

# Cerebral Specialization for Speech and Non-Speech Stimuli in Infants

G. Dehaene-Lambertz

Centre Hospitalier Universitaire Bicêtre (AP-HP) and CNRS, Paris

## Abstract

■ Early cerebral specialization and lateralization for auditory processing in 4-month-old infants was studied by recording high-density evoked potentials to acoustical and phonetic changes in a series of repeated stimuli (either tones or syllables). Mismatch responses to these stimuli exhibit a distinct topography suggesting that different neural networks within the temporal lobe are involved in the perception and representation of the different features of an auditory stimulus.

These data confirm that specialized modules are present within the auditory cortex very early in development. However, both for syllables and continuous tones, higher voltages were recorded over the left hemisphere than over the right with no significant interaction of hemisphere by type of stimuli. This suggests that there is no greater left hemisphere involvement in phonetic processing than in acoustic processing during the first months of life. ■

## INTRODUCTION

Adults' cognitive functions rely on specialized neural networks. In the auditory domain, in particular, different neural networks within the temporal lobe are involved in the representation of the different features of a sound. Giard et al. (1995) have shown that the intensity, frequency, and duration of a sound are separately and automatically coded in sensory memory: Different mismatch responses are thus generated depending on the deviant characteristics in the next sound heard as compared to the stored representations. Phoneme perception presents distinctive characteristics within auditory perception, as for example, categorical perception versus continuous perception. Whether a phonetic representation is calculated immediately in sensory memory or whether categorical perception is the result of a second processing stage, calculated after the acoustical deviance between stimuli is first coded, has been much debated. It was recently demonstrated that a phonetic representation is indeed calculated in addition to that for the acoustical features of the stimulus and that this phonetic representation depends on the subjects' native language. Dehaene-Lambertz (1997) recorded event-related potentials to synthetic syllables varying along a voiced place of articulation continuum. She showed that for an equal acoustical distance between the deviant and the previous syllables, the mismatch response (or MMN) was dependent on the linguistic relevance of the syllable change in the subjects' native language. For example, she observed no MMN for a Hindi phonetic contrast in French adults while a MMN was present for a similar acoustical change that crossed a French phonetic

boundary. Näätänen et al. (1997) demonstrated that, even when the acoustical distance between a deviant vowel and the standard was progressively increased, the MMN was significantly reduced when adults heard an acoustic change that did not signal a phoneme boundary relevant to their native language, as compared to a change that crossed a native phoneme boundary.

In order to study the developmental course of these specialized network, a first step would be to determine whether the auditory cortex is organized into functional networks at an early age. In 2-month-old infants, Dehaene-Lambertz and Dehaene (1994) have shown that the introduction of a deviant syllable after a succession of four identical syllables induces a mismatch response, which originated in the temporal lobes, around 400 msec after the syllable-onset. Does this mismatch response reflect the operation of a nonspecific "change detector" or are different mismatch responses elicited depending on the deviant parameter, as it is the case with adults?

One of the most evident proofs of brain specialization, and the first historical demonstration, is the left-hemisphere specialization for language in most right-handed adults (Broca, 1861). Recent papers in brain imaging have confirmed a century of neuropsychological data, showing left predominant activations in normal adults processing their native language (Dehaene et al., 1997; Kim, Relkin, Lee, & Hirsch, 1997; Mazoyer et al., 1993). The fact that language can develop to a normal level after an early left-brain lesion (Muter, Taylor, & Vargha-Khadem, 1997; Woods & Carey, 1979) has led some authors to postulate that the brain is initially equipoten-

tial and that the left advantage for processing language develops along with the acquisition of language (Lenneberg, 1967). However, the infant brain, and even the fetal brain during the last months of pregnancy, shows anatomical asymmetries similar to those described in adults, notably in the temporal lobes (Chi, Dooling, & Gilles, 1977; Witelson & Pallie, 1973). Anatomical asymmetry does not necessarily mean functional asymmetry. Several behavioral experiments have tried to address this point by using dichotic listening in infants. Bertoni et al. (1989), Entus (1977), and Glanville, Best, and Levenson (1977) have found an advantage in favor of the right ear in neonates and 2-month-old infants detecting a phonetic change, and one in favor of the left ear when they detect a change in musical notes. However, these results were neither replicated by Vargha-Khadem and Corballis (1979), nor by Best, Hofman and Glanville (1982). These contradictory results may be due to the combination of the difficulties of behavioral experiments in infants and of the lack of sensitivity of dichotic listening. Even in adults, the right-ear advantage for phoneme perception is variable across studies, from 60% to 80% of right handers (Ahoniska, Cantell, Tolvanen, & Lyytinen, 1993).

