spatially normalized (2 × 2 × 2 mm) by matching the segmented gray matter to a template gray matter probability map and were smoothed using an 8-mm FWHM Gaussian kernel.

fMRI data analysis

Regional brain activity for the various conditions was assessed on a voxel-by-voxel basis using SPM99b (Wellcome Department of Cognitive Neurology, University College, London). A mixed-effect model was employed. At the first level (fixed effect within subjects), the before training conditions, the after training conditions, the after-minus-before training conditions (taking into account of the SCN control condition implicitly in the design), and the before-minus-after training conditions for each phonetic contrast were assessed separately for each subject. The data were modeled using a boxcar function convolved with the hemodynamic response function. In addition, global normalization and grand mean scaling as well as high-pass filtering (minimum cutoff period of 96 s) were carried out. At the second level (random effect between subjects) the contrast

image of the parameter estimates of the first level analysis for each subject was used as input for a SPM model employing a basic one-sample *t* test. The problem of multiple comparisons was controlled for by adjusting the *T* threshold using the false discovery rate FDR procedure (Genovese et al., 2002). The FDR procedure provides an unbiased manner in which data sets of differing signal intensity can be compared based on the same corrected *P* value. In order to constrain the number of multiple comparisons, voxels falling in regions of the CSF were exclusively masked and only voxels within the gray matter were included in the analysis. The CSF and gray matter masks were defined by the Montreal Neurological Institute templates based on the average of 305 individuals, threshold at 0.2. The uncorrected threshold of *T* = 4.5, *P* < 0.001, was used under conditions in which the *T* value of the FDR-corrected threshold of *P* < 0.05 is greater than 4.5. The spatial extent threshold used during all analyses was 25 voxels.

Results

Behavioral performance

Subjects' ability to identify English /r/ and /l/ improved significantly, from a mean of 62.5% in the before training test to 80.4% in the after training test outside of the MRI scanner [*F*(1,8) = 78.3; *P* < 0.001]. This suggests that the training was indeed effective in enhancing Japanese listeners' ability to perceive the /r–l/ contrast. Subjects' performance collected during fMRI scanning on the /b–g/ contrast was, not surprisingly, very high (above 97% on average) both before and after training. For the /r–l/ contrast, performance again improved significantly, from 62.7% before training to 73.2% after training [*F*(1,6) = 10.5; *P* < .05]. Finally, the /b–v/ contrast also showed a significant improvement, from 60.1 to 68.4% [*F*(1,7) = 8.5; *P* < .05]. As a result of machine failures, identification performance during fMRI scanning for two subjects under the /r–l/ and /b–g/ conditions and one subject under the /b–v/ condition were lost. The difference in /r–l/ identification performance in and outside the scanner is likely a result of the speeded response task, anxiety, as well as the presence of scanner noise. Button press response times did not significantly differ for any of the conditions relative to the /r–l/ conditions (this is true both for experimental and for control SCN trials).

Brain imaging

Brain regions involved with phonetic processing of a contrast that is easily identified were determined by finding activated voxels common to the /b–g/ contrast both before and after training (for the /r–l/ contrast) relative to the SCN condition (/b–g/ before: *T* = 4.5, *P* < 0.001, uncorrected, *P* < 0.058, FDR corrected, *df* = 8; /b–g/ after: *T* = 4.29, *P* <

Table 1
Brain regions involved with phonetic processing

	Both /b–g/ (Fig. 1)		Before /r–l/ (Fig. 2a)		After /r–l/ FDR (Fig. 2b)		After /r–l/ 0.001 (Fig. 2c)		Enhancement /r–l/ (Fig. 3a)		Enhancement /b–v/ (Fig. 3b)		Enhancement /b–g/ (Fig. 3c)		Enhancement /b–v/ and /b–g/ (Fig. 3d)	
	Left	Right	Left	Right	Left	Right	Left	Right	Left	Right	Left	Right	Left	Right	Left	Right
Broca's (BA44–45)	56		255		814	648	410	233	157	400		3		218		13
PMC (BA6–8)	43		102		1114	1265	413	332	292	703		56		228		29
SMA			6		168	214	38	29	66	30						
BA8			8		56	49	18	10	10	24			1			
Precentral gyrus			3		84	8	7	1	10	1						
DLPFC (BA9–46)					252	603	93	163	33	420					13	
VIFC (BA47–11)			168	73	949	643	466	202	262	265			17	2		
Frontal polar cortex					56	115			1	40						
Medial frontal cortex			4	2	103	156	18	19	15	82						
ITG			3		163	23	46	6	41	1						
MTG	128	65	157	107	1052	917	589	553	426	236		17			2	
STG/S	68	27	122	43	766	572	428	185	541	181		18				
A1–A2 (BA41–42)		2		38	564	415	226	123	320	144						
Planum tempore		12	21	12	526	319	187	136	184	172		28			34	
Fusiform gyrus			63		226	18	68		56							
Parahippocampal gyrus					338	187	28	27	143	54						
Amygdala					72	58		3	45	26						
SMG (BA40)			1		987	429	334	50	415	324		1			187	
AG (BA39)			3		229	130	94	41	49	27					1	
BA43					58	1	1		42	1						
Inferior parietal lobule					86	74	38	13	56	30					13	
Superior parietal lobule				322	75	84	23	71	26				9			
Posterior central gyrus					76	4			37	4						
BA19			19	18	288	197	94	45	22	73					1	
Cuneus			82	4	489	254	170	70	59	82					36	
Insula	8		136	34	1394	1216	487	315	475	747		25	49	88		42
Clastrum					18	479	359	70	58	198	241	3	1	13		
Cingulate				2	722	861	117	162	192	219			60	43		
Hippocampus					91		2		64							
Globus pallidus					277	519	91	295	103	426					57	
Caudate				18	854	767	325	75	513	446					29	
Putamen					116	450	10	30	37	115					10	
Thalamus			43		577	490	142	94	185	293					42	
Cerebellum		142	88	558	5430	5613	1993	1943	3261	2987	404	326	449	68	315	57
Brainstem			12		1097	1347	256	234	279	414	26					
Red nucleus					6	45		10		27						
Substantia nigra			1		39	21	22	1	21	16						

Note. PMC, premotor cortex; SMA, supplementary motor area; DLPFC, dorsolateral prefrontal cortex; VIFC, ventral inferior frontal cortex; ITG, inferior temporal gyrus; MTG, medial temporal gyrus; STG/S, superior temporal gyrus/sulcus; A1–A2, primary and secondary auditory cortex; SMG, supramarginal gyrus; AG, angular gyrus. The basal ganglia is composed of the globus pallidus, caudate, and putamen. Brain regions were determined by nearest neighbor approximation based on coordinates given in the Talairach Daemon Database, Version 1.1, Research Imaging Center, University of Texas Health Science Center at San Antonio. Bold numbers denote areas referred to in the text. It should be noted that the large clusters of activity seen in the figures overlap with several brain regions defined in this table.

