

Effects of language experience: Neural commitment to language-specific auditory patterns

Yang Zhang,^{a,*} Patricia K. Kuhl,^a Toshiaki Imada,^{a,b} Makoto Kotani,^c and Yoh'ichi Tohkura^{d,1}

^a*Institute for Learning and Brain Sciences, University of Washington, Box 357988, Seattle, WA 98195, USA*

^b*Research Center for Advanced Technologies, Tokyo Denki University, Inzai-shi, Chiba 270-1382, Japan*

^c*Department of Electronic Engineering, Tokyo Denki University, Chiyoda-ku, Tokyo 101-8457, Japan*

^d*NTT Science and Core Technology Group, Nippon Telegraph and Telephone Corporation, Atsugi-shi, Kanagawa 243-0198, Japan*

Received 13 July 2004; revised 15 January 2005; accepted 25 February 2005
Available online 13 April 2005

Linguistic experience alters an individual's perception of speech. We here provide evidence of the effects of language experience at the neural level from two magnetoencephalography (MEG) studies that compare adult American and Japanese listeners' phonetic processing. The experimental stimuli were American English /ra/ and /la/ syllables, phonemic in English but not in Japanese. In Experiment 1, the control stimuli were /ba/ and /wa/ syllables, phonemic in both languages; in Experiment 2, they were non-speech replicas of /ra/ and /la/. The behavioral and neuromagnetic results showed that Japanese listeners were less sensitive to the phonemic /r-l/ difference than American listeners. Furthermore, processing non-native speech sounds recruited significantly greater brain resources in both hemispheres and required a significantly longer period of brain activation in two regions, the superior temporal area and the inferior parietal area. The control stimuli showed no significant differences except that the duration effect in the superior temporal cortex also applied to the non-speech replicas. We argue that early exposure to a particular language produces a "neural commitment" to the acoustic properties of that language and that this neural commitment interferes with foreign language processing, making it less efficient.

© 2005 Elsevier Inc. All rights reserved.

Keywords: Magnetoencephalography; Native Language Neural Commitment Theory; Speech perception

Speech perception is altered by an individual's language experience. Adults have substantial difficulty categorizing and discriminating many non-native speech contrasts (Strange, 1995). This phenomenon extends to early bilinguals as well as listeners with extensive experience in the second language (Guion et al., 2000; Pallier et al., 1997). Studies further indicate that listeners of

different languages respond to distinct acoustic aspects of the same speech stimulus (Fox et al., 1995; Iverson et al., 2003; Jongman and Moore, 2000). When listening to the American English /r-l/ phonemes, for example, multi-dimensional scaling analyses suggest that American listeners attend to the third formant (Fig. 1a), which reliably distinguishes /r/ from /l/, whereas Japanese listeners respond more strongly to the second formant, one critical for related Japanese phonemes but not helpful in distinguishing /r/ from /l/ (Iverson et al., 2003).

The effects of language experience on speech perception are presumed to be due to neural coding of the acoustic components that are critical to native-language processing (Kuhl, 2000; Werker and Tees, 1999), a process that begins as early as 6 months of age (Kuhl et al., 1992). Neurophysiological studies provide corroborating data that show the existence of language-specific "memory traces" in adult listeners (Dehaene-Lambertz, 1997; Dehaene-Lambertz et al., 2000; Näätänen et al., 1997; Rivera-Gaxiola et al., 2000; Sharma and Dorman, 2000; Winkler et al., 1999) and their emergence in young infants (Cheour et al., 1998; Rivera-Gaxiola et al., 2005). These studies used the mismatch negativity (MMN) and/or its magnetic counterpart, mismatch field (MMF), obtained using magnetoencephalography (MEG), to examine the discrimination of two speech sounds (Näätänen, 2001). Discriminative neural responses can be elicited irrespective of the listener's attention by the presentation of an infrequent stimulus (deviant) in the midst of a repetitive stimulus (standard). The MMN/MMF has been proven to be a reliable neurophysiological marker of speech discrimination and changes with learning (Kraus et al., 1996; Menning et al., 2002; Tremblay et al., 1997; see Näätänen, 1992, 2001 for reviews). The first goal of the present study, therefore, was to use MEG to examine neural activation for the detection of speech-sound differences. We hypothesized that the MMF would be reduced for non-native differences, reflecting listeners' decreased sensitivity to the acoustic cues important for non-native contrasts.

Our second goal was to go beyond the examination of speech-sound difference detection. We were interested in cross-language

* Corresponding author. Fax: +1 206 221 6472.

E-mail address: yazhang@u.washington.edu (Y. Zhang).

¹ Currently, Human and Social Information Research Division, National Institute of Informatics, Tokyo, 101-8430, Japan.

Available online on ScienceDirect (www.sciencedirect.com).

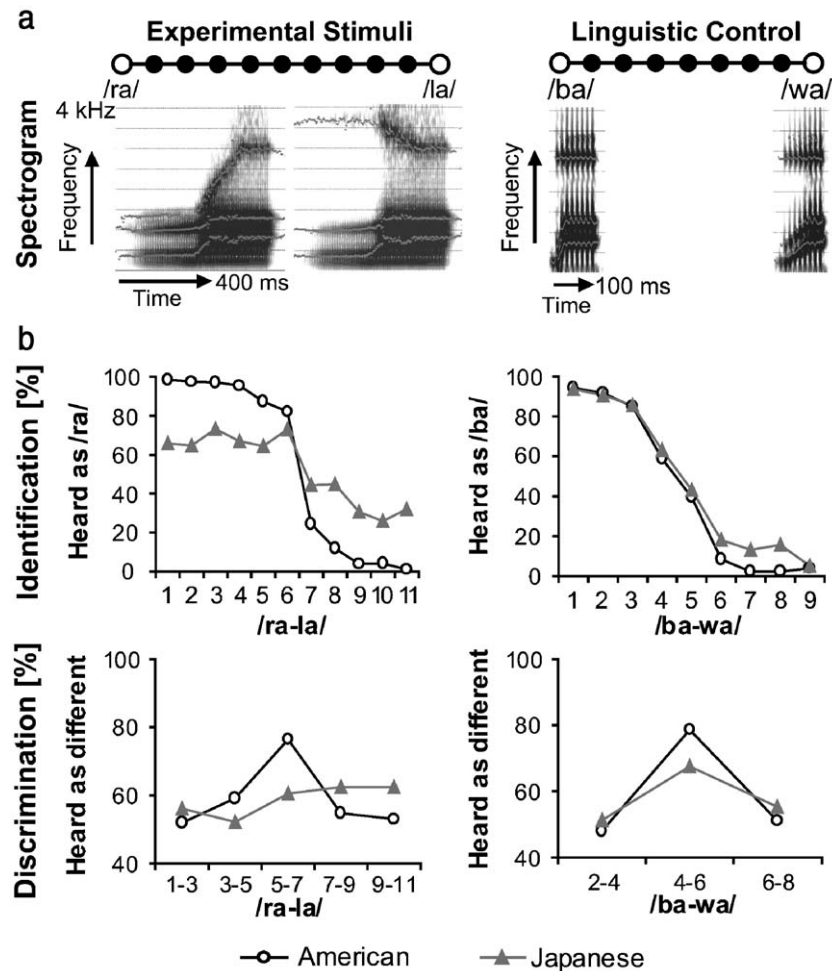


Fig. 1. (a) Spectrograms of the endpoint syllables of the /ra-la/ and /ba-wa/ series. The /ra-la/ syllables were 400 ms long and varied in the third formant (F3) transition in 10 steps; the starting frequencies of F3 varied from 1325 Hz to 3649 Hz and were spaced equally on the mel scale (Stevens et al., 1937). The /ba-wa/ syllables were 100 ms long and varied in the duration of the first and second formant transitions in 8 steps, from 16 to 48 ms with a step size of 4 ms. (b) Behavioral results of Experiment 1.

comparisons of brain activation for the individual speech stimuli. We have previously shown that, early in development, infants learn the prototypical patterns of phonemes in their language (Kuhl et al., 1992). Accordingly, we argue that brain activation to stimuli that are prototypical native-language examples for one population and non-prototypical exemplars for another population should show substantial differences in brain activation (Kuhl, 2004). More specifically, we hypothesized that processing native-language prototypes would show “neural efficiency,” defined as brain activation that was more focal, rather than highly distributed, and of shorter, rather than more sustained, duration. To test this hypothesis, we employed MEG to examine the properties of localized neural activities for the individual stimuli used in the experiments (Hämäläinen et al., 1993; Hari et al., 2000). This allowed us to assess the recruitment of brain regions, as well as the time course of activation, involved in processing individual speech sounds.

Neural efficiency in neurophysiological and brain imaging studies has been used to describe several desirable properties of the brain in connection with intelligence, expertise, and the effects of priming, practice, and learning (Ertl and Schafer, 1969; Haslinger et al., 2004; Jaušovec and Jaušovec, 2001; Landau et al., 2004; see

Deary and Caryl, 1997; Näätänen, 1973; Neubauer et al., 2004; Poldrack, 2000 for reviews). Neural efficiency in terms of activation is generally contained in two quantifiable properties: speed (the time it takes for a particular response to occur), and space, or the degree of activation (the spatiotemporal extent of brain resources used up by the construct). The faster the response is (presumably involving higher neural conduction speed and connectivity), the more efficient the neural system. Similarly, the smaller the extent of activation is (presumably involving less computational demand), the more efficient the system. In studies of speech perception and phonetic learning, recent fMRI data show that listening to non-prototype sounds activates greater auditory areas when compared to the over-learned prototypes (Guenther et al., 2004). Moreover, native-language listeners or faster learners of second language phonemes have more focal fMRI patterns of activation than non-native listeners or slower learners (Callan et al., 2004; Dehaene et al., 1997; Golestani and Zatorre, 2004). In view of these findings, we specifically tie the concept of neural efficiency to our “Native-Language-Neural-Commitment” (NLNC) hypothesis (Kuhl, 2004). Our MEG study differed from the previous fMRI studies by employing a passive listening condition and adding the activity duration measure. Thus, we were able to

examine neural efficiency in automatic processing of speech in both spatial and temporal domains.

Three sets of sounds were used in tests of Japanese and American listeners: /r/ and /l/, phonemic only in English; /b/ and /w/, phonemic in both Japanese and English, and non-speech analogs of /r/ and /l/, sounds that mimic the acoustic properties of speech without being perceived as speech. Previous behavioral and MEG research show robust differences in /r–l/ perception between the two populations (Bradlow et al., 1999; Goto, 1971; Guion et al., 2000; Iverson et al., 2003; Miyawaki et al., 1975; Phillips et al., 1995; Zhang et al., 2001). The use of both speech and non-speech control stimuli in our study allows assessment of auditory and phonetic processing in speech perception. The use of two controls also rules out possible confounding effects caused by potential inherent between-group differences. We predicted that brain and behavioral measures would show significant population differences for the /r–l/ contrast but not for the control stimuli.

Experiment 1

Methods

All experiments complied with the principles of research involving human subjects as stipulated by the University of Washington.

Subjects

Two groups of normal-hearing right-handed adults were tested, including ten Japanese (mean age = 22 ± 1 , 3 females and 7 males) and ten Americans (mean age = 27 ± 3 , 4 females and 6 males). Subjects were volunteers under informed consent. They were recruited after screening for hearing, handedness, and language background. The American subjects were monolingual English speakers. The Japanese subjects had received 8 years of English as a second language education in the middle school through college, which involved very limited exposure to spoken English.

Stimuli

The stimuli were two computer-synthesized series of sounds, /ra–la/ and /ba–wa/, adopted from previous behavioral studies (see Iverson and Kuhl, 1996; Miller and Liberman, 1979 for detailed description). Each contrast was created on a carefully controlled acoustic continuum (Fig. 1a). For each stimulus in the series, all acoustic factors were identical with the exception of the critical acoustic parameter distinguishing the phonemes.

Behavioral experiment procedure

The behavioral experiments started with a short practice phase immediately followed by the test phase. The practice phase familiarized listeners with the synthesized stimuli and testing procedure. The identification task required subjects to identify each stimulus in the series when randomly presented in isolation. The testing session presented 40 trials for each stimulus in random order. The discrimination task required subjects to judge whether pairs of presented stimuli were the same or different. The stimulus pairs had equal physical differences. The inter-stimulus interval was 250 ms. Stimulus presentation was randomized, and both directions of presentation order were tested, each for 20 trials. An equal number of control trials were used, in which the same sound was presented twice, to assess false positives.

MEG experiment procedure

Two experiments were conducted using the oddball paradigm (Näätänen, 2001), in which stimuli (two endpoints from either the /r–l/ or /b–w/ continuum) were delivered to the right ear through an echo-attenuated plastic tube system terminating in a non-magnetic foam earplug. The sound output intensity was at 80 dB sound pressure level (SPL). In each experiment, stimulus presentation consisted of two consecutive blocks with the standard and the deviant reversed in the second block; the block sequences were counter-balanced among subjects. The deviant and standard presentation ratio was 15:85, and the inter-stimulus interval was randomized between 800 ms and 1200 ms. Subjects read self-chosen books under the instruction to ignore the stimuli during the recording session.

MEG recording

MEG data were recorded using whole-scalp planar-type 122-channel sensor system (Neuromag Ltd., Finland) in a four-layered magnetically shielded room. The MEG signals were bandpass-filtered from 0.03 to 100 Hz and sampled at 497 Hz. Entire epochs with magnetic amplitude bigger than 3000 fT/cm or EOG (electro-oculogram) bigger than 150 μ V in any channel were rejected to exclude data with blinking/movement artifacts or other noise contamination. At least 100 epochs were averaged for the deviant.

MEG analysis

Epochs immediately preceding the deviant were averaged for the standard, resulting in a number of averages comparable with the deviant. The averaged data were digitally low-pass filtered at 40 Hz, and then the DC-offset during a prestimulus baseline of 100 ms was removed. The MMF calculation used standard and deviant responses to the identical stimulus from two consecutive presentation blocks in order to eliminate effects due to the inherent acoustic differences in different standard and deviant stimuli within a block. The waveform amplitude used in the statistics was defined as the vector sum of amplitudes at two orthogonal channels in the same sensor location. To confirm the validity of the MMF peak amplitude values, we calculated the baseline noise. We also calculated the noise level of the system when no subject was being tested for a comparison.