Electrophysiological studies are equally inconclusive. Molfese and Burger-Judish (1991) observed larger event-related potentials (ERPs) on a right-temporal electrode compared to the symmetrical left electrode for a phoneme change along the VOT dimension, and a reverse advantage for a change along the place of articulation dimension. Novak, Kurtzberg, Kreuzer, and Vaughan (1989) have recorded a higher ERP amplitude to syllables on a central-right electrode compared to the left one. The number of electrodes in these studies was very small making it difficult to distinguish between a genuine effect and a spurious artifactual effect. Lastly, Dehaene-Lambertz and Dehaene (1994) and Dehaene-Lambertz and Baillet (1998), using a high-density geodesic net of 65 electrodes, recorded higher voltages over the left hemisphere than over the right one when 2- and 3-month-old infants were listening to syllables. However, as noticed by the authors, this asymmetry could either be due to asymmetries in brain morphology, for instance, in the orientation of the left- and right-sylvian fissures, or to a genuine advantage in favor of the left hemisphere in phonetic processing. To determine whether there is an early left specialization to process speech stimuli, it is necessary to contrast speech and nonspeech stimuli within the same subjects.

To study these two questions—is the auditory cortex organized into functional networks at an early age, and is there any evidence of early left–right asymmetries in auditory processing especially in speech processing—high-density ERPs to tones and syllables were recorded in 4-month-old infants. Although ERPs do not allow one to localize all the cerebral regions activated during a

cognitive task precisely, different voltage cartographies indicate that the electrical response recorded on the scalp originates in the activation of distinct neural networks (Giard et al., 1995). Therefore, in both experiments presented in this paper, we used a high-density geodesic net with 65 electrodes, which allows a precise mapping of voltage over the entire scalp.

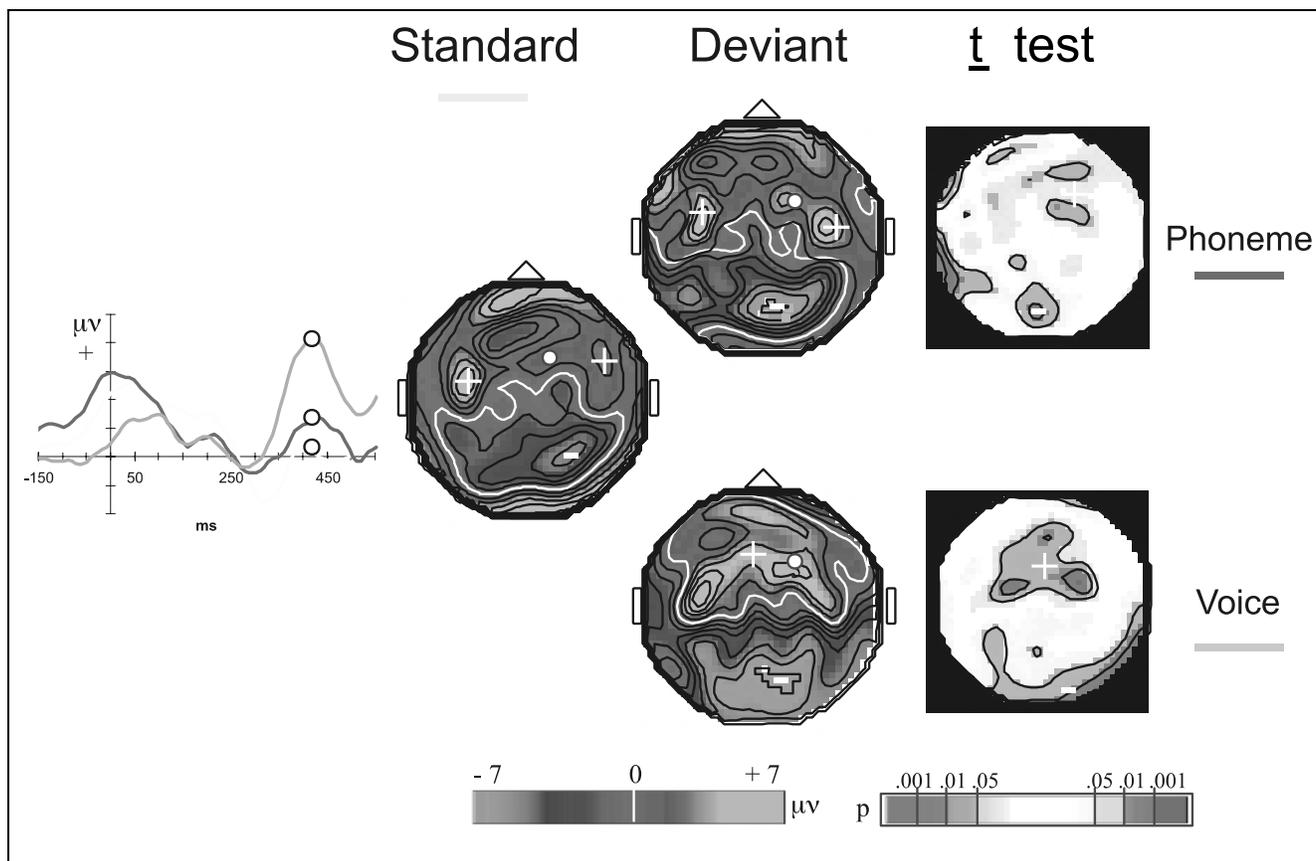
Stimuli were presented by blocks of four, the last stimulus being or not similar to the three previous ones. In the first experiment, the responses to a voice change and to a phoneme change were compared. To perceive a voice change, processing of pitch and timbre is essential. The processing of these two parameters is regularly attributed to the right hemisphere in studies in normal adults (Zatorre, Evans, Meyer, & Gjedde, 1992) or in brain-damaged patients (Chobor & Brown, 1987). Yet, studies of voice perception in normal adults with dichotic presentation of the stimuli do not find always a left-ear advantage suggesting a right-hemisphere superiority to process voices. The observation of this asymmetry varies depending on the task (discrimination or recognition) or on the familiarity of the subject with the voice (Kreiman & Lancker, 1988). Nevertheless, if the left-hemisphere advantage described for syllable perception is related to pure phonetic processing, it should be weaker for a voice change than for a phoneme change.