0.05, FDR corrected, $df = 8$) (Fig. 1). Brain regions primarily activated include the STG/S, and MTG bilaterally, right crus VIII of the cerebellum, left Broca's area, and left PMC (Table 1, Fig. 1).

Brain activity for /r–l/ before training ($T = 4.5$, $P < 0.001$, uncorrected, $P < 0.052$, FDR corrected, $df = 8$) and after training relative to the SCN condition ($T = 2.64$, $P < 0.05$, FDR corrected, $df = 8$) is shown in Figs. 2a–c. Brain regions primarily activated before training include the STG/S, MTG, VIFC, and insula bilaterally, right and to a lesser extent left crus VII–VIII of the cerebellum, left Bro-

ca's area, and left PMC (Table 1, Fig. 2a). After training there is a considerable expansion of the spatial extent of activity (increases) in all regions showing activity before training as well as additional activity (shifts) bilaterally in the SMG, IP, SP, claustrum, parahippocampal gyrus, cingulate, basal ganglia (globus pallidus, caudate, putamen), brainstem, substantia nigra SN, crus VI of the cerebellum, thalamus, primary and secondary auditory cortex A1–A2, PT, DLPFC, and SMA, as well as right Broca's area and right PMC (Table 1, Fig. 2b). This pattern of results is still present even when /r–l/ after training is analyzed with the

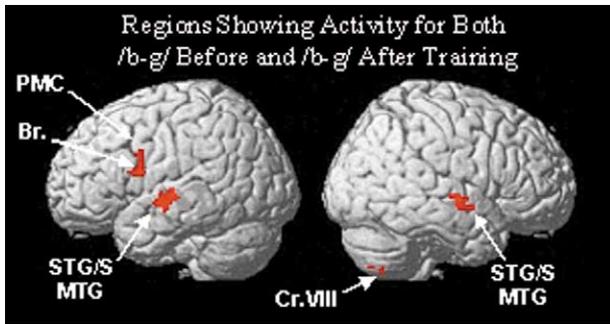


Fig. 1. SPM showing brain activity greater for both /b-g/ before and after training relative to the control SCN condition. Anatomical location was determined by reference to Talairach coordinates and atlas (Schmahmann et al., 2000; Talairach and Tournoux, 1988) and are noted on the figure. Significant differential activity rendered on the surface of the brain (red) present in the superior temporal gyrus, STG, superior temporal sulcus, STS, medial temporal gyrus, MTG, premotor cortex, PMC, Broca's area, Br., and the cerebellum crus Cr. VIII.

same T threshold ($T = 4.5$, $P < 0.001$) as that used for /r-l/ before training (Table 1, Fig. 2c).

Enhancement (both increases and shifts) in neural activity thought to reflect neural plasticity was determined by a statistical subtraction between the after training condition relative to the before training condition (Figs. 3a–e). Voxels included in the analysis were restricted to those that were significantly active for the /r-l/ after training condition relative to SCN (the results depicted in Fig. 2b). Brain regions showing enhancement in activity for the /r-l/ contrast ($T = 2.37$, $P < 0.05$, FDR corrected) include primarily the STG/S, MTG, A1-A2, PT, SMG, insula, claustrum, cingulate, basal ganglia (globus pallidus, caudate, putamen), thalamus, crus I, V, VI, VII, and VIII of the cerebellum, brain stem, SN, DLPFC, VIFC, SMA, PMC, and Broca's area bilaterally (Table 1, Figs. 3a–e). Greater enhancement in activity in the right hemisphere over that of the left was found in Broca's area and the PMC, DLPFC, insula, and putamen (Table 1). To ensure that enhancement in activity is not the result of changes in general processing strategies, task-related experience, or session-related effects, statistical subtraction analysis, after relative to before /r-l/ training, was conducted for both /b-v/ and /b-g/ contrasts (it should be noted that no explicit training occurred for the /b-v/ or /b-g/ contrasts). No enhancement in activity was found to be significant using a criterion of $T = 4.5$, $P < 0.001$, or $P < 0.05$, FDR corrected, restricting the analysis over the same voxels as those restricted for the analysis of the /r-l/ contrast. A more lenient analysis ($T = 1.86$, $P < 0.05$, uncorrected) was then conducted for both /b-v/ and /b-g/ contrasts restricted to voxels found to be active for the /r-l/ after-minus-before analysis. Brain regions showing enhancement in activity for the /b-v/ contrast include primarily crus VI, VII, and VIII of the cerebellum bilaterally as well as the right STG/S, MTG, PT, insula, and PMC (Table 1, Figs. 3b and e). Brain regions showing enhancement in activity for the /b-g/ contrast include primarily crus VI, VII,

and VIII of the cerebellum and the insula and cingulate gyrus bilaterally as well as the right PT, SMG, basal ganglia (globus pallidus, caudate), PMC, Broca's area, and left VIFC (Table 1, Figs. 3c and e). Brain regions showing enhancement in activity for both the /b-v/ and /b-g/ contrasts include primarily the right and to a lesser extent left crus VI, VII, and VIII of the cerebellum, as well as the right insula, PMC, and Broca's area (Table 1, Figs. 3d and e). Statistical analysis revealed no significant decrements in activity for any of the phonetic contrasts (criteria: $P < 0.001$, uncorrected, or $P < 0.05$, FDR corrected) such that before training was greater than after training.