MMF analysis

In the MMF waveform obtained by subtracting the response to the standard sound from the one to the deviant sound, the maximum peak amplitude was searched for, in each subject, in the latency range between the N1m peak latency and 500 ms among 9 channel locations (18 channels) covering the left or right auditory region; therefore, the channel that showed a maximum peak amplitude was different depending on the subject. The obtained amplitude and latency were defined as the MMF peak amplitude and latency.

ECD analysis

We assumed a single moving equivalent current dipole (ECD) in a spherical head model to estimate the source of neuromagnetic activities under a small region of selected channels. A total of 70 configurations of local channel selection were used (see Fig. 2 for examples). The number of channels in each configuration ranges as follows: 14 channels (11 configurations), 16 channels (33 configurations), 18 channels (25 configurations), and 20 channels (1 configuration). Our MEG system houses 122 planar gradi-

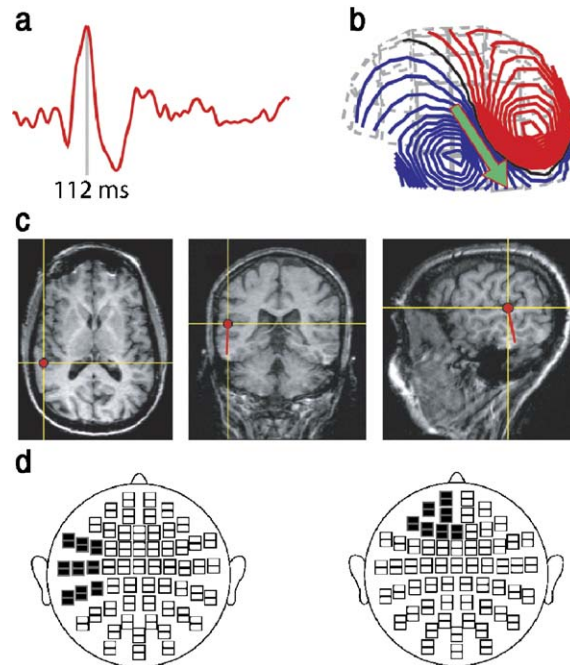


Fig. 2. Sample ECD analysis. (a) Waveform data from one MEG channel of an individual American subject's response to the standard /ra/ sound. (b) Field contour map with a dipolar pattern for the response at 112 ms (red curves indicating the magnetic flux coming out of the head, blue curves indicating the magnetic flux going into the head, and the green arrow schematically indicating the current source location, magnitude, and current direction). 18 channels over the left temporal lobe (solid line) were used for source estimation in this analysis. (c) 3-D location of the ECD activity (red dot at the crossing point of two yellow hair lines) in the auditory cortex, together with the current direction (red line from red dot), in axial, coronal, and sagittal sections. (d) Two examples of channel selection configuration (left temporal and left frontal) out of 70 used in estimating ECDs.

meters with a short baseline (16.5 mm). To try to make use of the advantages of this short baseline, we used a single ECD model to localize the activity presumably scattered in more than one brain region. In particular, we performed the following procedures.

1. Localize a single ECD as a function of time (spatiotemporal ECD) using each of 70 channel configurations. The ECDs are obtained every 2 ms for 680 ms from 20 ms after the stimulus onset.
2. Select the ECDs that satisfy the following criteria: (a) goodness-of-fit $\geq 80\%$, (b) 95%-confidence-volume $\leq 4188.8 \text{ mm}^3$ (corresponding to the volume of 20-mm diameter sphere), (c) location is further than 30 mm from the sphere center approximating the brain and is directly beneath the channel whose amplitude is closer than the baseline noise to the maximum vector-sum amplitude in the channel configuration for obtaining the ECD. Since our MEG system was the planar-type sensor system with a very low ambient noise level, the selected ECDs in our analysis were found directly beneath the channel with the maximum magnetic amplitude in most cases.
3. Make an ECD cluster by grouping ECDs obtained in 2 above, so that the distance between any pair of ECDs at the same sampling point within each new spatial ECD cluster is less than or equal to 20 mm (within a sphere of 20-mm diameter). These ECD clusters are called "spatial ECD clusters."
4. Make a new ECD cluster, from spatial ECD clusters or from single ECDs, by grouping temporally contiguous ECDs, so that the distance between any pair of ECDs of adjacent temporal sampling points within each new ECD cluster is less than or equal to 10 mm (within a sphere of 10-mm diameter), and so that the distance between any pair of ECDs, irrespective of their latencies, within each new ECD cluster, is less than or equal to 20 mm (within a sphere of 20-mm diameter). These ECD clusters are called "spatiotemporal ECD clusters."
5. From these spatiotemporal ECD clusters obtained in 4 above, select only the ECD clusters with the activity continuing at least 5 sampling points, approximately 10 ms given our sampling frequency at 497 Hz.
6. Hereafter, the spatiotemporal ECD clusters are simply referred to as "ECD clusters." The activity duration of an ECD cluster is an OR-sum of each element ECD of this ECD cluster. In this calculation, the active sampling point duplicated by more than one single ECD within a single ECD cluster is counted as a single active sampling point.
7. The scatter of ECD activities within a region of interest (ROI) is measured by the number of ECD clusters in the region.
8. The duration of ECD activities within an ROI is measured by a simple sum of the duration of each ECD cluster in the region.
9. The definitions of spatiotemporal ECD clusters shown in 7 and 8 above ensure that the activity within 20-mm distance is counted as a single spatiotemporal activity.
10. The anatomical locations of the ECD clusters are identified based on each subject's MR images. The superior temporal cortex (ST) is defined as the region below the lateral sulcus (inclusive) and above the superior temporal sulcus (not inclusive) within the temporal cortex. The middle temporal cortex (MT) is defined as the region below the superior temporal sulcus (inclusive) and above the inferior temporal sulcus (not inclusive) within the temporal cortex. The

inferior parietal region (IP) is defined as the region posterior to the post-central sulcus (not inclusive), below the intra-parietal sulcus (inclusive), and above the line connecting the point where the lateral sulcus starts ascending and the anterior end point of the lateral occipital sulcus. The inferior frontal cortex is defined as the region below the inferior frontal sulcus (inclusive), anterior to the inferior precentral sulcus (inclusive), and posterior to the lateral orbital sulcus (not inclusive).

Current source density analysis

To confirm the ECD analysis results, we performed two current density analyses, namely, MNE (minimum norm estimation obtaining a solution that has a minimum L2-norm, Hämäläinen and Ilmoniemi, 1984, 1994) and MCE (minimum current estimation obtaining a solution that has a minimum L1-norm, Uutela et al., 1999). Instead of point-like sources, MNE and MCE algorithms search for the best estimate of a distributed primary current and select the solution with the smallest norm from all current distribution to explain the measured magnetic field. The MNE/MCE solutions require no a priori assumptions about the nature of the source current distribution and are considered to be more appropriate when the activity distribution is poorly known (Hämäläinen et al., 1993; Hari et al., 2000).

MNE analysis

Specifically, we used the following procedure for MNE, an implementation of the minimum L2-norm estimate:

1. A sphere model with 75-mm radius is used throughout all subjects.
2. The electric current locations about 20 mm apart from each other are assumed.
3. The singular value whose magnitude is at least one twentieth of the maximum singular value is employed.
4. We define that the current location is active when it satisfies the following criteria: (a) The current strength is greater than $Q = Q_{na} + 2.326 * Q_{ns}$, where Q_{na} is an average current strength during baseline and Q_{ns} is its standard deviation (one-tailed analysis). Although the essential problem of what patches in the model should be counted as being active is not solved by this definition, the probability that the activity is derived from noise is less than 1% in this statistical definition. (b) The activity satisfying (a) continues at least 5 sampling points, which is approximately 10 ms.
5. We identify two anatomical regions in this analysis based on the same criteria as those used in a single ECD model analysis, auditory cortex (ST + MT) and inferior parietal region (IP).

MCE analysis

For MCE analysis (Neuromag MCE Version 1.3, an implementation of the minimum L1-norm estimate), we used the following procedure:

1. A realistic head model (standard brain in Boundary Element Model (BEM)) is used as the source space throughout all subjects.
2. Individual subjects' BEMs and their head origin parameters are used for the conductor model in the forward calculations.
3. Electric current locations about 10 mm apart from each other are assumed.

4. Thirty singular values are employed for regularization.
5. The estimate is calculated separately for each time point. The activity within a latency range from 20 to 700 ms is integrated and averaged across the subjects.

Statistical analysis

We performed two-tailed Student's *t* tests on behavioral measures. To statistically examine the differences in degrees of brain activation between two subject groups, we performed a repeated measures ANOVA (two subject groups [Am and Jp] \times two stimuli [either /ba/ and /wa/ or /la/ and /ra/] \times two hemispheres \times two stimulus roles [deviant and standard]), in each experiment (either /ba–wa/ or /ra–la/), on the following quantities: for MMF waveform analysis, we used the MMF peak amplitude and latency; for single ECD model analysis, we used the number of ECD clusters in the pre-determined regions of interest (ROIs) and the sum of active duration of ECD clusters in the ROIs; for MNE, we used the number of active locations in the ROIs normalized by the total number of current locations in the corresponding ROIs, the weighted number of active locations in the ROI (the weight was the current strength) normalized by the total number of current locations in the ROIs, and the sum of active duration of all active current locations in the ROIs normalized by the total number of current locations in the ROIs. Although the validity of our statistical definitions for duration and scatter of brain activities in the MNE analysis is unknown and remains to be tested against actual data, we take the stand that the current strength in MNE solutions often reflects the extent of the active area (personal communications with Uutela). Therefore, we included the current-weighted number of active locations in our statistical calculation of the MNE results. For MCE, no statistical calculation was done because we restricted the number of active current locations. When necessary, we also performed a repeated measures ANOVA separately in each subject group.

Results

Behavioral results

Subjects listening to native-language contrasts (American subjects listening to /r–l/ and /b–w/ and Japanese subjects listening to /b–w/) showed typical identification and discrimination functions as seen in classic studies of categorical perception (Fig. 1b) (Liberman et al., 1957). Two-tailed Student's *t* tests revealed a number of between-group differences for the /r–l/ continuum (Table 1). As expected, in the /r–l/ identification task, American listeners were significantly better at labeling the stimuli except for stimulus 6 at the phonetic boundary region and in the /r–l/ discrimination task, American listeners were significantly better at cross-category discrimination (pair 5–7). Japanese listeners showed higher sensitivity to within-category differences (pairs 7–9 and 9–11). For the /b–w/ continuum, there were no significant between-group differences in identification. However, in the /b–w/ discrimination task, the American listeners were also better at cross-boundary discrimination (pair 4–6), perhaps because they reflected American pronunciations of these sounds.

MEG noise results

Mean noise level (\pm standard deviation) in the shielded room, averaged across all blocks in Experiment 1 (across all channels, all subjects, all stimuli, and all stimulus roles), was 4.95 ± 2.63 fT/cm with no subject being tested (averaged across 100 epochs) and

Table 1

Summary statistics for behavioral identification and discrimination in Experiment 1 (Fig. 1b)

Identification	/ra-la/	1	2	3	4	5	6	7	8	9	10	11
	Am vs. Jp	***	***	***	***	**	N.S.	*	**	***	**	***
	/ba-wa/	1	2	3	4	5	6	7	8	9		
	Am vs. Jp	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.
Discrimination	/ra-la/	1-3		3-5		5-7			7-9			9-11
	Am vs. Jp	N.S.		N.S.		***			*			*
	/ba-wa/		2-4			4-6			6-8			
	Am vs. Jp		N.S.			*			N.S.			

Note. *** stands for $P < 0.001$, ** $P < 0.01$, * $P < 0.05$ (two-tailed t test). N.S. indicates no significant difference. Unless shaded, the American (Am) listeners had better accuracy than the Japanese (Jp) listeners.

5.87 ± 2.97 fT/cm with a subject for equivalent number of accepted epochs. There was no significant difference between noise levels with American and Japanese subjects, between noise levels of any combination from /ra/, /la/, /ba/, and /wa/, or between noise levels during standard and deviant stimulations.

MMF results

The MMF peak amplitudes from all subjects, four stimuli, and two hemispheres were significantly larger than the baseline noise level [$P < 0.001$]. Repeated measure ANOVA showed significant between-group differences in the mismatch responses for /r-l/ but not for /b-w/ (Figs. 3a, b, c). As predicted, the /r-l/ MMF peak amplitudes were significantly smaller in the Japanese listeners [39.78 (Am) vs. 32.40 (Jp) fT/cm; $F(1,18) = 14.54$, $P < 0.01$]. There was also an interaction effect of hemisphere by group [$F(1,18) = 5.36$, $P < 0.05$]: American subjects exhibited larger MMF peak amplitudes in the left hemisphere that approach significance [42.09 (LH) vs. 37.48 (RH) fT/cm; $F(1,9) = 4.44$, $P = 0.06$] whereas Japanese subjects showed bilaterally equivalent MMF peak amplitudes [31.82 (LH) vs. 32.99 (RH) fT/cm]. In addition, Japanese listeners showed a significantly earlier MMF response than American listeners [413.61 (Am) vs. 337.78 (Jp) ms; $F(1,18) = 4.79$, $P < 0.05$]. There was no significant effect of stimulus (/ra/ vs. /la/ or /ba/ vs. /wa/) in either subject group.