In the second experiment, a more drastic comparison was done: A timbre change in tones was opposed to the same phoneme change than in experiment 1. Some authors have attributed the asymmetry observed in phonetic processing to the presence of rapid transitions in consonants, the left hemisphere being more precise in the coding of temporal information (Johnsrude, Zatorre, Milner, & Evanc, 1997). Therefore, in order to avoid any processing favoring the left hemisphere, the tones opposed to syllables were continuous. The change in timbre was obtained by changing the number of harmonics (a spectral cue) and not by modifying the attack (a time cue). It was expected that the perception of these tones should induce at least a symmetrical response if not a right-hemispheric advantage.

## RESULTS

### Experiment 1

2-D temporal animations of the grand-average during the entire trial were examined: As previously described (Dehaene-Lambertz & Dehaene, 1994), the evoked response to each syllable comprised two peaks (maxima, respectively, at 258 and 402 msec postsyllable onset) that decreased in amplitude with repetition.<sup>1</sup> The introduction of a new syllable induced a recovery of the amplitude of peak 2, but not of peak 1, relative to the control trials. The voice discrimination response was longer in duration and more diffuse on the scalp than the phoneme discrimination response. The topography



**Figure 1.** Grand-averaged responses to the last syllable of the trials (S4) in experiment 1. Left: ERP from a fronto-central electrode (○ on maps). Right: maps of evoked responses to standard syllables, phonetic deviant and voice deviants at 400 msec following stimulus onset (● on ERP). Rightmost column: maps of statistical significance ( $t$  test) of deviant versus standard stimulus at the same time.

of both peaks showed an anterior positivity synchronous with a posterior negativity (Figure 1), compatible with previous descriptions that suggested temporal dipoles at the origin of these peaks (Dehaene-Lambertz & Baillet, 1998). However, the negativity was more posterior over the occipital regions for the voice change than for the phoneme change. After a 600-msec postsyllable onset, the two changes induced a similarly slow central negativity. Because this late response is probably an amodal response due to orientation to novelty (Dehaene-Lambertz & Dehaene, 1994), we have focused on the early discrimination response (peak 2) to study the lateralization of phonetic processes in infants.

Two successive time-windows of 80 msec each (320–400 and 400–480 msec after the syllable onset), centered on peak 2 maximum, were selected. Two analyses of variance (ANOVA) were performed on voltage averaged over the four sites (frontal, central, temporal, and occipital) during each of the two time-windows, with condition (control, phoneme change, and voice change), location (frontal, central, temporal, and occipital), and hemisphere as factors. The condition factor was analyzed using three planned contrasts: a voice contrast, control versus voice change, a phoneme contrast, control versus phoneme change, and a direct

comparison of voice change versus phoneme change in order to separate voice and phonetic effects. We have first examined the interactions location  $\times$  condition, location  $\times$  hemisphere, and the triple interaction location  $\times$  hemisphere  $\times$  condition. When they were significant, we then computed post hoc analyses at each location.

From 320 to 400 msec, only the condition  $\times$  location interaction was significant ( $F(6,90) = 3.53, p = .003$ ). This interaction was significant both for the phoneme contrast ( $F(3,45) = 2.83, p = .049$ ) and the voice contrast ( $F(3,45) = 7.63, p < .001$ ). Post hoc comparisons at each location showed that for the phoneme contrast, there was a contrast effect at the frontal site ( $F(1,15) = 7.37, p = .016$ ) and a significant hemisphere  $\times$  contrast interaction at the temporal site ( $F(1,15) = 4.63, p = .048$ ), due to a discrimination response present only over the left-temporal site ( $F(1,15) = 10.45, p = .005$  and  $F(1,15) < 1$ , respectively, over the left- and right-temporal sites). For the voice contrast, the contrast effect was significant over the frontal ( $F(1,15) = 9.16, p = .008$ ), the central ( $F(1,15) = 8.2, p = .012$ ), and the occipital sites ( $F(1,15) = 5.77, p = .030$ ), with no significant interaction of voice contrast with hemisphere at any location. When a direct comparison of voice

change versus phoneme change was computed, there was no significant contrast effect at the different locations, nor any significant hemisphere  $\times$  change (voice or phoneme) interaction. Moreover, this interaction was not significant at the temporal location ( $F(1,15) = 3.74$ ,  $p = .07$ ).