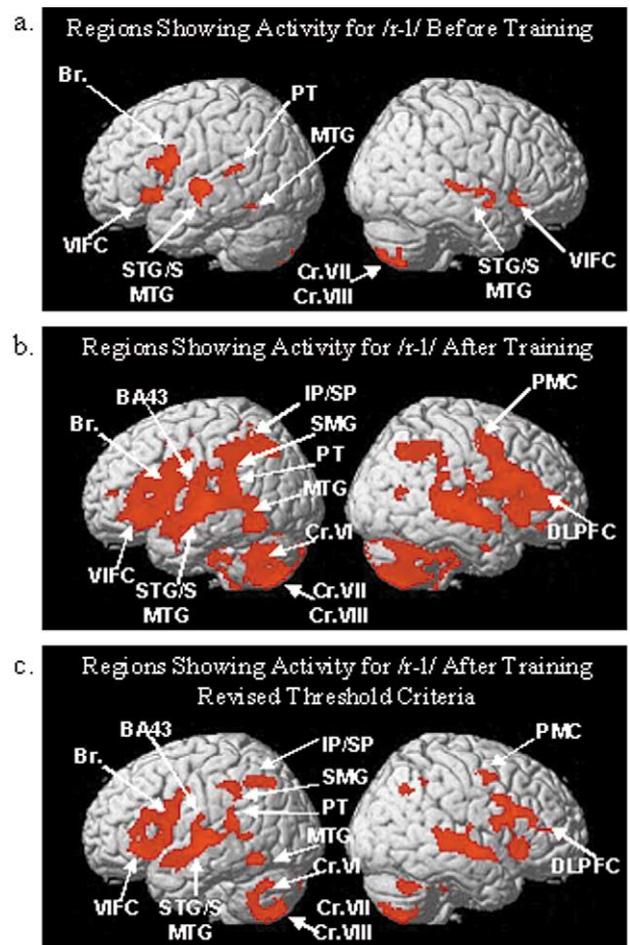


Fig. 2. SPM showing brain activity for (a) /r-l/ before training relative to the control SCN condition (b), /r-l/ after training relative to SCN condition using the FDR-corrected threshold, and (c) /r-l/ after training relative to SCN using the same threshold as that used for the before training /r-l/ contrast ($P < 0.001$). Regions activated were determined by reference to Talairach coordinates and atlas (Schmahmann et al., 2000; Talairach and Tournoux, 1988) and are noted on the figure. For some regions showing bilateral activity only one hemisphere is labeled. The abbreviations used include STG/S, superior temporal gyrus/sulcus; MTG, medial temporal gyrus; PT, planum temporale; SMG, supramarginal gyrus; IP, inferior parietal lobule; SP, superior parietal lobule; BA43, Brodmann area 43; Br., Broca's area; VIFC, ventral inferior frontal cortex; PMC, premotor cortex; DLPFC, dorsolateral prefrontal cortex; cerebellum crus Cr. VI, VII, and VIII.