ECD results

Significant between-group differences were found for /r-l/ but not for /b-w/. First, there were bilaterally more ECD clusters for /r-l/ in the Japanese listeners than in the American listeners [14.4 (Am) vs. 21.2 (Jp) in averaged total number of clusters; $F(1,18) = 4.74$, $P < 0.05$], indicating greater spread of activation in the processing of non-native speech stimuli (Figs. 3d, e, 4; also see Movie 1 in Appendix, provided as supplementary information, for ECD activities from two individual subjects). Second, we examined four regions of interest (ROIs) to locate the foci of between-group differences in terms of cumulative ECD duration. Significant differences were found in two regions (Fig. 3f): Japanese listeners had longer ECD durations for /r-l/ in ST (superior temporal) [44.42 (Am) vs. 64.43 (Jp) ms; $F(1,18) = 9.75$, $P < 0.01$] and IP (inferior parietal) [37.80 (Am) vs. 80.19 (Jp) ms; $F(1,18) = 9.42$, $P < 0.01$].

MCE/MNE results

The MCE results are shown in Fig. 5. Bilateral activations of the auditory cortex for /ra-la/ and for /ba-wa/ were consistently found for both subject groups. For the experimental /ra-la/ stimuli, the Japanese group appeared to show a larger area of activation than the American group. The control /ba-wa/ stimuli showed a similar tendency but to a lesser extent. These results were consistent with the ECD cluster results. Interestingly, in the MCE results, Japanese subjects seemed to show more anterior activation than American subjects, which could be a reflection of group difference in head shapes when mapping the activities onto the standard brain.

The MNE statistics showed the following results. For /r-l/, there were more active locations in Japanese than in American subjects, which approached significance [90.9% (Am) vs. 95.0% (Jp) in percentage of active locations; $F(1, 18) = 4.27$, $P = 0.054$]. On the other hand, there was no significant difference for /b-w/. Analyses on current-weighted number of active locations also showed consistent results. For /r-l/, there were significantly more locations in Japanese than in American subjects [62.1% (Am) vs. 71.4% (Jp) in weighted percentage; $F(1, 18) = 5.81$, $P < 0.05$] whereas there was no significant difference for /b-w/. Consistent with ECD analysis, there were no significant effects of hemispheric laterality or stimulus (/ra/ vs. /la/ or /ba/ vs. /wa/). Unlike the ECD results, however, the active duration as defined in our MNE analysis did not show any significant difference either for /r-l/ or for /b-w/ between Japanese and American subjects.

Discussion

The behavioral results obtained in Experiment 1 are consistent with earlier studies (Iverson et al., 2003; Miller and Liberman, 1979; Miyawaki et al., 1975; Zhang et al., 2001), showing that categorical perception of speech sounds is influenced by language experience. As expected, Japanese speakers had difficulty categorizing the /r-l/ sounds. However, strict categorical perception cannot fully explain the data. Japanese listeners' difficulty with /r-l/ was double-fold. Japanese listeners were *less* accurate at detecting the cross-category difference when compared to Americans. Meanwhile, they were *more* accurate at detecting within-category acoustic differences when compared to Americans. Better

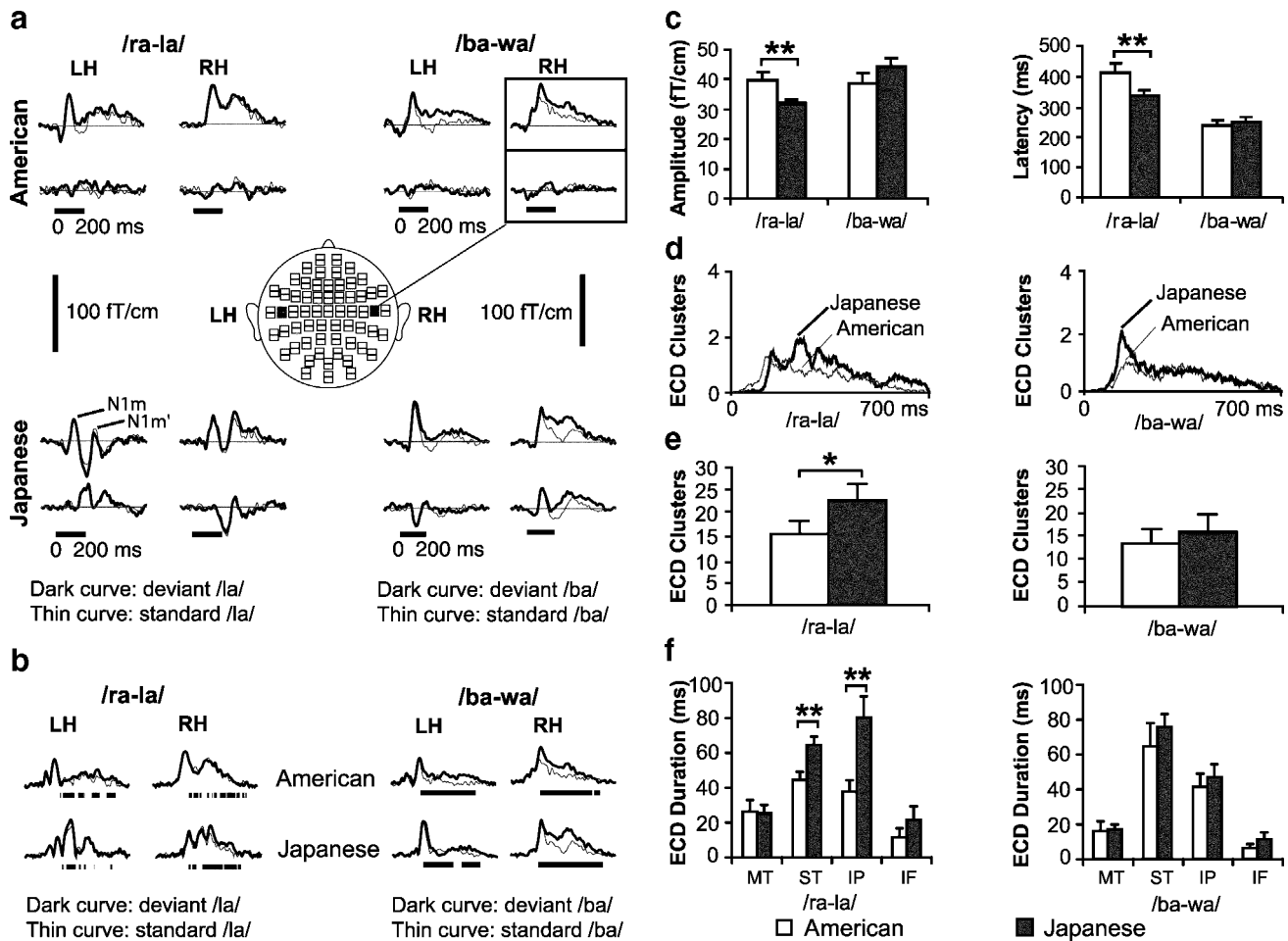


Fig. 3. MEG results of Experiment 1 (* $P < 0.05$; ** $P < 0.01$). (a) MEG waveform. Group averages are shown from one recording site of the left hemisphere (LH) and one recording site of the right hemisphere (RH) (thick line = deviant, thin line = standard) for /la/ and /ba/ syllables respectively. Each recording site has two channels in the planar MEG system. The double N1m responses to stimulus onset and F3 transition onset of /r-l/ are marked in the Japanese listeners' waveform. Group averages are obtained as follows. First, the center of the subject's sphere approximating his/her own brain is moved to the origin of the device coordinate system. Second, x, y, and z axes of the subject's head coordinate system are made in parallel with x, y, and z axes of the device coordinate system, respectively. Third, the subject's head is rotated by 15° upward around x axis connecting from the left preauricular point to the right one. By this procedure, each subject's head position is moved to the same standardized position. Fourth, accordingly, the MEG waveforms are recalculated using MNE estimates employing the singular values up to the magnitude of one twentieth of the maximum singular value. Fifth, these virtual waveforms are averaged across all subjects in a channel-wise fashion. (b) Vector sum amplitude waveforms of the selected channels in panel (a). The dark horizontal bars underneath the waveforms indicate the time range(s) where the time-point-to-time-point differences between deviant and standard were significant [$P < 0.01$] compared to baseline noise. (c) Mean maximum peak MMF amplitude and latency. Note that the maximum MMF peaks did not appear in the same channel at the same latency across subjects. Thus, the values may appear larger than those reflected by the average waveform in panel (a). (d) Mean total number of active ECD clusters in the whole brain as a function of time. (e) Mean total number of ECD clusters in the whole brain during the 20–700 ms time window. (f) Mean total ECD duration in four brain regions. The data in panels (c, d, e, f) were collapsed over the two hemispheres and the two stimuli for simplicity. No statistically significant effects for hemisphere or stimulus were found unless specified in the text.

within-category discrimination for Japanese listeners is predicted by a model in which linguistic experience results in acquired similarity among the stimuli within a category, as shown in experiments on infants and adults, and consistent with Kuhl's theorizing that exposure to a specific language in early infancy alters perception to make category exemplars appear more similar to the category prototype (Iverson and Kuhl, 1996; Iverson et al., 2003; Kuhl, 1991; Kuhl et al., 1992; see Guenther and Gjaja, 1996 for a neural model of the effect).

The MMF results were consistent with previous studies showing reduced neural sensitivity to non-native speech contrasts (Dehaene-Lambertz, 1997; Dehaene-Lambertz et al., 2000; Näätänen et al., 1997; Phillips et al., 1995; Rivera-Gaxiola et al., 2000;

Sharma and Dorman, 2000; Winkler et al., 1999; Zhang et al., 2001; see Cheour et al., 1998 and Rivera-Gaxiola et al., 2005 for infant data). These findings suggest that early language experience may have reshaped the state of the neural system at a pre-attentive level, altering perception in a way that assists native-language processing (Näätänen, 2001). The fact that adult listeners show mismatch responses to phonemic distinctions irrelevant to their native language suggests that acquired language-specific representations do not completely "filter" out phonetically irrelevant information. Koyama et al. (2000b, 2003) proposed that Japanese listeners' reduced sensitivity for /r-l/ could result from backward masking of the foreign consonants by subsequent vowels. However, the same research group also found similar acoustic

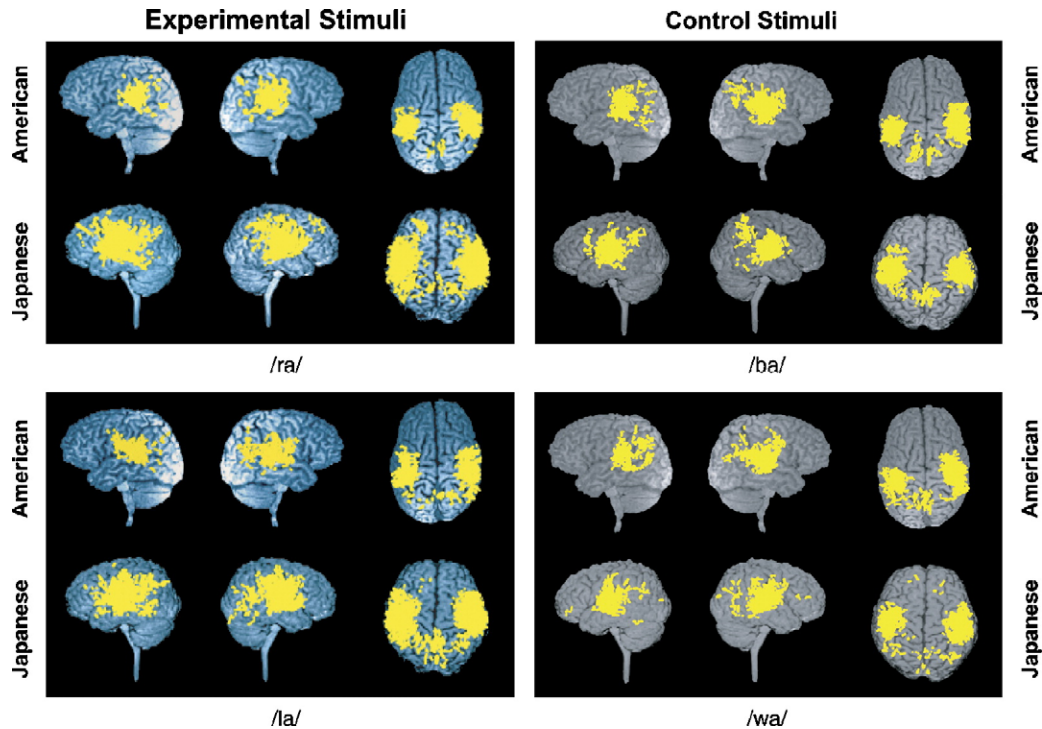


Fig. 4. ECD spatial distribution data. Composite category-specific ECD activities (yellow dots) from 10 American listeners (top in each panel) and 10 Japanese listeners (bottom in each panel) are shown for each of the two experimental stimuli, /ra/ and /la/ (blue, left side), and each of the two control stimuli, /ba/ and /wa/ (gray, right side). Left brain, right brain, and top-of-the-head view are shown in each row. The ECDs were initially calculated using each individual subject's head model based on their MR images. For illustration purpose, the ECD coordinates were linearly transformed, irrespective of the cortical regions, to fit within the spherical head model of one individual subject's brain from each subject group to represent their respective groups. As individuals vary in the configuration of cortical structures, head size and shape, the shown ECDs were a very rough approximation of the actual activities. Note also that the lack of ECD depth information in these plots may give a false interpretation of the exact brain regions of activation. For example, an active yellow dot on the Sylvian fissure (auditory region) may be projected along the central sulcus of the top-view brain.

backward masking effects in the American listeners and relatively smaller MMFs in Americans compared with Japanese listeners (personal communication with Koyama). It remains to be explored whether the experience-dependent changes in perception are due to changes in lower-level auditory analyzers, attentional changes, or higher-order linguistic processes, like word learning. It is also noteworthy that Koyama et al.'s /ra-la/ stimuli included acoustic differences not only in F3 but also in F1 and F2. Given the fact that native and non-native listeners perceptually weigh acoustic dimensions differently (Fox et al., 1995; Iverson et al., 2003; Jongman and Moore, 2000), there could have been confounding contributions to the MMF response from F1 and F2 differences that are not critical for the /r-l/ distinction.