From 400 to 480 msec, the condition  $\times$  location  $\times$  hemisphere interaction and the condition  $\times$  location interaction were significant ( $F(6,90) = 2.72$ ,  $p = .018$  and  $F(6,90) = 3.15$ ,  $p = .007$  respectively). The interaction with location was significant only for the voice contrast ( $F(3,45) = 6.86$ ,  $p < .001$ ;  $F(3,45) < 1$  for the phoneme contrast). Post hoc comparisons at each location showed that the voice contrast effect was significant over the frontal and occipital locations ( $p < .015$ ). There was also a significant hemisphere  $\times$  voice contrast interaction at the occipital site ( $F(1,15) = 5.13$ ,  $p = .039$ ), due to a discrimination response present only at the left-occipital site ( $F(1,15) = 22.43$ ,  $p < .001$  and  $F(1,15) < 1$ , respectively, over the left- and right-occipital sites). When a direct comparison of voice change versus phoneme change was computed, there was no significant contrast effect at any location. The hemisphere  $\times$  contrast (voice or phoneme) interaction was significant at the occipital site ( $F(1,15) = 5.53$ ,  $p = .033$ ). This contrast effect was significant at the left but not at the right-occipital site ( $F(1,15) = 18.73$ ,  $p < .001$  and  $F(1,15) < 1$ , respectively, over the left- and right-occipital sites). The voltage was more negative over the left-occipital region than over the right one for the voice change. It was the reverse for both the control and the phoneme-change conditions.

In conclusion, discrimination responses were recorded for both changes. The voltage topographies for both responses showed an anterior positivity synchronous with a posterior negativity with an inversion plane above the central regions. This suggests a main source of brain activation within the temporal regions, similar to what was indicated in a previous modelling (Dehaene-Lambertz & Dehaene, 1994; Dehaene-Lambertz & Baillet, 1998). However, the dense scalp coverage used in this experiment allows us to observe significant differences between the topographies of the discrimination responses: The voice-discrimination response is more medial than the phoneme one. Asymmetries between the left and right hemispheres were found for both changes and were in favor of the left hemisphere for both discrimination responses. No interaction of hemisphere by type of change was significant.

## Experiment 2

The examination of the 2-D temporal animations showed that the evoked response to tones, as well as to syllables, comprised two peaks. However, their latencies were different: The maxima were at 176 and 328

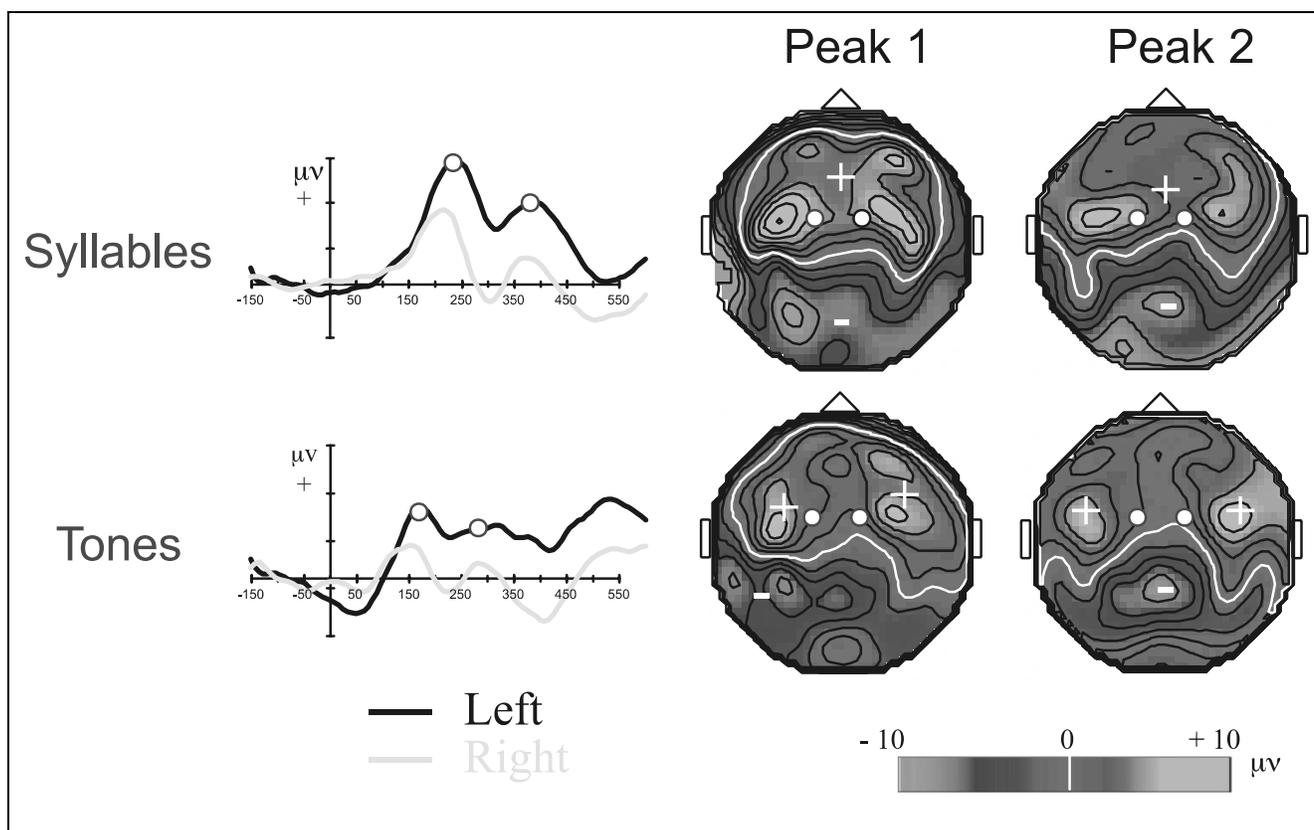
msec after the tone onset, compared to 248 and 400 msec after the syllable onset. This difference of about 80 msec was probably related to the 85 msec of formant transition that delayed the energy maximum in syllables while it was immediate in tones. For both types of stimuli, the peak amplitude decreased with repetition, and the introduction of a deviant stimulus induced a recovery of peak 2. In this second experiment, we can compare the responses to the verbal and non-verbal stimuli from the first stimulus of the trials. Therefore, we have analyzed the evoked responses to the first stimulus of the trials (S1), then, the effect of repetition on the ERP to tones and syllables (S1 to S3) and, lastly, the effect of the change at S4.