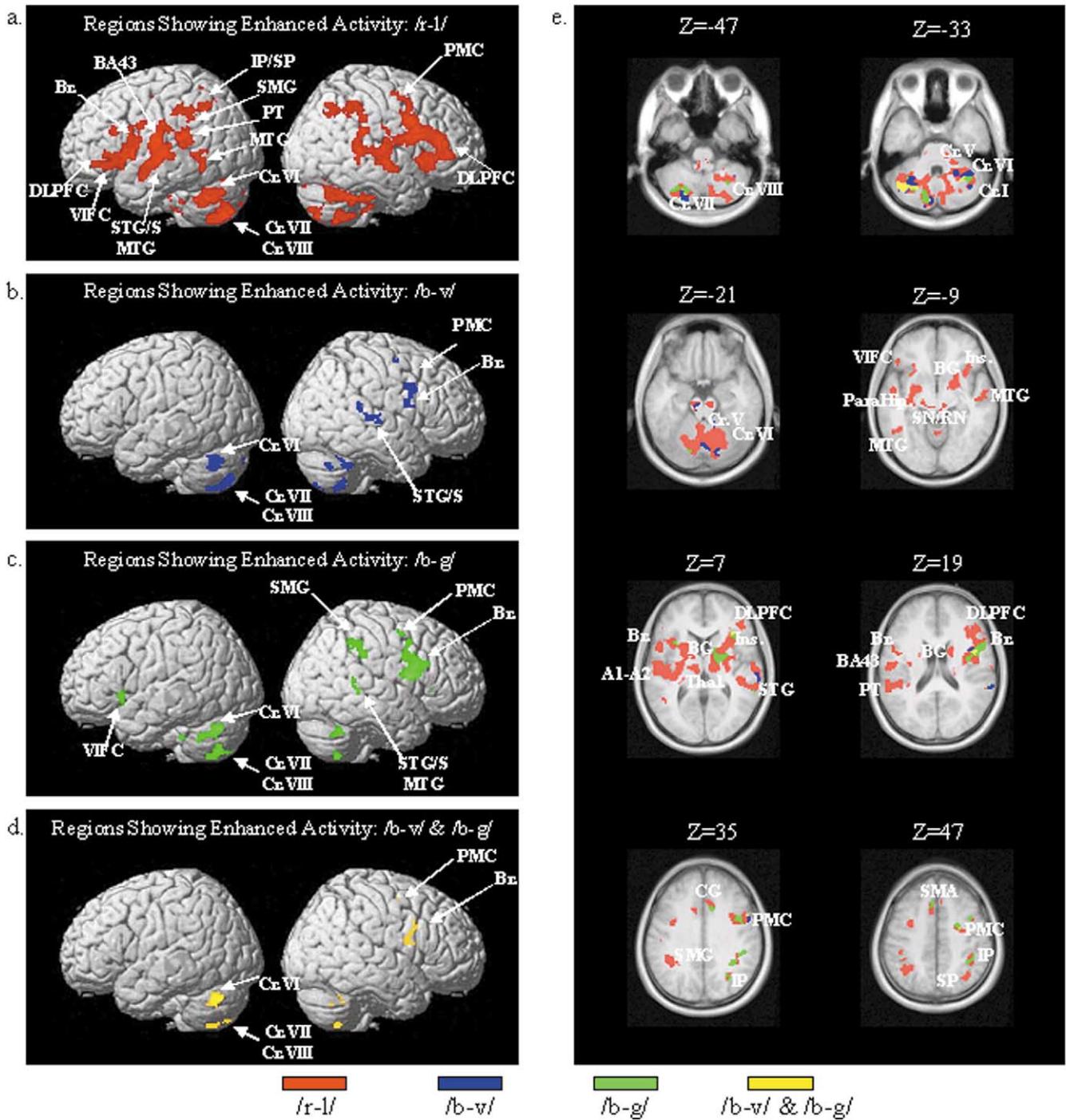


Fig. 3. SPM of enhancement in activity after relative to before training for the /r-l/ contrast (red), /b-v/ contrast (blue), /b-g/ contrast (green), and /b-v/ and /b-g/ contrasts (yellow). Analysis of the /b-v/ and /b-g/ contrasts was restricted to regions that show enhancement for the /r-l/ contrast. It should be noted that the /r-l/ contrast is significant at $T = 2.37$, $P < 0.05$, FDR corrected, and the /b-v/ and /b-g/ contrasts are significant at $T = 1.86$, $P < 0.05$, uncorrected. Regions activated were determined by reference to Talairach coordinates and atlas (Schmahmann et al., 2000; Talairach and Tournoux, 1988) and are noted on the figure. For some regions showing bilateral activity only one hemisphere is labeled. (a-d) Significant enhancement in activity rendered on the surface of the brain. (e) Eight transverse slices showing significant enhancement in activity. The abbreviations used include STG/S, superior temporal gyrus/sulcus; MTG, medial temporal gyrus; A1-A2, primary and secondary auditory cortex; PT, planum temporale; SMG, supramarginal gyrus; IP, inferior parietal; SP, superior parietal; BA43, Brodmann area 43; Br., Broca's area; VIFC, ventral inferior frontal cortex; PMC, premotor cortex; DLPFC, dorsolateral prefrontal cortex; cerebellum crus Cr. I, V, VI, VII, and VIII; SN, substantia nigra; BG, basal ganglia; ParaHip., parahippocampal gyrus; Thal., thalamus; CG, cingulate gyrus; SMA, supplementary motor area.

Discussion

The results are consistent with the hypothesis that acquisition of auditory–articulatory (perceptual–motor) mappings are used to facilitate /r–l/ phonetic identification and are contrary to the prediction that brain activity for the /r–l/ contrast will come to resemble that present during processing of an easy phonetic contrast (/b–g/) as training proceeds. It can be clearly seen in Fig. 2 that /r–l/ after training does not come to better resemble processing of an easy phonetic contrast, /b–g/ (Fig. 1), relative to /r–l/ before training (Table 1, Figs. 1 and 2a–c). Brain activity for the easy phonetic contrast (/b–g/) both before and after training as well as brain activity for the /r–l/ contrast before training relative to SCN is relatively circumscribed in regions known to be involved with phonetic processing, including the STG/S, MTG, Broca's area, and PMC (with the /r–l/ contrast showing slightly more extensive activity as well as activity in regions not active for the /b–g/ contrast) in relation to the /r–l/ contrast after training relative to SCN, which shows considerably more activity in brain regions thought to be involved with aspects of acquisition and processing of speech (Table 1, Figs. 1 and 2a–c). This pattern of results can still be seen in Fig. 2 even when the /r–l/ after training analysis is conducted using the same T threshold as that used for the /r–l/ before training analysis instead of the more lenient FDR-corrected threshold (Table 1, Figs. 2a–c). It is possible that the results seen in this study represent an intermediate stage of learning and that after native-like proficiency is attained that the neural processes involved may become more circumscribed, as is seen for the /b–g/ contrast. It should also be pointed out that although activity in the PMC most likely reflects speech-related processes, differences resulting from random versus directed button press under the control SCN condition versus the experimental conditions (Van Oostende et al., 1997) cannot be ruled out.

The greater enhancement in activity for the /r–l/ contrast after training relative to before training was confirmed using an after-minus-before analysis. Considerable enhancement was found to be present denoted by expansion in the spatial extent and/or magnitude of activity (increases) as well as by the emergence of activity present after but not before training (shifts). Increases in activity occurred in regions involved with complex speech-related acoustic features (STG/S, MTG bilaterally), and phonetic processing (left Broca's area and PMC) (Table 1, Figs. 2–3). The recruitment of new activity (shifts) occurred in regions involved with phonetic processing (including right PMC, right Broca's area, and PT bilaterally) as well as regions involved with acquisition and processing of speech related perceptual-motor mappings (cerebellum, SN, basal ganglia, SMG, insula, VIFC, DLPFC, SMA, PMC, and Broca's area bilaterally), (Table 1, Figs. 2–3). Brain regions that were also active for the /b–v/ and/or the /b–g/ contrasts included predominantly the right PMC, right Broca's area, right PT,

right insula, and the cerebellum bilaterally (Table 1, Fig. 3, blue, green, yellow). Changes in brain activity in these regions may reflect differences in general processing strategies, task-related experience, and/or session-related effects after relative to before training. Improved identification performance for the /b–v/ contrast may have resulted from the use of general processing strategies or task familiarity acquired during perceptual learning of the /r–l/ contrast that are reflected in this enhancement of brain activity (it should be noted that no explicit training was conducted for the /b–v/ or /b–g/ contrasts). The use of general processing strategies or task familiarity acquired as a result of training on the /r–l/ contrast may also account for enhancement in brain activity for the /b–g/ contrast even though behavioral performance was at ceiling and no explicit training was conducted for either the /b–v/ or /b–g/ contrast. It should be noted that there is a great deal of activity after training relative to before training for the /r–l/ contrast that is not common to the /b–v/ and /b–g/ contrasts (Table 1, Fig. 3, red), even though a much more conservative threshold was employed (/r–l/, $T = 2.37$, $P < 0.05$, FDR corrected; /b–v/ and /b–g/, $T = 1.86$, $P < 0.05$, uncorrected). This activity probably does not reflect the use of general processing strategies nor task familiarity, but rather changes in neural processes related to representation and acquisition of the /r–l/ contrast.

Enhancement of activity found in primary and secondary auditory cortex A1–A2 (Table 1, Figs. 3a and c) suggests that the processing carried out in structures comparatively early in the auditory pathway can be modified as a result of training. The enhancement in activity may reflect induced changes in the processing of these structures or/and may reflect additional top-down influence from other cortical processes.