A result not previously observed was the significantly earlier MMF for Japanese listeners for /r-l/; this is noteworthy, given that a smaller mismatch response is typically associated with a later latency (Näätänen, 1992). We speculate that the MMF latency difference may reflect Japanese listeners' failure to process the acoustic information differentiating /r/ and /l/ as an integrated whole including the late-occurring F3 transition, which is critical for the /r-l/ distinction. The temporal integration window of successive auditory information is purported to be within the range of 200 ms of the initial stimulus (Cowan, 1984; Inouchi et al., 2003; Loveless and Hari, 1993; Näätänen, 1992; Winkler et al., 1998; Yabe et al., 1998). Significant cross-linguistic differences were found in processing speech sound duration information (Minagawa-Kawai et al., 2004; Nenonen et al., 2003). The /r-l/

stimuli used in the present tests contained two acoustic sources with well-defined onsets in the consonant, the 155 ms steady F3 portion (not critical to /r-l/ differentiation) and the 100 ms F3 transition (Fig. 1a) (critical to the /r-l/ distinction) (Iverson and Kuhl, 1996; Iverson et al., 2003). The acoustic differences following both onsets in the stimuli can contribute to the MMF. The data indicate that Japanese listeners strongly reacted to both onsets, resulting in a double N1m in both hemispheres (Fig. 3a). American listeners, on the other hand, did not show clear second N1m response in the left hemisphere. We interpret the MMF and N1m differences as suggesting that the American responses reflect phonetic processing of unitary /r/ and /l/ percepts whereas the Japanese responses reflect acoustic processing of two distinct physical changes in the stimulus. Essentially, we claim that, when Americans listen to /r/ and /l/, their automatic processing strategies cause them to "wait" for the late-occurring F3 transition information that is critical to the /r-l/ distinction. This reasoning implies that the neural processing window for speech is affected by linguistic experience.

The ECD and MCE/MNE results suggest that non-native speech processing is less efficient. The first discussion on the relationship between processing efficiency and activation level can probably be found in Näätänen (1973) where he describes learning as an increased differentiation of the activation pattern, so that when performance is highly efficient, only mechanisms absolutely necessary for the occurrence of performance are activated (also see Hebb, 1949 for discussion on efficiency of neuronal firing). In

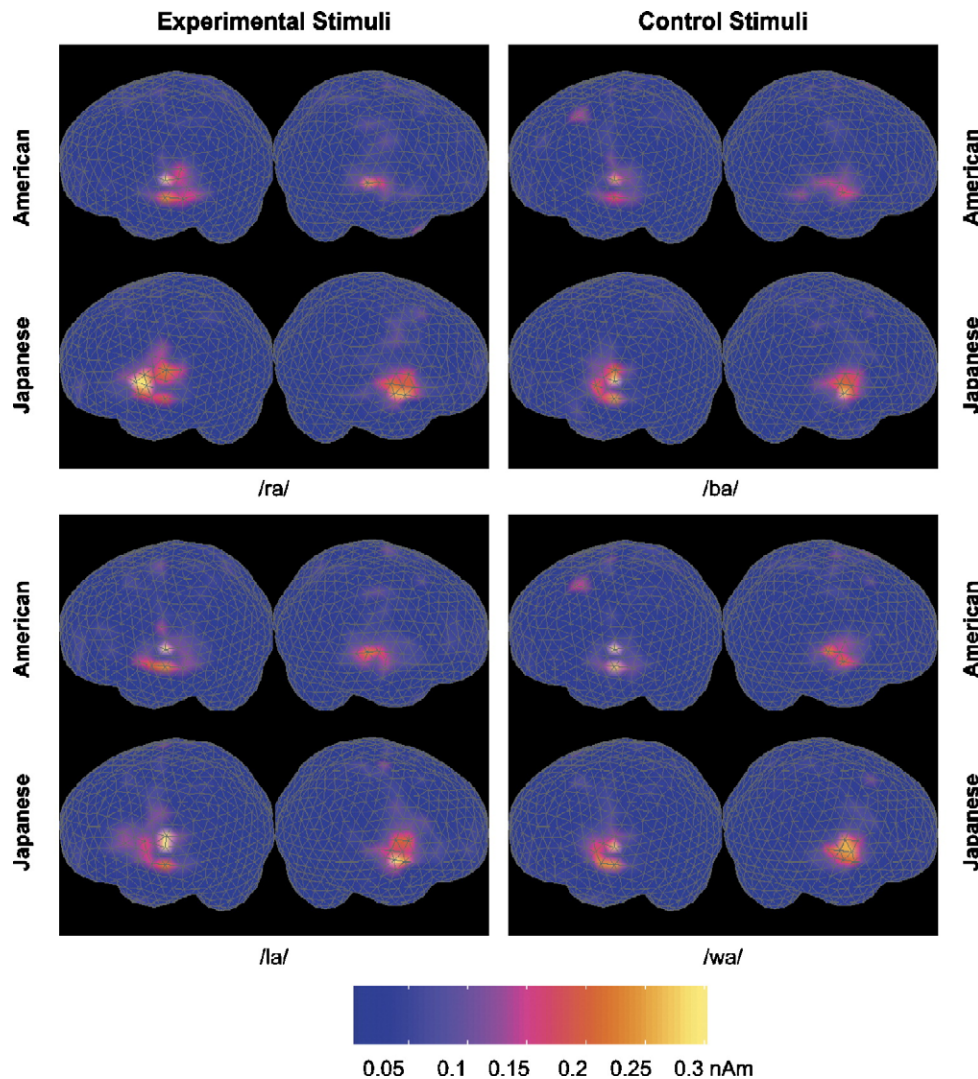


Fig. 5. Minimum current estimates for each stimulus in Experiment 1. Samples between 20 and 700 ms were integrated and averaged for each subject group.

particular, our ECD results here show evidence in the domain of language that, when processing non-native speech, as opposed to native, there is a significantly greater spread of activation and longer duration of activation. The spatial distribution patterns are also consistent with the MCE and MNE results. While previous fMRI studies have suggested the lack of efficiency in processing non-native or non-prototypical speech (Callan et al., 2003, 2004; Golestani and Zatorre, 2004; Guenther et al., 2004; Wang et al., 2003), the results could largely be a reflection of greater attention required of the task to recognize the non-native/non-prototypical speech. In our MEG study, the neuromagnetic experiments preceded the behavioral tests. A passive listening condition was used for MEG recordings with explicit instruction for subjects to focus on reading. Although it is possible that Japanese listeners might not be able to ignore the /r-l/ stimuli as easily as Americans, our design allowed a minimization of the attentional influences on the observed cross-linguistic differences. We argue that language experience results in neural networks dedicated to the automatic processing of native-language acoustic information and that this gives native speakers an advantage in processing efficiency (Guenther et al., 2004; Kuhl, 2004; Kuhl et al., 2001). In the absence of neural networks dedicated to the properties of the

incoming sounds, brain activation is not as focal and requires a longer processing time. Regional brain analyses suggest further that the auditory cortex and the inferior parietal area may be involved in comparing incoming acoustic information with the neural networks shaped by experience to react to native or newly acquired speech information (Hickok and Poeppel, 2000; Näätänen, 2001; Scott and Johnsrude, 2003; Wang et al., 2003).

The reconstruction of the underlying sources from MEG/EEG signals is known to be a severe inverse problem. Although the ECD and MCE/MNE approaches have very different assumptions in source modeling, we did not find contradictory results regarding the spatial scatter of activation. Spatial ECD clustering techniques were previously applied to modeling multiple sources (e.g., Hellstrand et al., 1993; Lewine et al., 1999). Our parameter choice for ECD clustering by a 20 mm diameter was in agreement with the minimum requirement for a reasonably good separation of multiple generator sources that could be temporally and spatially overlapping (Fujimaki et al., 2002; Hari et al., 1988, 2000; Imada et al., 2001; Lütkenhöner, 1998; Liu et al., 1998; Lütkenhöner et al., 1991). It has also been shown that, when activities are several centimeters apart, the localization results using the ECD model and MNE, MCE, or other kinds of current density models are largely

consistent (Jensen and Vanni, 2002; Komssi et al., 2004; Singh et al., 1994; Stenbacka et al., 2002). For our purpose of comparing the scatter and duration of activation in two subject groups, the current source density approach does not appear as sensitive as the single ECD analysis. This is probably due to the distributed nature of the current source density solutions and the inadequacy of mathematical definitions for scatter and duration of activation in our analysis. In most of the distributed current source analyses such as MNE, the locations for current activities are pre-determined under the assumption that the activities are spread over all of these locations. Thus, in a hypothetical no-noise condition with a point source activity, the MNE results would be such that all pre-determined source locations have non-zero current. Because there is no noise, it would be hard to define which source locations are really active. This holds true even when the real activity has a small or large area rather than a single point. Therefore, one needs to define some objective threshold to obtain the “active” pre-determined locations, and theoretically, this threshold should not be affected by the existence of noise. When noise exists, there should be another threshold on top of the noiseless threshold. Previous research suggests that, when compared to the ECD analysis, noise tends to be more easily confounded with the real sources in the current density solutions (Stenbacka et al., 2002).

A number of MEG studies of speech perception on healthy subjects or clinical patients have demonstrated the feasibility of assessing the “degree of engagement” for a brain region in terms of the cumulative number of ECD sources as an index for the duration of activation (Breier et al., 2000; Maestú et al., 2002; Papanicolaou et al., 2003; Patariaia et al., 2004; Simos et al., 2000). For instance, Papanicolaou et al. (2003) applied the single moving ECD analysis to the study of hemispheric laterality in syllable perception for time windows in the range of 60–1400 ms. Our sequential ECD analysis used a shorter window of 20–700 ms, and our clustering approach is a further step towards a better evaluation of the extent of activation. Although our method is computationally more intensive, it has the advantages of separating the potentially different source generators in the whole cortex and calculating the duration of activation for each ECD cluster for ROI analyses. We were cautious about avoiding redundancy of sampling time points within an ECD cluster that might derive from the use of multiple selection configurations including partially overlapping local channels in our calculation.

Finally, the results of Experiment 1 did not suggest a strong pattern of laterality. In the MMF amplitude results, Americans showed a tendency toward left-hemisphere processing for the /r–l/ stimuli, but this was not the case for the /b–w/ stimuli. Japanese listeners did not show laterality effects for either contrast.

In summary, the results of Experiment 1 suggest that Japanese and American listeners neurally process the same acoustic information (American /r–l/ sounds) very differently. Japanese listeners process the non-native sounds less efficiently, activating larger areas of the brain for longer periods of time. Moreover, they neurally react to each acoustic change in the stimuli in a left-to-right fashion. In contrast, American listeners process the native-language sounds more efficiently, showing focal activation of short duration. In addition, they respond to the acoustic changes in speech events in a more integrated and holistic fashion; neural detection of change in the /r–l/ stimuli is delayed until the critical acoustic information occurs. In both cases, we attribute the population differences to linguistic experience.

This conclusion would be strengthened if both populations responded identically to a non-speech contrast that closely mimicked the /r–l/ sounds. In Experiment 2, American and Japanese listeners were tested with non-speech analogs of the /r/ and /l/ stimuli as well as the /r–l/ speech stimuli themselves. We hypothesized that Japanese listeners would show positive correlations between speech and non-speech measures, both on the behavioral and neural measures, and that American listeners would not.

Experiment 2

Methods

The identical procedure, equipment, and analysis methods of Experiment 1 were used. A subset of seven American listeners from Experiment 1 (3 females, 4 males) and a new group of eight Japanese listeners (age 22 ± 1 , 3 females, 5 males) participated in the experiment. The experimental stimuli were the /ra–la/ syllables of Experiment 1. The control stimuli were non-speech analogs of the /ra/ and /la/ syllables, created by replacing the fundamental frequency and four formants with sinusoidal components at the exact same frequencies (Fig. 6a). The subjects uniformly reported that the sine wave replicas were not perceived as speech. Unlike in Experiment 1, we simplified the behavioral identification task by using stimuli separated by two steps on the continuum. For the non-speech stimuli, subjects were asked to label each stimulus according to whether it contained a noticeable rising tone or a falling tone.

In addition to the same statistics as those in Experiment 1, simple pair-wise Pearson's correlation analyses were performed for the speech and non-speech data (behavioral measure and MEG measure) using Bonferroni adjusted probabilities.

Results

Behavioral results

American listeners were significantly better at identifying all the /r–l/ stimuli except for stimulus 5 (Fig. 6b and Table 2). In Experiment 1, the exception was stimulus 6 instead of 5. This is understandable given the context of omitting stimuli 2, 4, 6, 8, and 10 in the identification task, especially stimulus 6 at the phonetic boundary. Moreover, the /r–l/ discrimination task again showed that American listeners had better performance for the cross-category stimulus pair (5–7) and that Japanese subjects showed higher sensitivity to within-category discrimination (pair 9–11, not to pair 7–9, though). By contrast, the two subject groups had nearly identical performance for the non-speech continuum.

MMF results

The /r–l/ results replicated the findings of Experiment 1. There was no significant effect of stimulus (/ra/ vs. /la/ or their counterparts of acoustic analogs) in either subject group. Japanese listeners had reduced MMF amplitude and earlier MMF latency for /r–l/ (Figs. 7a, b, c). They also responded to the /r–l/ stimuli with a clear double N1m in both hemispheres, whereas Americans did not. No significant between-group differences in MMF amplitude and latency were found for the non-speech control. In addition, no significant laterality effects were found for /r–l/ in either population. However, the non-speech stimuli showed RH (right hemisphere) dominance in both subject groups [42.80 (LH) vs. 49.97 (RH) fT/cm; $F(1,6) = 4.81$, $P = 0.07$ for Americans, and 46.33 (LH) vs. 57.58 (RH) fT/cm; $F(1,7) = 10.56$, $P < 0.05$ for Japanese].