To analyze the first three stimuli of the trials, we have examined two successive time-windows of 160 msec each (160–320 and 320–480 msec after the stimulus onset), corresponding to peaks 1 and 2 of the evoked response to syllables. For S1, ANOVA was performed on voltage averaged over the four sites during each of the two time-windows, with stimulus type (tone and syllable), location (frontal, central, temporal, and occipital), and hemisphere (right and left) as factors. We have first examined the interactions location  $\times$  hemisphere, location  $\times$  stimulus type, and location  $\times$  hemisphere  $\times$  stimulus type. When they were significant, we then computed post hoc analyses at each location.

### S1 Analyses

For the two peaks, there was a significant location  $\times$  stimulus-type interaction ( $F(3,42) = 2.95$ ,  $p = .043$  for peak 1 and  $F(3,45) = 2.85$ ,  $p = .048$  for peak 2) indicating that the topography of the evoked response was different between tones and syllables (Figure 2). The hemisphere  $\times$  location interaction was significant due to higher voltages above the left hemisphere ( $F(3,42) = 4.23$ ,  $p = .011$  for peak 1 and  $F(3,42) = 3.59$ ,  $p = .021$  for peak 2), i.e., the central electrodes were more positive ( $p < .011$  for both peaks) and the temporal, occipital electrodes tended to be more negative over the left than the right hemisphere. Voltage was similar over the left- and right-frontal sites. The triple interaction hemisphere  $\times$  stimulus type  $\times$  location was not significant ( $F(3,42) < 1$  for the two peaks). Post hoc analyses were computed for each location to check for any significant hemisphere  $\times$  stimulus-type interaction. None were significant (all  $F(1,14) < 1$ ).

Some authors (Shucard, Shucard, Cummings, & Campos, 1981) have suggested that sex could influence the hemispheric lateralization, female infants showing more asymmetry than males. In both experiments, the syllable stimuli were similar. It was therefore possible to analyze the hemispheric effect for the first syllable relative to sex across these two experiments (12 fe-



**Figure 2.** Grand-averaged responses to the first stimulus of the trials (S1) in experiment 2. Left: ERP from two symmetrical central electrodes (○ on maps). Right: maps of evoked responses at the maxima of peaks 1 and 2 (● on ERP) for tones and syllables. For both types of stimuli, the response voltage is greater over the left than over the right hemisphere.

males, mean age 115 days, vs. 19 males, mean age 114 days). An analysis of variance was carried out for the same two time-windows (peaks 1 and 2) with location and hemisphere as within-subject factors and sex as a between-subjects factor. In both time-windows, there was no main effect of sex ( $F(1,29) < 1$ ), nor any significant interaction of sex with any other factor.

This analysis also confirmed that higher voltages were recorded above the left than above the right hemisphere, yielding a significant hemisphere  $\times$  location interaction ( $F(3,87) = 7.5, p = .0002$  for peak 1, and  $F(3,87) = 3.27, p = .025$  for peak 2). A hemisphere effect was significant over the central ( $p < .001$ ) and occipital locations ( $p = .020$ ) for peak 1 and over the central ( $p = .019$ ) and temporal locations ( $p = .019$ ) for peak 2. This difference in the topography of the two peaks confirms the description of Dehaene-Lambertz and Dehaene (1994). They recorded higher voltages over the posterior regions of the left hemisphere for both peaks and described peak 2 as more lateral above the temporal region than peak 1.