The enhancement in Broca's area and the SMG for the /r–l/ contrast after relative to before training (Table 1, Fig. 3) is also interesting given implications that these regions are concerned with establishment of an auditory–articulatory mapping involved with speech production acquisition as well as phonological working memory (Hickok and Poeppel, 2000). Additional activity present after training but not present before training (shifts) in the right hemisphere superior temporal and speech motor regions (Table 1, Figs. 2 and 3) are in line with verbal processing found in children showing greater bilateral activity than adults (Gaillard et al., 2000), suggesting that similar right hemisphere neural processes used during acquisition of one's native language may later be used for L2 acquisition. Involvement of right hemisphere activity for the /r–l/ contrast after training in regions not active for /b–g/ (Table 1, Figs. 1 and 2) is consistent with studies showing that the second language is processed to some degree in different brain regions from the native language (Chee et al., 2001; Dehaene et al., 1997; Kim et al., 1997; Nakai et al., 1999) and is contrary to studies showing no difference in brain regions involved with native and L2 processing (Chee et al., 1999; Illes et al., 1999;

Klein et al., 1995). This is not to say that L2 is distinctly represented in these regions, but rather that activity in these regions reflects the recruitment of additional neural processes to carry out phonetic perception. In this study it is proposed that shifts in activity reflect acquisition of perceptual–motor mappings used to facilitate speech perception.

The pattern of results found in this study is consistent with some aspects of L2 processing found in other studies in which passive listening comprehension tasks were used (Nakai et al., 1999; Perani et al., 1998). In the experiment conducted by Perani et al. (1998) greater brain activity was found in superior and middle temporal areas during passive listening for high proficiency, late acquisition bilinguals than for low proficiency, late acquisition bilinguals. This is consistent with the greater enhancement in brain activity found in this study for the /r–l/ contrast after relative to before training (Table 1, Fig. 3). However, there are several discrepancies between this study and that conducted by Perani et al. (1998). Speech motor areas showed a large degree of enhancement in this study but not in the Perani et al. (1998) study. Lack of speech motor activity in the Perani et al. (1998) study may have resulted from the passive listening task used or perhaps from differences in the analysis techniques for PET that use group averages (used in their study) compared to that of fMRI that are able to extract differences at a single subject level. Another discrepancy was the much greater degree of activity in our study for the /r–l/ contrast after training (Table 1, Figs. 2b and c) relative to the circumscribed activity for the /b–g/ contrast (Table 1, Fig. 1), compared to the lack of a difference in brain activity between native and L2 processing in the Perani et al. (1998) study and greater activity found for the native than for the second language in the Perani et al. (1996) study. Given that there was near perfect performance for the /b–g/ contrast and that the English /b–g/ contrast is easily assimilated to native Japanese categories of the same phonemes it could be expected, based on the results of the Perani et al. (1996) study, that the /b–g/ contrast should show greater activity than the /r–l/ contrast. This line of reasoning follows from the hypothesis that the neural mechanisms used for perception of a difficult L2 phonetic contrast are the same as those for perception of a native phonetic contrast (or a L2 contrast easily assimilated to a native phonetic contrast). However, if additional neural mechanisms are recruited, as is conjectured to be the case in this study, the greater degree of activity for the /r–l/ contrast after training compared to the /b–g/ contrast is to be expected. Consistent with the results of this study, suggesting that L2 phonetic processing is mediated by acquisition and processing of speech-related perceptual–motor mappings, a passive listening comprehension fMRI study conducted by Nakai et al. (1999) also reports greater activity in Broca's area and the SMA, SMG, and AG for the nonnative comprehensive language than for native and nonnative noncomprehensive languages.

One possible confound may result from more efficient lexical access after learning because of the better ability to

disambiguate the semantic content of the word pair that occurs with better /r–l/ identification performance. However, the instructions given to the subjects to identify the initial phoneme of the word presented, the speeded task for the subjects, and the difference in activity from the /b–g/ condition, for which lexical access is even more efficient, makes it unlikely that this potential confound is responsible for the enhancement in activity after training found for the /r–l/ contrast. In addition, although it is possible that the results are due to greater attention being paid to the /r–l/ contrast after training relative to the /b–v/ and /b–g/ contrasts, it is unlikely given that behavioral performance for the /b–v/ contrast improved for the after training session as well.

It was proposed above that /r–l/ phonetic identification performance is facilitated in part by recruitment of cortical and subcortical neural mechanisms involved with the establishment and representation of auditory–articulatory (perceptual–motor) mappings related to verbal speech processing and learning. These areas were shown to be active in this study (Table 1, Fig. 3). The cortex, cerebellum, and the basal ganglia may all be differentially involved with establishment of auditory–articulatory mappings that allow for context-dependent prediction of phonetic identity to be made based on ongoing auditory input and internally generated phonetic production. In the context of this study, distinct functions and learning algorithms underlying acquisition of perceptual–motor mappings related to verbal speech processing are conjectured to be carried out in these brain regions: the cortex is specialized for unsupervised learning (learning based on statistical regularities in the input space); the cerebellum is specialized for supervised learning (learning that uses explicit error-feedback as a teaching signal); and the basal ganglia is specialized for reinforcement learning (learning that is based on reward) (Doya, 1999).

The cortex contains representational maps that are in part formed in an unsupervised manner based solely on statistical regularities present in sensory and motor stimulation/activation. This correlation-based learning is referred to as Hebbian (Hebb, 1949). There are a considerable number of studies demonstrating that cortical map formation during development is dependent on the statistical pattern inherent in input stimulation (Miller et al., 1999; Sur et al., 1990; Wiesel, 1982). Cortical auditory speech maps are likely formed during development based on statistical regularities inherent in the native language environment (Guenther and Bohland, 2002; Kuhl, 2000). Cortical maps between articulatory and auditory representations of speech could also form during development based on properties of statistical correspondence. There have been several studies demonstrating that cortical map representation can be reorganized in adults based on the pattern of input stimulation (Recanzone, 2000). Although mere exposure to the pattern of input stimulation is sufficient to reorganize adult cortical map representations under some conditions, others require train-

ing using reward feedback in which attention to the stimulus is necessary (Recanzone, 2000). It is possible that reward feedback training and/or attention involves modulation and enhancement of Hebbian learning processes such that value is associated with specific properties in input stimulation and not others. One potential source of this modulatory effect could be by reinforcement-based learning thought to involve the basal ganglia (see below). In the context of this study, the cortex establishes speech related auditory and articulatory representational maps. A considerable degree of neural plasticity in cortical regions involved with both speech perception and production was revealed in this study (see above for details as well as Table 1, Figs. 2 and 3).