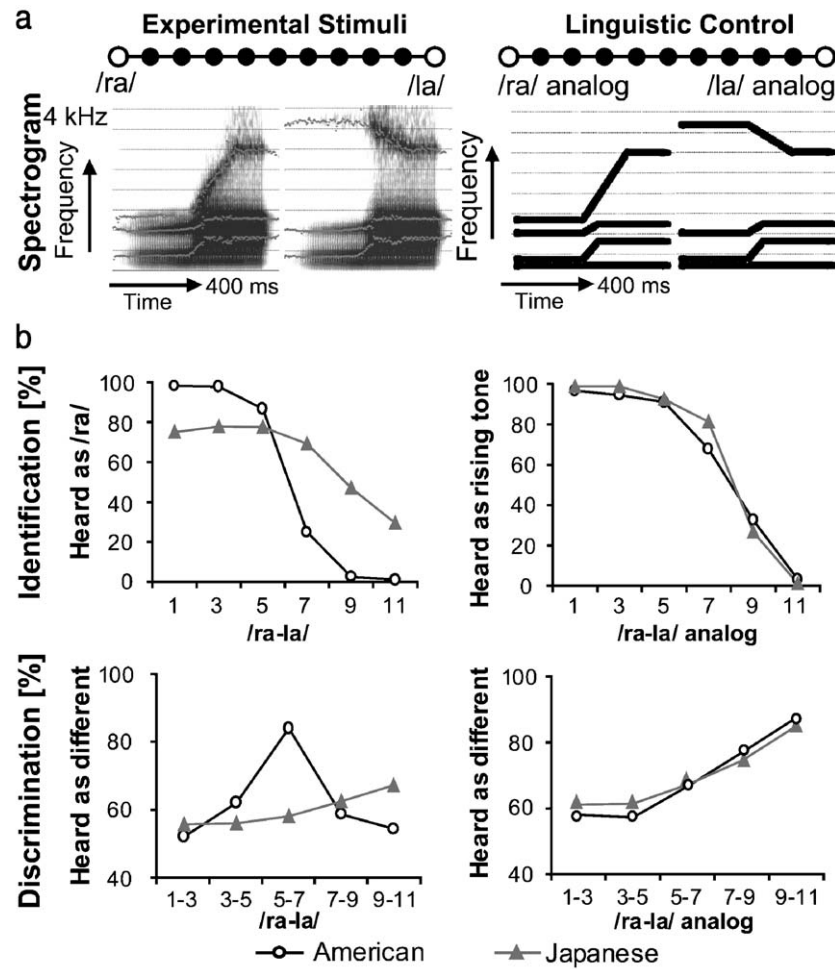


Fig. 6. (a) Spectrograms of the endpoint syllables of the /ra-la/ and /ra-la/ sine wave analog series. (b) Behavioral results of Experiment 2.

ECD results

As in Experiment 1, ECD cluster and duration measures showed between-group differences for /r-l/ (Figs. 7d, e, f, 8). Japanese subjects had more clusters in processing the /r-l/ sounds [13.75 (Am) vs. 23.09 (Jp); $F(1,13) = 3.68$, $P = 0.08$] with longer durations

in ST [48.46 (Am) vs. 116.98 ms (Jp); $F(1,13) = 7.31$, $P < 0.05$] and IP [37.63 (Am) vs. 65.43 ms (Jp); $F(1,13) = 4.19$, $P = 0.06$]. The reduced significance level here was likely due to the smaller sizes of subject groups. For the non-speech control, the ECD clusters did not show significant between-group differences. However, there were

Table 2

Summary statistics for behavioral identification and discrimination in Experiment 2 (Fig. 6b)

Identification	/ra-la/	1	3	5	7	9	11
	Am vs. Jp	***	***	N.S.	**	***	***
	Analog	1	3	5	7	9	11
	Am vs. Jp	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.
Discrimination	/ra-la/	1-3	3-5	5-7	7-9	9-11	
	Am vs. Jp	N.S.	N.S.	***	N.S.	*	
	Analog	1-3	3-5	5-7	7-9	9-11	
	Am vs. Jp	N.S.	N.S.	N.S.	N.S.	N.S.	

Note. *** stands for $P < 0.001$, ** $P < 0.01$, * $P < 0.05$ (two-tailed t test). N.S. indicates no significant difference. Unless shaded, the American (Am) listeners had better accuracy than the Japanese (Jp) listeners.

differences in ECD duration: Japanese listeners had longer activity in ST [87.30 (Am) vs. 160.12 (Jp) ms; $F(1,13) = 4.93$, $P < 0.05$], and American listeners had longer activity in IF [31.51 (Am) vs. 15.37 (Jp) ms; $F(1,13) = 4.51$, $P = 0.05$]. No significant effects of hemisphere or stimulus were observed in either subject group.

Speech–non-speech correlations

Simple pair-wise Pearson's correlation analyses showed the following results. For behavioral discrimination, Japanese listeners' responses to speech stimuli exhibited the same pattern as their responses to non-speech [$r = 0.79$, $P < 0.0001$]. The neural data showed similar results; in Japanese listeners, MMF amplitude in the right hemisphere was strongly correlated for speech and non-speech [$r = 0.71$; $P < 0.001$]. In addition, and as predicted, Japanese subjects also showed significant bilateral speech–non-speech correlations in ECD clusters [$r = 0.59$, $P < 0.01$ for LH; $r = 0.62$, $P < 0.01$ for RH] as well as in ECD durations [in ST, $r = 0.61$, $P < 0.01$; and in IP, $r = 0.52$, $P < 0.05$]. No significant correlations were seen between speech and non-speech, either in the behavioral or neural data, for the American listeners.

MCE/MNE results

The MCE results are shown in Fig. 9. As in Experiment 1, bilateral activation of the auditory cortex for /ra-la/ and for the

acoustic analogs was consistently found for both American and Japanese groups. The general patterns in the spread of activation between the American and Japanese groups also appeared consistent with those of Experiment 1 and the ECD results of Experiment 2. Unlike Experiment 1, however, the MNE statistics did not show any significant between-group differences in either the number of active locations or the active duration. Nevertheless, we still observed some consistency in the spatial distributions of activities. For /r-l/, there were more current-weighted number of active locations in the Japanese listeners than in Americans at a marginally significant level [60.9% (Am) vs. 72.6% (Jp) in weighted percentage; $F(1, 13) = 4.22$, $P = 0.061$]. On the other hand, no significant difference was found for the tonal analog stimuli in the weighted or non-weighted calculations. As in the ECD analysis, no significant effects of hemisphere or stimulus effects were observed in the MNE results.

Discussion

Experiment 2 essentially replicated the main results for /r-l/ in Experiment 1. First, American and Japanese listeners show large behavioral differences in the ability to categorize the /r-l/ stimuli. Second, Japanese listeners showed earlier and reduced mismatch responses to the /r-l/ stimuli. Third, Japanese listeners recruited

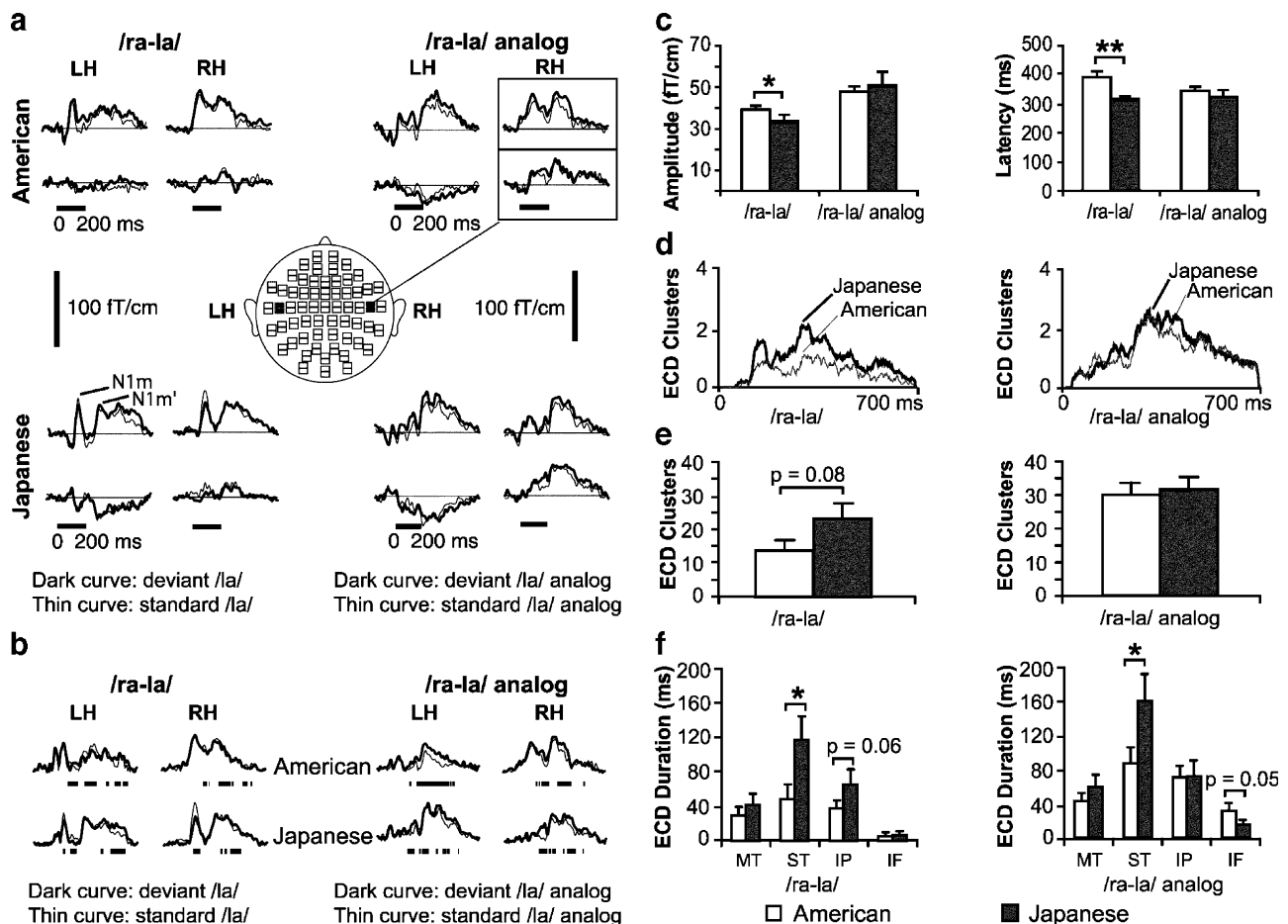


Fig. 7. MEG results of Experiment 2 plotted in the same convention as Fig. 3. (a) MEG waveform. (b) Vector sum waveforms and dark horizontal bars underneath the waveforms indicating the time range(s) where significant differences between deviant and standard were found. (c) Mean maximum peak MMF amplitude and latency. (d) Mean total number of active ECD clusters in the whole brain as a function of time. (e) Mean total number of ECD clusters in the whole brain during the 20–700 ms time window. (f) Mean total ECD duration in four brain regions.

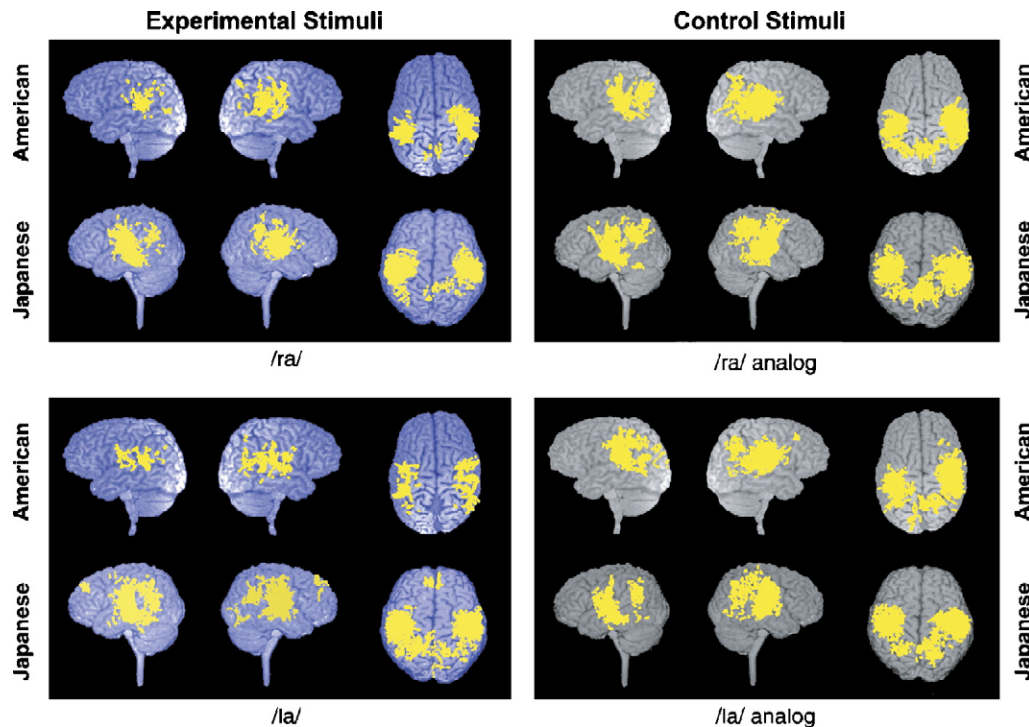


Fig. 8. ECD spatial distribution data plotted in the same convention as Fig. 4. Composite category-specific ECD activities (yellow dots) from 7 American listeners (top in each panel) and 8 Japanese listeners (bottom in each panel) are shown for each of the two experimental stimuli, /ra/ and /la/ (blue, left side), and each of the two control stimuli, /ra/ analog and /la/ analog (gray, right side). Left brain, right brain, and top-of-the-head view are shown in each row.

bilaterally more brain resources with relatively longer activities in ST and IP regions. The /r–l/ analogs did not show such differences. The ECD and MCE/MNE results in Experiment 2 further indicate that brain activities in both populations are bilaterally distributed for the speech stimuli as well as the non-speech control. Both populations also show overlapping brain regions activated for speech and non-speech, consistent with earlier fMRI studies (Binder et al., 2000; Joanisse and Gati, 2003).