#### *Analyses of the Stimulus Repetition (S1 to S3)*

Habituation (S1 to S3) was studied for each location on the same two time-windows (160–320 and 320–480 msec

after the stimulus onset), with the stimulus number (1, 2, and 3), stimulus type (tone and syllable), and hemisphere (right and left) as factors. On each time-window and for all locations except the central site, there was a main effect of stimulus number ( $p < .001$ ), due to a significant decrease in voltage between stimuli 1 and 2 ( $p < .002$ ). There was a main effect of hemisphere over the central and occipital locations for peak 1 ( $p < .002$ ), and for peak 2 over the central location only ( $F(1,14) = 7.52, p = .016$ ) due to higher voltages recorded above the left than above the right hemisphere. Each time a main effect of hemisphere was present, post hoc analyses showed that this effect was significant for both syllables and tones and in the same direction for both types of stimuli. A main effect of stimulus-type was present over the central location for both peaks ( $p < .002$ ). At this location, evoked responses to tones were less positive than those to syllables (Figure 2).

#### *Analyses of S4*

To analyze S4, three successive time-windows of 80 msec (240–480 msec after the stimulus onset) were selected in order to include the onset and maximum of peak 2 for both stimulus types: Peak 2 maximum is at 328 msec

**Table 1.** Statistical Analyses of the Mismatch Response in Experiment 2 (*p* values)

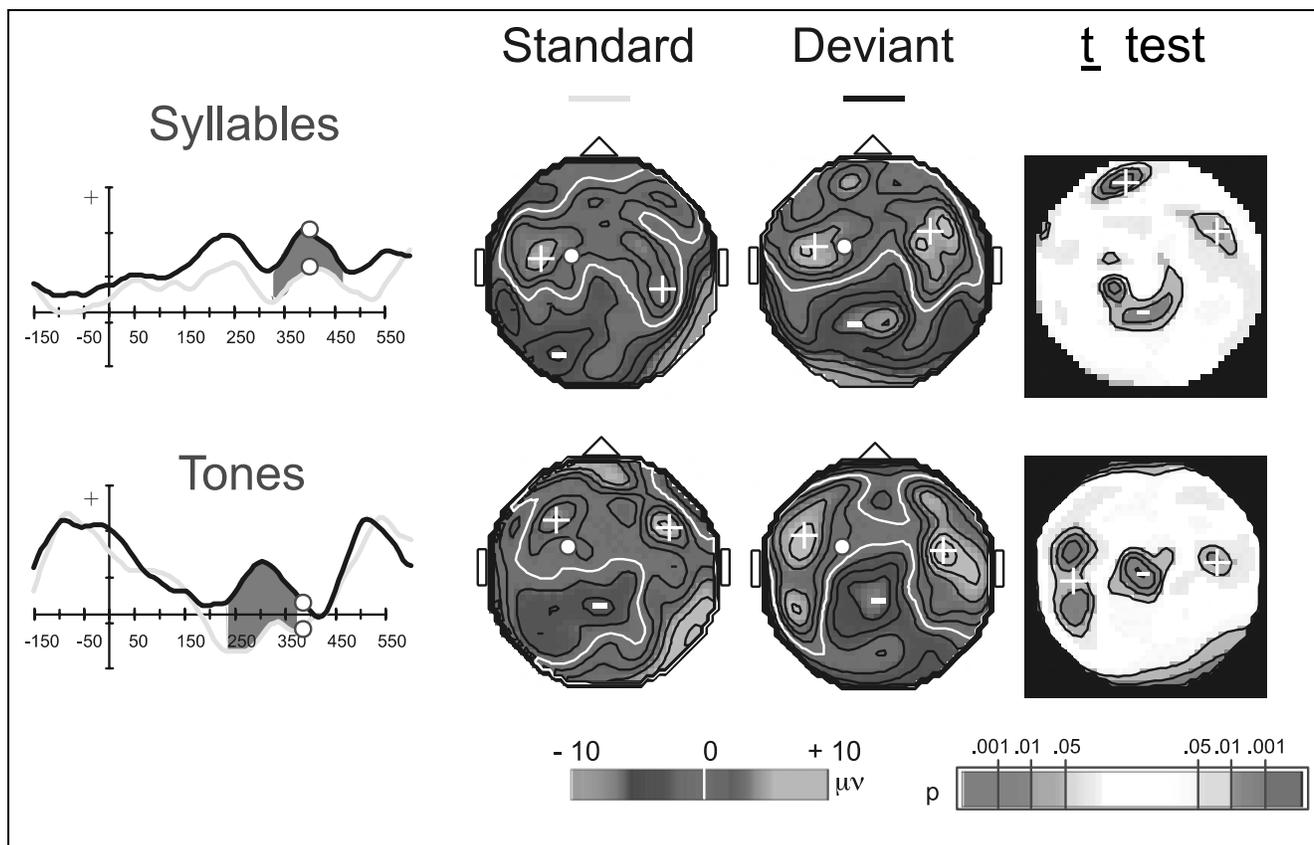
| <i>Analyses</i>                      | <i>Time-windows</i> |                     |                     |
|--------------------------------------|---------------------|---------------------|---------------------|
|                                      | <i>240–320 msec</i> | <i>320–400 msec</i> | <i>400–480 msec</i> |
| Stimulus type × location × condition | .020                | n.s.                | n.s.                |
| Condition × location                 | n.s.                | .001                | .003                |
| Analyses restricted to syllables     |                     |                     |                     |
| Location × condition                 | n.s.                | .069                | n.s.                |
| Condition/frontal                    | n.s.                | .048                | n.s.                |
| Condition/central                    | n.s.                | n.s.                | .023                |
| Condition/temporal                   | n.s.                | .015                | .028                |
| Condition/occipital                  | n.s.                | n.s.                | n.s.                |
| Analyses restricted to tones         |                     |                     |                     |
| Location × condition                 | <.001               | .004                | .007                |
| Condition/frontal                    | .031                | .006                | .013                |
| Condition/central                    | n.s.                | .027                | .039                |
| Condition/temporal                   | n.s.                | n.s.                | n.s.                |
| Condition/occipital                  | .003                | .070                | .077                |