The cerebellum, known to be involved with motor as well as cognitive processes (Desmond and Fiez, 1998; Desmond et al., 1997; Fabbro, 2000; Ivry and Fiez, 2000; Liu et al., 2000; Middleton and Strick, 2000; Silveri et al., 1998), has been in the acquisition of internal models of perceptual–motor mappings (Imamizu et al., 2000). Internal models in the cerebellum are thought to provide for both forward and inverse mappings between sensory–motor representations (Kawato, 1999). In the context of this paper, the cerebellum uses supervised learning to acquire internal forward and inverse models that reciprocally map between auditory speech representations and articulatory speech representations. The error signal for supervised learning could be generated via feedback by internally generated speech during training. Internal models function as a coordinate transform between auditory space and articulatory space that can be used to constrain phoneme selection to facilitate perception. In support of this, we found learning-induced plasticity in crus VI of the cerebellum in the region known to be involved with lip and tongue motor representation as well as speech production (Grodz et al., 2001; Wildgruber et al., 2001) (Fig. 3). In addition, regions of the cerebellum that show learning-induced plasticity in this study (Fig. 3), crus VI, VII, and VIII, are implicated in verbal and cognitive processing (Desmond and Fiez, 1998; Desmond et al., 1997).

The basal ganglia, SN, and thalamus together with cortical regions compose a dopaminergic neural circuit involved with sensory–motor and cognitive reinforcement-based learning (Doya, 1999). All of these sites show enhancement in activity for the /r–l/ contrast after relative to before training (Table 1, Fig. 3). It is interesting to point out that interruption of this circuit by lesioning of the basal ganglia prevents learning-induced plasticity of birdsong in adult zebra finches (Brainard and Doupe, 2000). Similarly, damage to this circuit in humans is associated with deficits in verbal learning and memory (Muller et al., 2000; Trepanier et al., 1998). Recent brain imaging studies have shown learning-related changes in activity on both motor tasks (Grafton, Hazeltine, and Ivry, 2002) and nonmotor tasks (Poldrack et al., 2001). With reference to this study, it is conjectured that the basal ganglia together with the dopaminergic neural circuit use reinforcement-based learning

to modify and select phonetic representations based on context-sensitive prediction and subsequent evaluation of ongoing cortical auditory and articulatory representations.

The results of this experiment and hypotheses concerned with various types of learning algorithms being involved with reorganization and recruitment of neural processes responsible for acquisition of a difficult phonetic contrast are relevant to findings by McCandliss et al. (2002) concerned with the success and failure of teaching native Japanese speakers the English /r–l/ phonetic contrast. The McCandliss et al. (2002) study found improved /r–l/ perception using stimuli with exaggerated acoustic features relevant to distinguishing /r/ from /l/ that were adaptively adjusted gradually during training without feedback, but not for stimuli that were nonexaggerated and unadjusted during training without feedback. These results are consistent with unsupervised Hebbian-based learning that may serve to reorganize acoustic–phonetic cortical representations. Under conditions in which error-feedback was used during training, for both adaptive and nonadaptive stimuli, it resulted in improved perception that generalized to novel stimuli to a much greater extent than the adaptive no-feedback training. These results are difficult to explain from a purely unsupervised Hebbian-based approach in which one would expect greater learning to occur in the adaptive condition over the non-adaptive condition regardless of feedback. One potential explanation for these results is the recruitment of neural processes that utilize reinforcement and supervised based learning (as described above) to establish perceptual–motor mappings related to verbal speech processing. Reinforcement-based learning involving the basal ganglia may be particularly robust at adaptation to contextual information necessary to generalize to perception of novel stimuli.

This is the first brain imaging study to explore localized changes in brain activity associated with learning of a L2 phonetic contrast. Speech perception of a difficult L2 phonetic contrast appears to be facilitated by learning-induced plasticity bilaterally in cortical as well as subcortical regions that are potentially involved with formation of forward and inverse perceptual–motor mappings between auditory representations and articulatory representations of speech. Facilitation of phonetic identification by use of auditory–articulatory mappings is consistent with the idea that perception is made in reference to potential action, either internally generated, as in this case, or externally generated. Although the use of auditory–articulatory mappings is compatible with aspects of the revised motor theory of speech perception (Lieberman and Mattingly, 1985), it should be pointed out that these mappings do not necessarily exist as predefined gestures. As put forward in this study, they could potentially be learned based on context-dependent statistical regularities inherent in stimulation that afford adaptive value. Given the results of this study it follows that there may be some benefit in using auditory–articulatory based training for acquisition of difficult L2 phonetic contrasts.

Acknowledgments

This research was conducted as part of “Research on Human Communication,” with funding from the Telecommunications Advancement Organization of Japan. We thank Mitsuo Kawato, Kenji Doya, Nicolas Schweighofer, and Kevin Munhall for valuable comments on drafts of this paper. In addition we thank the fMRI technicians Yasuhiro Shimada and Ichiro Fujimoto.