The non-speech ECD duration results will be interesting to explore in future studies as we found significantly longer ST activity in Japanese listeners and significantly longer IF activity in American listeners. Language experience may change the neural responses for non-speech spectral patterns that mimic those found in speech (Salmelin et al., 1999; Vihla et al., 2002). When listening to such non-speech stimuli, listeners may demonstrate increased sensitivities to the learned auditory patterns (Hsieh et al., 2001). In a sense, non-speech that closely maps native speech may “fool” neural networks dedicated to speech.

A noteworthy finding of Experiment 2 is that only Japanese listeners showed significant correlations between speech and non-speech. This was the case for behavioral discrimination, MMF, ECD cluster, and duration measures. Thus, for Japanese listeners, /r–l/ speech and their non-speech analogs appear to draw on very similar auditory processing strategies except when they were asked to label the stimuli. In contrast, American listeners use distinct processing strategies for /r–l/ speech and /r–l/ non-speech analogs and no correlations between speech and non-speech were observed.

General discussion

The two experiments in the present study demonstrate strong effects of language experience on speech perception. In both

Experiments, Japanese listeners exhibit reduced sensitivity in discriminating American English /r–l/ (Figs. 1b, 6b and Tables 1, 2 for behavioral data; Figs. 3a, b, c and 7a, b, c for MEG data). However, the results go beyond demonstrating this well-known difficulty. Our findings show that Japanese listeners not only have difficulty discriminating cross-boundary /r–l/ stimuli but they also show significantly better within-category discrimination for /r–l/, which would make /r–l/ categorization very difficult. There is also an asymmetry in within-category discrimination as the cross-linguistic difference was observed only for /l/ but not for /r/. Previous studies have also observed this phenomenon (e.g., Miyawaki et al., 1975) and demonstrated further that, while American listeners focus attention on the critical F3 difference in the /r–l/ stimuli, Japanese listeners essentially ignore F3 and attend instead to F2 variations in the stimuli, ones that are irrelevant to the /r–l/ distinction (Iverson et al., 2003). In contrast, neither the linguistic control stimuli (the /b–w/ of Experiment 1) nor the acoustic control stimuli (the /r–l/ analogs of Experiment 2) showed the same patterns of subject-group differences.

What produces these cross-language processing differences? We attribute these perceptual differences to early linguistic experience (Kuhl, 2000, 2004; Werker and Tees, 1984), and more specifically, to infants’ well-demonstrated sensitivities to the statistical distributions of auditory speech patterns present in language input (Jusczyk, 1997; Kuhl et al., 1997; Maye et al., 2002; Saffran et al., 1996). We argue that early experience results in a *neural commitment* to the frequently occurring patterns in native-language speech (Kuhl, 2000, 2004). For instance, as early as 6 months of age, infants raised in the United States and Sweden in monolingual families already show that they are less sensitive to within-category variations in native-language vowels when contrasted with comparable changes in foreign-language vowels (Kuhl

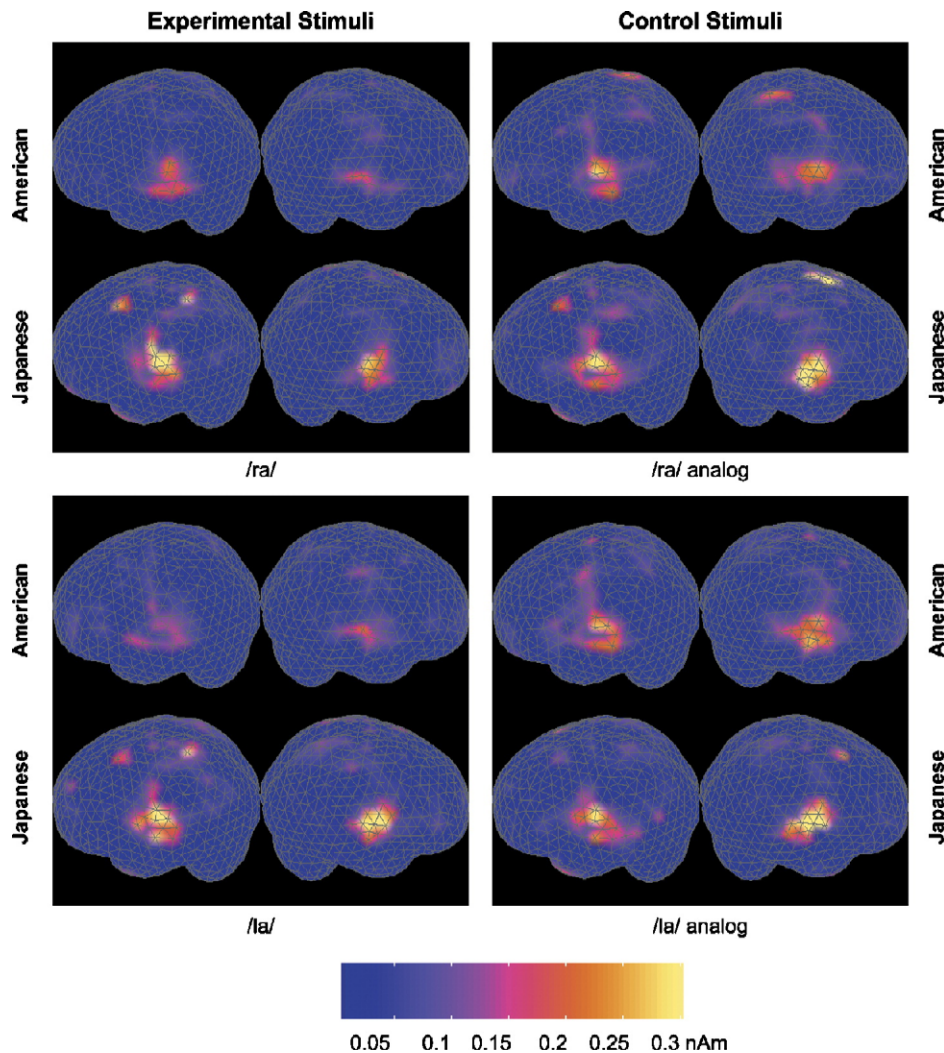


Fig. 9. Minimum current estimates for each stimulus in Experiment 2. Samples between 20 and 700 ms were integrated and averaged for each subject group. Note that there was a smaller sample size in Experiment 2 for each group than in Experiment 1.

et al., 1992). This effect, named the *perceptual magnet effect*, has been demonstrated both in infants (Kuhl, 1991; Kuhl et al., 1992) and adults (Guenther et al., 2004; Iverson and Kuhl, 1996; Iverson et al., 2003) and confirms the effects of native-language experience in warping the mental representations of acoustic space for phonetic categories. Moreover, the results of our study are in line with these behavioral data as well as the accumulated MMN/MMF data (Dehaene-Lambertz, 1997; Dehaene-Lambertz et al., 2000; Näätänen et al., 1997; Nenonen et al., 2003; Rivera-Gaxiola et al., 2000; Sharma and Dorman, 2000; Winkler et al., 1999), showing that the roots of cross-language speech processing differences appear to lie in the sensory (pre-attentive) processing of acoustic distinctions, with experience affecting not only the acoustic dimensions that are attended to but also the temporal window over which the phonetic information is processed.

What evidence in support of the early experience/neural commitment hypothesis can be gleaned from the neural efficiency data? First, the ECD cluster measures suggest that linguistic experience alters the brain's deployment of cortical resources during speech processing. The same /r-l/ speech sounds produced more focal cortical representations with shorter duration of activation in the Americans when compared with the Japanese ((Figs. 3d, e, f, 4, 7d,

e, f, 8); also see Movie 1 in Appendix as supplementary information). The spatial patterns of brain activities showing the cross-language differences were also supported by MCE/MNE results (Figs. 5 and 9). These effects were not seen with either linguistic or acoustic control stimuli. Second, regional analyses from ECD results suggest a neural circuitry that subserves both speech and non-speech perception, consisting of the ST, MT, IP, and IF cortex. Linguistic experience appears to impact ST and IP significantly, where processing non-native speech information involved longer durations of neural activation. The general brain activation patterns are consistent with earlier imaging studies of phonetic perception (Callan et al., 2003, 2004; Golestani and Zatorre, 2004; Guenther et al., 2004; Wang et al., 2003) and of language comprehension as well. In one, word recognition in a second language elicited more diffuse and variable brain activation than in the native language (Dehaene et al., 1997). In another, cochlear-implant users, after language training, showed increased brain activation in the auditory cortex compared to normal subjects (Naito et al., 2000). Our results integrate these findings by suggesting that the deployment of additional brain resources associated with non-native language processing or newly acquired language comprehension ability may have its roots in native-

language learning at the phonetic level, which results in a neural commitment to learned patterns in the absence of explicit attention (Kuhl, 2000, 2004). We suppose that the attainment process of neural commitment early in life may involve fundamental changes in cortical synaptic connections and connectivity in neural pathways (Golestani et al., 2002; Kandel, 2001). It remains to be explored how early language exposure, as characterized by the prevalent use of *motherese* in the vital context of social interaction (Kuhl et al., 1997, 2003), fine tunes the unsupervised self-organizing neural system (de Boer and Kuhl, 2003; Guenther and Gjaja, 1996; McCandliss et al., 2002; Munakata and Pfaffly, 2004), giving rise to the neural commitment to language-specific auditory patterns.

A cautionary note is necessary regarding the interpretation of “neural efficiency.” An alternative account is simply that the observed differences in brain activity may not be due to neural efficiency, but they could reflect rather fundamentally different types of neural processes used by native and non-native speakers. Relatively speaking, the construct of efficiency has a strong intuitive appeal in cognitive neuroscience—higher ability in any domain of intelligence (linguistic, musical, logical–mathematical, spatial, bodily kinesthetic, etc.) means more efficient use of the underlying neural mechanisms (Deary and Caryl, 1997; Näätänen, 1973; Neubauer et al., 2004). However, this notion is not without problems in characterizing the brain activation patterns in association with certain traits of the subjects under study. For instance, smaller activation does not necessarily mean greater efficiency in cases involving deficiency of the neural system (e.g., Brunswick et al., 1999), emotional involvement (e.g., Sanfey et al., 2003), attentional modulation (e.g., Shomstein and Yantis, 2004), or overlearning (e.g., Grabner et al., 2003). We take the position that either the brain as a whole, or particular task-related ROIs, can be considered less efficient if more and longer brain activity is involved in executing the SAME task. Generally, one can compare efficiency across subjects rather than within a subject; if it is within a subject, it would have to be measures at different times for the SAME task—for example, before and after training or at different ages in a longitudinal design to look at developmental changes. Under these specifications, for example, interpreting hemispheric laterality effects in terms of efficiency would be inappropriate because the two hemispheres are parts of the same system that could be engaged in different aspects of the task. What makes “neural efficiency” difficult in some cases is that it does not necessarily correspond to “behavioral efficiency” in a linear sense. Conflicting findings have been reported in a number of studies using linguistic, cognitive, or motor tasks (see Landau et al., 2004; Poldrack, 2000; Van Mier, 2000 for reviews). Changes in brain activation may reflect not only prior learning experience but also the time course of learning, subject characteristics, and the nature of task. It appears that only more empirical facts will allow us to examine the validity of the neural efficiency concept and to develop and test specific ideas of how efficiency of a neural network is related to its activation.

Despite the fact that our stimuli were presented only to the right ear, our study did not show the clear left-hemisphere dominance for speech perception that is documented in the literature (e.g., Breier et al., 1999; Dehaene-Lambertz et al., 2005; Frost et al., 1999; Koyama et al., 2000a; Liégeois-Chauvel et al., 1999; Näätänen et al., 1997). Bilateral activation patterns for speech under a monaural condition have been reported previously (Ackermann et al., 2001; Wong et al., 2002). A number of other

studies also support the idea that speech is processed bilaterally at the phonetic level (Binder et al., 2000; Hickok and Poeppel, 2000; Joanisse and Gati, 2003; Kasai et al., 2001; Mathiak et al., 2000; Scott and Johnsrude, 2003). On the other hand, consistent with many other studies (see Tervaniemi and Hugdahl, 2003 for a review), we observed relatively stronger MMFs for the tonal stimuli in the right hemisphere than in the left in both of our subject groups. Among the major factors, such as subject attributes of gender and handedness (e.g., Salmelin et al., 1999; Tzourio et al., 1998), experimental task (e.g., Szymanski et al., 1999), and stimulus properties (e.g., Kasai et al., 2001; Rinne et al., 1999; Shtyrov et al., 2000b), the assessment of hemispheric laterality may depend on the component of the brain responses and the time window being examined (Hertrich et al., 2002; Papanicolaou et al., 2003; Rinne et al., 2000; Sams et al., 1990; Shtyrov et al., 2000a). Further research is necessary to determine what factors (see Näätänen, 2001; Tervaniemi and Hugdahl, 2003; Tzourio-Mazoyer et al., 2004; Vihla and Salmelin, 2003; Zatorre and Belin, 2001 for discussions), including the attentional demands and the levels of linguistic processing involved, contribute to left hemispheric dominance for speech or the lack thereof.

Collectively, the behavioral and MEG measures in the present study provide substantiating evidence for a neural commitment to the native language that underlies the observed accuracy and efficiency in native as opposed to non-native speech categorization. Our results specifically indicate that this neural commitment reflects the linguistic aspects of the sound patterns in a listener's primary language and thus may serve as a “neural marker” of language learning. We predict that, like the MMN/MMF, the ECD cluster and duration measures should be sensitive to the effects of second-language learning. In other words, neuromagnetic measures taken before and after successful training on non-native phonetic contrasts should show increases in efficiency. Furthermore, these measures should be sensitive to learning during development; as children learn the acoustic cues that reliably distinguish phonemic categories in their language, neural efficiency should systematically increase.