poststimulus onset for tones and at 400 msec for syllables. ANOVA was performed on voltage averaged over the four sites during each time-window, with stimulus-type (tones and syllables), condition (control and deviant), location (frontal, central, temporal, and occipital), and hemisphere as factors. We have first examined the discrimination responses for tones and syllables. For each time-window, we have therefore computed the condition × location, and stimulus type × condition × location interactions. When they were significant, we computed post hoc analyses for each stimulus-type at each location (Table 1). Secondly, we have studied how the hemisphere factor interacts with the other factors, and computed for each time-window all the interactions that included the hemisphere and location factors. When they were significant, we have studied whether or not the left-hemisphere involvement was different in tones and syllables-discrimination responses by computing the hemisphere × condition × stimulus-type interaction at each location.

The discrimination response began earlier for tones than for syllables (see Table 1) and was present over the frontal and the occipital sites. The first time-window (240–320 msec) corresponded to the onset of peak 2 for tones, the maximum of which was at 328 msec. For syllables, the onset of peak 2 was delayed by 80 msec and the discrimination response was similarly delayed by 80 msec. It appeared in the 320–400-msec window and

was present over the frontal and temporal sites. For tones, the discrimination response was still significant over the frontal site but also over the central site and decreased over the occipital location relative to the previous time-window. These topographical differences between tones and syllables-discrimination responses yielded a significant stimulus-type effect over the central site ( $p = .005$ ). The discrimination response for both stimuli was still present during the last time-window (400–480 msec). The voltage topographies were again different yielding a significant stimulus-type effect over the frontal ( $p = .031$ ) and temporal sites ( $p = .037$ ) (Figure 3).

For all time-windows, a hemisphere × location interaction was present ( $p < .015$ ) due to higher voltages above the left than above the right hemisphere. None of the other interactions was significant. In particular, the stimulus-type did not interact with the other factors, indicating that the syllables did not induce greater asymmetries than tones. The only indication of a possible difference in hemisphere involvement for tones and syllables was a triple hemisphere × condition × stimulus-type interaction at the frontal electrodes during the 320–400 msec window ( $F(1,13) = 5.14, p = .041$ ). Post hoc analyses showed a condition effect over the left-frontal site for tones ( $F(1,13) = 13.88, p = .002$  over the left-frontal location and  $F(1,13) = 1.99, p = .182$  over the right-frontal location), and over the right-frontal site



**Figure 3.** Grand-averaged responses to the last stimulus of the trials (S4) in experiment 2. Left: ERP from a fronto-central electrode (○ on maps). Right: maps of evoked responses to standard and deviant stimuli at 400 msec following stimulus onset (● on ERP). Rightmost column: maps of statistical significance (*t* test) of deviant versus standard stimulus at the same time.

for syllables ( $F(1,13) = 1.49$ ,  $p = .244$  over the left-frontal location and  $F(1,13) = 5.00$ ,  $p = .043$  over the right-frontal location). Although in experiment 1, the hemisphere  $\times$  condition interaction was not significant, post hoc analyses showed similar results: a phoneme-discrimination response was present over the right-frontal site ( $F(1,15) = 7.28$ ,  $p = .016$ ) and not over the left-frontal site ( $F(1,15) = 3.33$ ,  $p = .088$ ). By contrast, the voice-discrimination response was slightly more important over the left than the right-frontal site ( $F(1,15) = 8.36$ ,  $p = .011$  for the left and  $F(1,15) = 6.57$ ,  $p = .022$  for the right site). Note that in experiment 1, the phoneme-discrimination response was significant over the left but not the right-temporal site. This was also the case in experiment 2 ( $F(1,13) = 6.92$ ,  $p = .021$  over the left-temporal location and  $F(1,13) = 1.13$ ,  $p = .31$  over the right-temporal site).