References

- Akahane-Yamada, R., 1996. Learning non-native speech contrasts: what laboratory training studies tell us. *Proceedings of the Acoustical Society of America and Acoustical Society of Japan Third Joint Meeting*, pp. 953–958.
- Belin, P., Zatorre, R.J., Lafaille, P., Ahad, P., Pike, B., 2000. Voice-selective areas in human auditory cortex. *Nature* 403, 309–312.
- Binder, J.R., Frost, J.A., Hammeke, T.A., Bellgowan, P.S.F., Springer, J.A., Kaufman, J.N., Possing, E.T., 2000. Human temporal lobe activation by speech and nonspeech sounds. *Cereb. Cortex* 10, 512–520.
- Binder, J.R., Frost, J.A., Hammeke, T.A., Rao, S.M., Cox, R.W., 1996. Function of the left planum temporale in auditory and linguistic processing. *Brain* 119, 1239–1247.
- Bradlow, A., Pisoni, D., Akahane-Yamada, R., Tohkura, Y., 1997. Training Japanese listeners to identify English /r/ an /l/: IV. Some effects of perceptual learning on speech production. *J. Acoust. Soc. Am.* 101, 2299–2310.
- Brainard, M.S., Doupe, A.J., 2000. Interruption of a basal ganglia-forebrain circuit prevents plasticity of learned vocalizations. *Nature* 404, 762–766.
- Celsis, P., Celsis, Boulanouar, K., Doyon, B., Ranjeva, J. P., Berry, I., Nespoulous, J. L., Chollet, F., 1999. Differential fMRI responses in the left posterior superior temporal gyrus and left supramarginal gyrus to habituation and change detection in syllables and tones. *NeuroImage* 9, 135–144.
- Chee, M.W., Hon, N., Lee, H.L., Soon, C.S., 2001. Relative language proficiency modulates BOLD signal change when bilinguals perform semantic judgments. *NeuroImage* 13, 1155–1163.
- Chee, M.W., Tan, E.W., Thiel, T., 1999. Mandarin and English single word processing studied with functional magnetic resonance imaging. *J. Neurosci.* 19, 3050–3056.
- Chollet, F., 2000. Plasticity of the adult human brain, in: Toga, A., Mazziotta, J. (Eds.), *Brain Mapping: The Systems*, Academic Press, San Diego, pp. 621–638.
- Damasio, H., Damasio, A.R., 1980. The anatomical basis of conduction aphasia. *Brain* 103, 337–350.
- Dehaene, S., Dupoux, E., Mehler, J., Cohen, L., Paulesu, E., Perani, D., Perani, D., van de Moortele, P.F., Lehericy, S., Le Bihan, D., 1997. Anatomical variability in the cortical representation of first and second languages. *Neuroreport* 8, 3809–3815.
- Démonet, J.-F., Chollet, F., Ramsay, S., Cardebat, D., Nespoulous, J.L., Wise, R., Rascol, A., Frackowiak, R., 1992. The anatomy of phonological and semantic processing in normal subjects. *Brain* 115, 1753–1768.
- Desmond, J.E., Fiez, J.A., 1998. Neuroimaging studies of the cerebellum: language, learning and memory. *Trends Cogn. Sci.* 9, 355–362.
- Desmond, J.E., Gabrieli, J.D.E., Wagner, A.D., Ginier, B.L., Glover, G.H., 1997. Lobular patterns of cerebellar activation in verbal working-memory and finger-tapping tasks as revealed by functional MRI. *J. Neurosci.* 15, 967–985.
- Doya, K., 1999. What are the computations of the cerebellum, the basal ganglia and the cerebral cortex? *Neural Networks* 12, 961–974.
- Fabbro, F., 2000. Introduction to language and cerebellum. *J. Neurolinguistics* 13, 83–94.
- Gaillard, W.D., Hertz-Pannier, L., Mott, S.H., Barnett, A.S., LeBihan, D., Theodore, W.H., 2000. Functional anatomy of cognitive development: fMRI of verbal fluency in children and adults. *Neurology* 54, 180–185.
- Genovese, C.R., Lazar, N.A., Nichols, T., 2002. Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *NeuroImage* 15, 870–878.
- Grafton, S.T., Hari, R., Salenius, S., 2000. The human motor system, in: Toga, A., Mazziotta, J. (Eds.), *Brain Mapping The Systems*, Academic Press, San Diego, pp. 331–363.
- Grafton, S.T., Hazeltine, E., Ivry, R.B., 2002. Motor sequence learning with the nondominant left hand: a PET functional imaging study. *Exp. Brain Res.* 146, 369–378.
- Grodd, W., Hulsmann, E., Lotze, M., Wildgruber, D., Erb, M., 2001. Sensorimotor mapping of the human cerebellum: fMRI evidence of somatotopic organization. *Hum. Brain Mapp.* 13, 55–73.
- Guenther, F.H., Bohland, J.W., 2002. Learning sound categories: a neural model and supporting experiments. *Acoust. Sci. Technol.* 23, 213–220.
- Hebb, D.O., 1949. *The Organization of Behavior*. Wiley, New York.
- Hickok, G., Poeppel, D., 2000. Towards a functional neuroanatomy of speech perception. *Trends Cogn. Sci.* 4, 131–138.
- Illes, J., Francis, W.S., Desmond, J.E., Gabrieli, J.D., Glover, G.H., Poldrack, R., Lee, C.J., Wagner, A.D., 1999. Convergent cortical representation of semantic processing in bilinguals. *Brain Language* 70, 347–363.
- Imamizu, H., Miyauchi, S., Tamada, T., Sasaki, Y., Takino, R., Putz, B., Yoshioka, T., Kawato, M., 2000. Human cerebellar activity reflecting an acquired internal model of a new tool. *Nature* 403, 192–195.
- Ivry, R.B., Fiez, J.A., 2000. Cerebellar contributions to cognition and imagery, in: Gazzaniga, M.S. (Ed.), *The New Cognitive Neurosciences*, second ed. MIT Press, Cambridge, MA, pp. 999–1011.
- Jancke, L., Gaab, N., Wustenberg, T., Scheich, H., Heize, H.-J., 2001. Short-term functional plasticity in the human auditory cortex: an fMRI study. *Cogn. Brain Res.* 12, 479–485.
- Kawato, M., 1999. Internal models for motor control and trajectory planning. *Curr. Opin. Neurobiol.* 9, 718–727.
- Kent, R.D., Tjaden, K., 1997. Brain functions underlying speech, in: Hardcastle, W.J., Lavers, J. (Eds.), *The Handbook of the Phonetic Sciences*, Blackwell, London, pp. 220–255.
- Kim, K.H., Relkin, N.R., Lee, K.M., Hirsch, J., 1997. Distinct cortical areas associated with native and second languages. *Nature* 388, 171–174.
- Klein, D., Milner, B., Zatorre, R.J., Meyer, E., Evans, A.C., 1995. The neural substrates underlying word generation: a bilingual functional-imaging study. *Proc. Nat. Acad. Sci. USA* 92, 2899–2903.
- Kuhl, P.K., 2000. Language, mind, and brain: Experience alters perception, in: Gazzaniga, M.S. (Ed.), *The New Cognitive Neurosciences*, second ed. MIT Press, Cambridge MA, pp. 99–115.
- Liberman, A.M., Mattingly, I.G., 1985. The motor theory of speech perception revised. *Cognition* 21, 1–36.
- Lively, S., Pisoni, D., Yamada, R., Tohkura, Y., Yamada, T., 1994. Training Japanese listeners to identify English /r/ and /l/. III. Long-term retention of new phonetic categories. *J. Acoust. Soc. Am.* 96, 2076–2087.
- McCandliss, B.D., Fiez, J.A., Protopapas, A., Conway, M., McClelland, J.L., 2002. Success and failure in teaching the [r]-[l] contrast to Japanese adults: tests of a Hebbian model of plasticity and stabilization in spoken language perception. *Cogn. Affective, Behav. Neurosci.* 2, 89–108.
- Middleton, F.A., Strick, P.L., 2000. Basal ganglia and cerebellar loops: motor and cognitive circuits. *Brain Res. Rev.* 31, 236–250.
- Miller, K.D., Erwin, E., Kayser, A., 1999. Is the development of orientation selectivity instructed by activity. *J. Neurobiol.* 41, 44–57.
- Miyawaki, K., Strange, W., Verbrugge, R., Liberman, A., Jenkins, J.J., Fujimura, O., 1975. An effect of linguistic experience: The discrimi-