Acknowledgments

Funding was provided by NTT Communication Science Laboratories (Nippon Telegraph and Telephone Corporation), the NIH (HD 37954), the University of Washington's Institute for Learning and Brain Sciences, the Talaris Research Institute, and the Apex Foundation, the family foundation of Bruce and Jolene McCaw. We thank M. Kawakatsu for MRI scanning and D. Gammon for his assistance in data analysis. We also sincerely thank the two anonymous reviewers for their insightful questions and comments.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.neuroimage.2005.02.040](https://doi.org/10.1016/j.neuroimage.2005.02.040).

References

- Ackermann, H., Hertrich, I., Mathiak, K., Lutzenberger, W., 2001. Contralaterality of cortical auditory processing at the level of the M50/M100 complex and the mismatch field: a whole-head magnetoencephalography study. *NeuroReport* 12, 1683–1687.

- Binder, J.R., Frost, J.A., Hammeke, T.A., Bellgowan, P.S., Springer, J.A., Kaufman, J.N., Possing, E.T., 2000. Human temporal lobe activation by speech and nonspeech sounds. *Cereb. Cortex* 10, 512–528.
- Bradlow, A., Ahakane-Yamada, R., Pisoni, D., Tokhura, Y., 1999. Training Japanese listeners to identify English /r/ and /l/: long-term retention of learning in perception and production. *Percept. Psychophys.* 61, 977–987.
- Breier, J.I., Simos, P.G., Zouridakis, G., Papanicolaou, A.C., 1999. Lateralization of cerebral activation in auditory verbal and non-verbal memory tasks using magnetoencephalography. *Brain Topogr.* 12, 89–97.
- Breier, J.I., Simos, P.G., Zouridakis, G., Papanicolaou, A.C., 2000. Lateralization of activity associated with language function using magnetoencephalography: a reliability study. *J. Clin. Neurophysiol.* 17, 503–510.
- Brunswick, N., McCrory, E., Price, C.J., Frith, C.D., Frith, U., 1999. Explicit and implicit processing of words and pseudowords by adult developmental dyslexics: a search for Wernicke's Wortschatz? *Brain* 122, 1901–1917.
- Callan, D.E., Tajima, K., Callan, A.M., Kubo, R., Masaki, S., Akahane-Yamada, R., 2003. Learning-induced neural plasticity associated with improved identification performance after training of a difficult second-language phonetic contrast. *NeuroImage* 19, 113–124.
- Callan, D.E., Jones, J.A., Callan, A.M., Akahane-Yamada, R., 2004. Phonetic perceptual identification by native- and second-language speakers differentially activates brain regions involved with acoustic phonetic processing and those involved with articulatory-auditory/orosensory internal models. *NeuroImage* 22, 1182–1194.
- Cheour, M., Ceboniene, R., Lehtokoski, A., Luuk, A., Allik, J., Alho, K., Näätänen, R., 1998. Development of language-specific phoneme representations in the infant brain. *Nat. Neurosci.* 1, 351–353.
- Cowan, N., 1984. On short and long auditory stores. *Psychol. Bull.* 96, 341–370.
- Deary, I.J., Caryl, P.G., 1997. Neuroscience and human intelligence differences. *Trends Neurosci.* 20, 365–371.
- de Boer, B., Kuhl, P.K., 2003. Investigating the role of infant-directed speech with a computer model. *Acoust. Res. Lett. Online* 4, 129–134.
- Dehaene, S., Dupoux, E., Mehler, J., Cohen, L., Paulesu, E., Perani, D., van de Moortele, P., Lehericy, S., Le Bihan, D., 1997. Anatomical variability in the cortical representation of first and second language. *NeuroReport* 8, 3809–3815.
- Dehaene-Lambertz, G., 1997. Electrophysiological correlates of categorical phoneme perception in adults. *NeuroReport* 8, 919–924.
- Dehaene-Lambertz, G., Dupoux, E., Gout, A., 2000. Electrophysiological correlates of phonological processing: a cross-linguistic study. *J. Cogn. Neurosci.* 12, 635–647.
- Dehaene-Lambertz, G., Pallier, C., Serniclaes, W., Sprenger-Charolles, L., Jobert, A., Dehaene, S., 2005. Neural correlates of switching from auditory to speech perception. *NeuroImage* 24, 21–33.
- Ertl, J.P., Schafer, E.W., 1969. Brain response correlates of psychometric intelligence. *Nature* 223, 421–422.
- Fox, R.A., Flége, J.E., Munro, M.J., 1995. The perception of English and Spanish vowels by native English and Spanish listeners: a multidimensional scaling analysis. *J. Acoust. Soc. Am.* 97, 2540–2550.
- Frost, J.A., Binder, J.R., Springer, J.A., Hammeke, T.A., Bellgowan, P.S.F., Rao, S.M., Cox, R.W., 1999. Language processing is strongly left lateralized in both sexes: evidence from functional MRI. *Brain* 122, 199–208.
- Fujimaki, N., Hayakawa, T., Nielsen, M., Knosche, T.R., Miyauchi, S., 2002. An fMRI-constrained MEG source analysis with procedures for dividing and grouping activation. *NeuroImage* 17, 324–343.
- Golestani, N., Zatorre, R.J., 2004. Learning new sounds of speech: reallocation of neural substrates. *NeuroImage* 21, 494–506.
- Golestani, N., Paus, T., Zatorre, R.J., 2002. Anatomical correlates of learning novel speech sounds. *Neuron* 35, 997–1010.
- Goto, H., 1971. Auditory perception by normal Japanese adults of the sounds 'l' and 'r'. *Neuropsychologia* 9, 317–323.
- Grabner, R.H., Stern, E., Neubauer, A.C., 2003. When intelligence loses its impact: neural efficiency during reasoning in a familiar area. *Int. J. Psychophysiol.* 49, 89–98.
- Guenther, F.H., Gjaja, M.N., 1996. The perceptual magnet effect as an emergent property of neural map formation. *J. Acoust. Soc. Am.* 100, 1111–1121.
- Guenther, F.H., Nieto-Castanon, A., Ghosh, S.S., Tourville, J.A., 2004. Representation of sound categories in auditory cortical maps. *J. Speech Lang. Hear. Res.* 47, 46–57.
- Guion, S.G., Flége, J.E., Akahane-Yamada, R., Pruitt, J.C., 2000. An investigation of current models of second language speech perception: the case of Japanese adults' perception of English consonants. *J. Acoust. Soc. Am.* 107, 2711–2724.
- Hämäläinen, M.S., Ilmoniemi, R.J., 1984. Interpreting measured magnetic fields of the brain: estimates of current distributions. Technical Report TKK-F-A559, Helsinki University of Technology, Espoo.
- Hämäläinen, M.S., Ilmoniemi, R.J., 1994. Interpreting magnetic fields of the brain: minimum norm estimates. *Med. Biol. Eng. Comput.* 32, 35–42.
- Hämäläinen, M.S., Hari, R., Ilmoniemi, R.J., Knuutila, J., Lounasmaa, O.V., 1993. Magnetoencephalography—theory, instrumentation, and applications to noninvasive studies of the working human brain. *Rev. Mod. Phys.* 65, 413–497.
- Hari, R., Joutsiniemi, S., Sarvas, J., 1988. Spatial resolution of neuro-magnetic records: theoretical calculations in a spherical model. *Electroencephalogr. Clin. Neurophysiol.* 71, 64–72.
- Hari, R., Levänen, S., Raij, T., 2000. Timing of human cortical functions during cognition: role of MEG. *Trends Cogn. Sci.* 4, 455–462.
- Haslinger, B., Erhard, P., Altenmüller, E., Hennenlotter, A., Schwaiger, M., Graf von Einsiedel, H., Rummeny, E., Conrad, B., Ceballos-Baumann, A.O., 2004. Reduced recruitment of motor association areas during bimanual coordination in concert pianists. *Hum. Brain Mapp.* 22, 206–215.
- Hebb, D.O., 1949. *The Organization of Behavior*. Wiley, New York.
- Hellstrand, E., Abraham-Fuchs, K., Jernberg, B., Kihlstrom, L., Knutsson, E., Lindquist, C., Schneider, S., Wirth, A., 1993. MEG localization of interictal epileptic focal activity and concomitant stereotactic radiosurgery. A non-invasive approach for patients with focal epilepsy. *Physiol. Meas.* 14, 131–136.
- Hertrich, I., Mathiak, K., Lutzenberger, W., Ackermann, H., 2002. Hemispheric lateralization of the processing of consonant-vowel syllables (formant transitions): effects of stimulus characteristics and attentional demands on evoked magnetic fields. *Neuropsychologia* 40, 1902–1917.
- Hickok, G., Poeppel, D., 2000. Towards a functional neuroanatomy of speech perception. *Trends Cogn. Sci.* 4, 131–138.
- Hsieh, L., Gandour, J., Wong, D., Hutchins, G.D., 2001. Functional circuits engaged in speech perception depend on language experience. *Brain Lang.* 76, 227–252.
- Imada, T., Mashiko, T., Sekine, H., 2001. Estimation errors of single dipole model applied to twin dipole activity: computer simulation study. In: Nenonen, J., Ilmoniemi, R.J., Katila, T. (Eds.), *Proceedings of the 12th International Conference on Biomagnetism*. Helsinki University of Technology, Espoo, Finland, pp. 733–737.
- Inouchi, M., Kubota, M., Ferrari, P., Roberts, T.P., 2003. Magnetic mismatch fields elicited by vowel duration and pitch changes in Japanese words in humans: comparison between native- and non-speakers of Japanese. *Neurosci. Lett.* 353, 165–168.
- Iverson, P., Kuhl, P.K., 1996. Influences of phonetic identification and category goodness on American listener's perception of /r/ and /l/. *J. Acoust. Soc. Am.* 99, 1130–1140.
- Iverson, P., Kuhl, P.K., Akahane-Yamada, R., Diesch, E., Tohkura, Y., Kettermann, A., Siebert, C., 2003. A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition* 87, B47–B57.
- Jaušovec, N., Jaušovec, K., 2001. Differences in EEG current density related to intelligence. *Cogn. Brain Res.* 12, 55–60.
- Jensen, O., Vanni, S., 2002. A new method to identify multiple sources of