- nation of [r] and [l] by native speakers of Japanese and English. *Percept. Psychophys.* 18, 331–340.
- Muller, U., Wächter, T., Barthel, H., Reuter, M., von Cramon, D.Y., 2000. Striatal [123I]beta-CIT SPECT and prefrontal cognitive functions in Parkinson's disease. *J. Neural Transm.* 107, 303–319.
- Nakai, T., Matsuo, K., Kato, C., Matsuzawa, M., Okada, T., Glover, G.H., Moriya, T., Inui, T., 1999. A functional magnetic resonance imaging study of listening comprehension of languages in human at 3 tesla-comprehension level and activation of the language areas. *Neurosci. Lett.* 263, 33–36.
- Paulesu, E., Frith, C., Frackowiak, R.S., 1993. The neural correlates of the verbal component of working memory. *Nature* 362, 342–345.
- Perani, D., Dehaene, S., Grassi, F., Cohen, L., Cappa, S.F., Dupoux, E., Fazio, F., Mehler, J., 1996. Brain processing of native and foreign languages. *NeuroReport* 7, 2439–2444.
- Perani, D., Paulesu, E., Galles, N.S., Dupoux, E., Dehaene, S., Bettinardi, V., Cappa, S.F., Fazio, F., Mehler, J., 1998. The bilingual brain: proficiency and age of acquisition of the second language. *Brain* 121, 1841–1852.
- Poldrack, R., 2000. Imaging brain plasticity: conceptual and methodological issues—a theoretical review. *NeuroImage* 12, 1–13.
- Poldrack, R.A., Clark, J., Pare-Blagoev, E.J., Shohamy, D., Creso Moano, J., Myers, C., Gluck, M.A., 2002. Interactive memory systems in the human brain. *Nature* 414, 546–550.
- Price, C.J., Wise, R.J., Warburton, E.A., Moore, C.J., Howard, D., Patterson, K., Frackowiak, R.S., Friston, K.J., 1996. Hearing and saying: The functional neuro-anatomy of auditory word processing. *Brain* 119, 919–931.
- Recanzone, G.H., 2000. Cerebral cortical plasticity: perception and skill acquisition, in: Gazzaniga, M.S. (Ed.), *The New Cognitive Neurosciences*, second ed. MIT Press, Cambridge, MA, pp. 237–247.
- Schmahmann, J., Doyon, J., Toga, A.W., Petrides, M., Evans, A.C., 2000. MRI Atlas of the Human Cerebellum. Academic Press, San Diego.
- Scott, S.K., Blank, C.C., Rosen, S., Wise, R.J.S., 2000. Identification of a pathway for intelligible speech in the left temporal lobe. *Brain* 123, 2400–2406.
- Silveri, M.C., Di Betta, A.M., Filippini, V., Leggio, M.G., Molinari, M., 1998. Verbal short-term store-rehearsal system and the cerebellum. *Brain* 121, 2175–2187.
- Smith, E.E., Jonides, J., 1999. Storage and executive processes in the frontal lobes. *Science* 283, 1657–1661.
- Sur, M., Pallas, S.L., Roe, A.W., 1990. Cross-modal plasticity in cortical development: differentiation and specification of sensory neocortex. *Trends Neurosci.* 13, 227–233.
- Talairach, J., Tournoux, P., 1988. *Co-planar Stereotactic Atlas of the Human Brain*. Thieme, New York.
- Trepanier, L.L., Saint-Cyr, J.A., Lozano, A.M., Lang, A.E., 1998. Neuropsychological consequences of posteroventral pallidotomy for the treatment of Parkinson's disease. *Neurology* 51, 207–215.
- Van Mier, H., 2000. Human learning. in: Toga, A., Mazziotta, J. (Eds.), *Brain Mapping: The Systems*. Academic Press, San Diego, pp. 605–620.
- Van Oostende, S., Van Hecke, P., Sunaert, S., Nuttin, B., Marchal, G., 1997. FMRI studies of the supplementary motor area and the premotor cortex. *Neuroimage* 6, 181–190.
- Wiesel, T.N., 1982. Postnatal development of the visual cortex and the influence of environment. *Nature* 299, 583–591.
- Wildgruber, D., Ackermann, H., Grodd, W., 2001. Differential contributions of motor cortex, basal ganglia, and cerebellum to speech motor control: effects of syllable repetition rate evaluated by fMRI. *NeuroImage* 13, 101–109.
- Wise, R.J.S., Scott, S.K., Blank, C., Mummery, C.J., Murphy, K., Warburton, E.A., 2001. Separate neural subsystems within 'Wernicke's area.' *Brain* 124, 83–95.
- Zatorre, R.J., Evans, A.C., Meyer, E., Gjedde, A., 1992. Lateralization of phonetic and pitch discrimination in speech processing. *Science* 256, 846–849.