- oscillatory activity from magnetoencephalographic data. *NeuroImage* 15, 568–574.
- Joanisse, M.F., Gati, J.S., 2003. Overlapping neural regions for processing rapidtemporal cues in speech and nonspeech signals. *NeuroImage* 19, 64–79.
- Jongman, A., Moore, C., 2000. The role of language experience in speaker and rate normalization processes. *Proceedings of the 6th International Conference on Spoken Language Processing*, vol. 1, pp. 62–65.
- Jusczyk, P.W., 1997. *The Discovery of Spoken Language*. MIT Press, Cambridge, MA.
- Kandel, E.R., 2001. The molecular biology of memory storage: a dialogue between genes and synapses. *Science* 294, 1030–1038.
- Kasai, K., Yamada, H., Kamio, S., Nakagome, K., Iwanami, A., Fukuda, M., Itoh, K., Koshida, I., Yumoto, M., Iramina, K., Kato, N., Ueno, S., 2001. Brain lateralization for mismatch response to across- and within-category change of vowels. *NeuroReport* 12, 2467–2471.
- Komssi, S., Huttunen, J., Aronen, H.J., Ilmoniemi, R.J., 2004. EEG minimum-norm estimation compared with MEG dipole fitting in the localization of somatosensory sources at S1. *Clin. Neurophysiol.* 115, 534–542.
- Koyama, S., Gunji, A., Yabe, H., Oiwa, S., Akahane-Yamada, R., Kakigi, R., Näätänen, R., 2000a. Hemispheric lateralization in an analysis of speech sounds. Left hemisphere dominance replicated in Japanese subjects. *Brain Res., Cogn. Brain Res.* 10, 119–124.
- Koyama, S., Gunji, A., Yabe, H., Yamada, R.A., Oiwa, S., Kubo, R., Kakigi, R., 2000b. The masking effect in foreign speech sounds perception revealed by neuromagnetic responses. *NeuroReport* 11, 3765–3769.
- Koyama, S., Akahane-Yamada, R., Gunji, A., Kubo, R., Roberts, T.P., Yabe, H., Kakigi, R., 2003. Cortical evidence of the perceptual backward masking effect on /l/ and /r/ sounds from a following vowel in Japanese speakers. *NeuroImage* 18, 962–974.
- Kraus, N., McGee, T.J., Carrell, T.D., Zecker, S.G., Nicol, T.G., Koch, D.B., 1996. Auditory neurophysiological responses and discrimination deficits in children with learning problems. *Science* 273, 971–973.
- Kuhl, P.K., 1991. Human adults and human infants show a “perceptual magnet effect” for the prototypes of speech categories, monkeys do not. *Percept. Psychophys.* 50, 93–107.
- Kuhl, P.K., 2000. Language, mind, and brain: experience alters perception. In: Gazzaniga, M.S. (Ed.), *The New Cognitive Neurosciences*, 2nd ed. MIT Press, Cambridge, MA, pp. 99–115.
- Kuhl, P.K., 2004. Early language acquisition: cracking the speech code. *Nat. Rev., Neurosci.* 5, 831–843.
- Kuhl, P.K., Williams, K.A., Lacerda, F., Stevens, K.N., Lindblom, B., 1992. Linguistic experience alters phonetic perception in infants by 6 months of age. *Science* 255, 606–608.
- Kuhl, P.K., Andruski, J.E., Chistovich, I.A., Chistovich, L.A., Kozhevnikova, E.V., Ryskina, V.L., Stolyarova, E.I., Sundberg, U., Lacerda, F., 1997. Cross-language analysis of phonetic units in language addressed to infants. *Science* 277, 684–686.
- Kuhl, P.K., Tsao, F.M., Liu, H.M., Zhang, Y., de Boer, B., 2001. Language/culture/mind/brain: progress at the margins between disciplines. *Ann. N. Y. Acad. Sci.* 935, 136–174.
- Kuhl, P.K., Tsao, F.M., Liu, H.M., 2003. Foreign-language experience in infancy: effects of short-term exposure and social interaction on phonetic learning. *Proc. Natl. Acad. Sci. U. S. A.* 100, 9096–9101.
- Landau, S.M., Schumacher, E.H., Garavan, H., Druzgal, T.J., D’Esposito, M., 2004. A functional MRI study of the influence of practice on component processes of working memory. *NeuroImage* 22, 211–221.
- Lewine, J.D., Davis, J.T., Sloan, J.H., Koditwakk, P.W., Orrison Jr., W.W., 1999. Neuromagnetic assessment of pathophysiologic brain activity induced by minor head trauma. *Am. J. Neuroradiol.* 20, 857–866.
- Liberman, A., Harris, K., Hoffman, B., Griffith, K., 1957. The discrimination of speech sounds within and across phoneme boundaries. *J. Exp. Psychol.* 54, 358–368.
- Liégeois-Chauvel, C., de Graaf, J.B., Laguitton, V., Chauvel, P., 1999. Specialization of left auditory cortex for speech perception in man depends on temporal coding. *Cereb. Cortex* 9, 484–496.
- Liu, A.K., Belliveau, J.W., Dale, A.M., 1998. Spatiotemporal imaging of human brain activity using functional MRI constrained magnetoencephalography data: Monte Carlo simulations. *Proc. Natl. Acad. Sci. U. S. A.* 95, 8945–8950.
- Loveless, N.E., Hari, R., 1993. Auditory evoked fields covary with perceptual grouping. *Biol. Psychol.* 35, 1–15.
- Lütkenhöner, B., 1998. Dipole separability in a neuromagnetic source analysis. *IEEE Trans. Biomed. Eng.* 45, 572–581.
- Lütkenhöner, B., Lehnertz, K., Hoke, M., Pantev, C., 1991. On the biomagnetic inverse problem in the case of multiple dipoles. *Acta Oto-Laryngol., Suppl.* 491, 94–104.
- Maestú, F., Ortiz, T., Fernandez, A., Amo, C., Martin, P., Fernández, S., Sola, R.G., 2002. Spanish language mapping using MEG: a validation study. *NeuroImage* 17, 1579–1586.
- Mathiak, K., Hertrich, I., Lutzenberger, W., Ackermann, H., 2000. Encoding of temporal speech features (formant transients) during binaural and dichotic stimulus application: a whole-head magnetoencephalography study. *Brain Res., Cogn. Brain Res.* 10, 125–131.
- Maye, J., Werker, J.F., Gerken, L., 2002. Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition* 82, B101–B111.
- McClelland, 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. Affect. Behav. Neurosci.* 2, 89–108.
- Menning, H., Imaizumi, S., Zwitserlood, P., Pantev, C., 2002. Plasticity of the human auditory cortex induced by discrimination learning of non-native, mora-timed contrasts of the Japanese language. *Learn. Mem.* 9, 253–267.
- Miller, J.L., Liberman, A.M., 1979. Some effects of later-occurring information on the perception of stop consonant and semivowel. *Percept. Psychophys.* 25, 457–465.
- Minagawa-Kawai, Y., Mori, K., Sato, Y., Koizumi, T., 2004. Differential cortical responses in second language learners to different vowel contrasts. *NeuroReport* 15, 899–903.
- Miyawaki, K., Strange, W., Verbrugge, R., Liberman, A., Jenkins, J.J., Fujimura, O., 1975. An effect of linguistic experience: the discrimination of [r] and [l] by native speakers of Japanese and English. *Percept. Psychophys.* 18, 331–340.
- Munakata, Y., Pfaffly, J., 2004. Hebbian learning and development. *Dev. Sci.* 7, 141–148.
- Näätänen, R., 1973. The inverted-U relationship between activation and performance: a critical review. In: Kornblum, S. (Ed.), *Attention and Performance IV*. Academic Press, New York, pp. 155–174.
- Näätänen, R., 1992. *Attention and Brain Function*. Erlbaum, Hillsdale, NJ.
- Näätänen, R., 2001. The perception of speech sounds by the human brain as reflected by the mismatch negativity (MMN) and its magnetic equivalent (MMNm). *Psychophysiology* 38, 1–21.
- Näätänen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huottilainen, M., Iivonen, A., Valnio, A., Alku, P., Ilmoniemi, R.J., Luuk, A., Allik, J., Sinkkonen, J., Alho, K., 1997. Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature* 385, 432–434.
- Naito, Y., Tateya, I., Fujiki, N., Hirano, S., Ishizu, K., Nagahama, Y., Fukuyama, H., Kojima, H., 2000. Increased cortical activation during hearing of speech in cochlear implant users. *Hear. Res.* 143, 139–146.
- Neenen, S., Shestakova, A., Huottilainen, M., Näätänen, R., 2003. Linguistic relevance of duration within the native language determines the accuracy of speech-sound duration processing. *Cogn. Brain Res.* 16, 492–495.
- Neubauer, A.C., Grabner, R.H., Freudenthaler, H.H., Beckmann, J.F., Guthke, J., 2004. Intelligence and individual differences in becoming neurally efficient. *Acta Psychol.* 116, 55–74.
- Pallier, C., Bosch, L., Sebastián-Gallés, N., 1997. A limit on behavioral plasticity in vowel acquisition. *Cognition* 64, B9–B17.

- Papanicolaou, A.C., Simos, P.G., Castillo, E., Breier, J.I., Davis, R.N., Diehl, R.L., 2003. Differential brain activation patterns during perception of voice and tone onset time series: a MEG study. *NeuroImage* 18, 448–459.
- Patarelia, E., Simos, P.G., Castillo, E.M., Billingsley-Marshall, R.L., McGregor, A.L., Breier, J.I., Sarkari, S., Papanicolaou, A.C., 2004. Reorganization of language-specific cortex in patients with lesions or mesial temporal epilepsy. *Neurology* 63, 1825–1832.
- Phillips, C., Marantz, A., McGinnis, M., Pesetsky, D., Wexler, K., Yellin, A., Poeppel, D., Roberts, T., Rowley, H., 1995. Brain mechanisms of speech perception: a preliminary report. MIT Work. Pap. Linguist. 26, 153–192.
- Poldrack, R., 2000. Imaging brain plasticity: conceptual and methodological issues—A theoretical review. *NeuroImage* 12, 1–13.
- Rinne, T., Alho, K., Alku, P., Holli, M., Sinkkonen, J., Virtanen, J., Bertrand, O., Näätänen, R., 1999. Analysis of speech sounds is left-hemisphere predominant at 100–150 ms after sound onset. *NeuroReport* 10, 1113–1117.
- Rinne, T., Alho, K., Alku, P., Holli, M., Sinkkonen, J., Virtanen, J., Bertrand, O., Näätänen, R., 2000. Analysis of speech sounds is left-hemisphere predominant at 100–150 ms after sound onset. *NeuroReport* 10, 1113–1117.
- Rivera-Gaxiola, M., Csibra, G., Johnson, M.H., Karmiloff-Smith, A., 2000. Electrophysiological correlates of cross-linguistic speech perception in native English speakers. *Behav. Brain Res.* 112, 1–11.
- Rivera-Gaxiola, M., Silva-Pereyra, J., Kuhl, P.K., 2005. Brain potentials to native and non-native speech contrasts in 7- and 11-month-old American infants. *Dev. Sci.* 8, 162–172.
- Saffran, J.R., Aslin, R.N., Newport, E.L., 1996. Statistical learning by 8-month-old infants. *Science* 274, 1926–1928.
- Salmelin, R., Schnitzler, A., Parkkonen, L., Biermann, K., Helenius, P., Kiviniemi, K., Kuikka, K., Schmitz, F., Freund, H.-J., 1999. Native language, gender, and functional organization of the auditory cortex. *Proc. Natl. Acad. Sci. U. S. A.* 96, 10460–10465.
- Sams, M., Aulanko, R., Aaltonen, O., Näätänen, R., 1990. Event-related potentials to infrequent changes in synthesized phonetic stimuli. *J. Cogn. Neurosci.* 2, 344–357.
- Sanfey, A.G., Rilling, J.K., Aronson, J.A., Nystrom, L.E., Cohen, J.D., 2003. The neural basis of economic decision-making in the Ultimatum Game. *Science* 300, 1755–1758.
- Scott, S., Johnsrude, I., 2003. The neuroanatomical and functional organization of speech perception. *Trends Neurosci.* 26, 100–107.
- Sharma, A., Dorman, M.F., 2000. Neurophysiologic correlates of cross-language phonetic perception. *J. Acoust. Soc. Am.* 107, 2697–2703.
- Shomstein, S., Yantis, S., 2004. Control of attention shifts between vision and audition in human cortex. *J. Neurosci.* 24, 10702–10706.
- Shtyrov, Y., Kujala, T., Lyytinen, H., Ilmoniemi, R.J., Näätänen, R., 2000a. Auditory cortex evoked magnetic fields and lateralization of speech processing. *NeuroReport* 11, 2893–2896.
- Shtyrov, Y., Kujala, T., Palva, S., Ilmoniemi, R.J., Näätänen, R., 2000b. Discrimination of speech and complex nonspeech sounds of different temporal structure in the left and right cerebral hemispheres. *NeuroImage* 12, 657–663.
- Simos, P.G., Breier, J.I., Fletcher, J.M., Bergman, E., Papanicolaou, A.C., 2000. Cerebral mechanisms involved in word reading in dyslexic children: a magnetic source imaging approach. *Cereb. Cortex* 10, 809–816.
- Singh, K.D., Ioannides, A.A., Gray, N., Kober, H., Pongratz, H., Daun, A., Grummich, P., Vieth, J., 1994. Distributed current analyses of bi-hemispheric magnetic N1m responses to ipsi/contralateral monaural stimuli from a single subject. *Electroencephalogr. Clin. Neurophysiol.* 92, 365–368.
- Stenbacka, L., Vanni, S., Uutela, K., Hari, R., 2002. Comparison of minimum current estimate and dipole modeling in the analysis of simulated activity in the human visual cortices. *NeuroImage* 16, 936–943.
- Stevens, S.S., Volkman, J., Newman, E.B., 1937. The Mel scale equates the magnitude of perceived differences in pitch at different frequencies. *J. Acoust. Soc. Am.* 8, 185–190.
- Strange, W. (Ed.), 1995. *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*. York Press, Timonium, MD.
- Szymanski, M.D., Yund, E.W., Woods, D.L., 1999. Phonemes, intensity and attention: differential effects on the mismatch negativity (MMN). *J. Acoust. Soc. Am.* 106, 3492–3505.
- Tervaniemi, M., Hugdahl, K., 2003. Lateralization of auditory-cortex functions. *Brain Res. Rev.* 43, 231–246.
- Tremblay, K., Kraus, N., Carrell, T.D., McGee, T., 1997. Central auditory system plasticity: generalization to novel stimuli following listening training. *J. Acoust. Soc. Am.* 102, 3762–3773.
- Tzourio, N., Crivello, F., Mellet, E., Nkanga-Ngila, B., Mazoyer, B., 1998. Functional anatomy of dominance for speech comprehension in left handers vs right handers. *NeuroImage* 8, 1–16.
- Tzourio-Mazoyer, N., Josse, G., Crivello, F., Mazoyer, B., 2004. Interindividual variability in the hemispheric organization for speech. *NeuroImage* 21, 422–435.
- Uutela, K., Hämäläinen, M., Somersalo, E., 1999. Visualization of magnetoencephalographic data using minimum current estimates. *NeuroImage* 10, 173–180.
- Van Mier, H., 2000. Human learning. In: Toga, A., Mazziotta, J. (Eds.), *Brain Mapping: The Systems*. Academic Press, San Diego, pp. 605–620.
- Vihla, M., Salmelin, R., 2003. Hemispheric balance in processing attended and non-attended vowels and complex tones. *Cogn. Brain Res.* 16, 167–173.
- Vihla, M., Kiviniemi, K., Salmelin, R., 2002. Auditory cortical activation in Finnish and Swedish speaking Finns: a magnetoencephalographic study. *Neurosci. Lett.* 322, 141–144.
- Wang, Y., Sereno, J.A., Jongman, A., Hirsch, J., 2003. fMRI evidence for cortical modification during learning of mandarin lexical tone. *J. Cogn. Neurosci.* 15, 1019–1027.
- Werker, J.F., Tees, R.C., 1984. Cross-language speech perception: evidence for perceptual reorganization during the first year of life. *Inf. Behav. Dev.* 7, 49–63.
- Werker, J.F., Tees, R.C., 1999. Influences on infant speech processing: toward a new synthesis. *Annu. Rev. Psychol.* 50, 509–535.
- Winkler, I., Czigler, I., Jaramillo, M., Paavilainen, P., Näätänen, R., 1998. Temporal constraints of auditory event synthesis: evidence from ERP. *NeuroReport* 9, 495–499.
- Winkler, I., Kujala, T., Tiitinen, H., Sivonen, P., Alku, P., Lehtokoski, A., Czigler, I., Csépe, V., Ilmoniemi, R.J., Näätänen, R., 1999. Brain responses reveal the learning of foreign language phonemes. *Psychophysiology* 36, 638–642.
- Wong, D., Pisoni, D.B., Learn, J., Gandour, J.T., Miyamoto, R.T., Hutchins, G.D., 2002. PET imaging of differential cortical activation to monaural speech and nonspeech stimuli. *Hear. Res.* 166, 9–23.
- Yabe, H., Tervaniemi, M., Sinkkonen, J., Huottilainen, M., Ilmoniemi, R.J., Näätänen, R., 1998. Temporal window of integration of auditory information in the human brain. *Psychophysiology* 35, 615–619.
- Zatorre, R.J., Belin, P., 2001. Spectral and temporal processing in human auditory cortex. *Cereb. Cortex* 11, 946–953.
- Zhang, Y., Kuhl, P.K., Imada, T., Kotani, M., Stevens, E., 2001. MMFs to native and nonnative syllables show an effect of linguistic experience. In: Nenonen, J., Ilmoniemi, R.J., Katila, T. (Eds.), *Proceedings of the 12th International Conference on Biomagnetism*. Helsinki University of Technology, Espoo, Finland, pp. 379–382.