In conclusion, evoked responses to tones and syllables differed in latency and topography. Habituation to repeated stimuli was similar for both types of stimuli with a significant amplitude decrease between the first and the succeeding stimuli. Lastly, the discrimination response appeared at the latency of peak 2. It was therefore delayed by 80 msec for syllables relative to tones. The topography of the discrimination response

was different between the two types of stimuli: that is, over the central and occipital clusters of the electrodes for tones and over the frontal and temporal clusters for syllables, suggesting that close but different temporal regions process these two types of stimuli. In both cases, higher voltages were recorded above the left than above the right hemisphere. No indication was found of any greater left involvement in syllable processing than in tone processing.

## DISCUSSION

In these two experiments, we have compared evoked responses to acoustical and phonetic stimuli in 4-month-old infants. In both cases, the evoked responses comprised two peaks, separated by a fixed delay of about 150 msec. The amplitude of both peaks decreases with stimulus repetition but only peak 2 is affected by a stimulus change. The temporal dependency of these two peaks and their different response to a stimulus change suggest that they represent two successive steps of the auditory processing. This confirms the functional description proposed by Dehaene-Lambertz and Dehaene (1994), who related these peaks to two proces-

sing stages, corresponding to an increasingly refined analysis of the auditory input.

Dense coverage of the scalp allows us to observe significant differences between the topographies of the discrimination responses in both experiments (Figures 1 and 3). These differences were not due to variations in the degree of response of an unspecific novelty detector. In experiment 2 for example, the discrimination response came later and was shorter and less intense for syllables than for tones. A condition effect was recorded over the temporal cluster of electrodes for syllables. This was not the case for tones although the discrimination response was more intense. This suggests that different neural networks in sensory memory were involved in the representation of the features of the tones and syllables. In addition, in experiment 1 where the stimuli were exclusively syllables, the discrimination-response cartographies were different depending on the changed-syllable parameter, voice or phoneme.

In both experiments when a phonetic change was processed, the discrimination-response maxima were localized over the right-frontal area (positive polarity) and the left-temporal area (negative polarity). A negative difference between standard and deviant trials was never significant over the temporal cluster of electrodes for acoustical-discrimination responses in both experiments. The same frontal and temporal maxima for phonetic-discrimination responses were also recorded in Dehaene-Lambertz and Dehaene's (1994) and Dehaene-Lambertz and Baillet's (1997) studies.<sup>2</sup> These two studies used high-density recordings for phonetic changes. The dipole modelling presented in these two papers proposed bilateral generators within the superior-temporal gyrus. Furthermore, Dehaene-Lambertz and Baillet (1997) have recorded the response to a change of phoneme that crosses or does not cross a phonetic boundary. The dipoles were about 1 cm more posterior and dorsal when the phoneme change was across category than when it was within category. This result, together with the experiments presented in this paper, reveals an auditory cortex organized into functional networks that code the different features of an auditory stimulus.

In adults, the mismatch response (or MMN) has a voltage distribution similar to that of infants, with a maximum over the fronto-central regions and a polarity reversal over the temporal regions. The main intracerebral generators of this response in adults have been localized by the dipole-source analyses of electrical responses (Giard et al., 1995; Sherg, Vajsar, & Picton, 1989) and the magnetic fields (Tiitinen, May, Reinikainen, & Näätänen, 1994) in the auditory cortex. The mismatch response seems thus very close in adults and infants: Both responses originate in similar cortices and are elicited by a change of stimulus after a succession of repeated stimuli, while the subjects' attention is

not actively oriented toward the auditory stimuli. Here, the infants were looking at interesting visual attractors while the stimuli were presented. Note, however, that the response polarity is reversed in infants as compared to adults. In adults, the mismatch-negativity maximum is located over the right-frontal site with positivity over the left mastoid. This change in polarity in infants may be due to orientation changes of the planum temporale secondary to the frontal-lobe growth, or more probably, because the ERPs are mainly influenced by synaptic activity, this polarity inversion might be the consequence of the different rates of maturation of the cerebral layers and of their connectivity (Novak et al., 1989).

Our second aim was to study whether there was any early left-hemisphere advantage for processing speech stimuli. We opposed a place of articulation contrast (/ba/ /ga/) and a timbre contrast in voices or tones. The perception of this phonetic contrast is generally described as being left lateralized in adults, whereas timbre perception is right lateralized. This has been demonstrated either in adult brain-damaged patients (Sidtis & Volpe, 1988; Chobor & Brown, 1987; Blumstein, Baker, & Goodglass, 1977), or in normal adults using dichotic listening (Shankweiler & Studdert-Kennedy, 1967). In all these studies, behavioral responses were measured. Subjects' attention was directed toward the stimuli and discrimination was an active task. With brain-imaging techniques, Zatorre, Evans, Meyer, and Gjedde (1992) described a similar lateralization in brain activations measured with PET while the subjects were engaged in an active phonetic or pitch-discrimination task. A recent fMRI study with adults using a paradigm very close to the one we used in our second experiment has shown that syllable perception induces greater left-hemisphere activation than tone perception. However, sequences with a stimulus change compared to sequences with repeated stimuli activated left-hemispheric areas for both syllables and tones: left-posterior superior-temporal gyrus for tones and left-supramarginal gyrus for syllables (Celsis et al., in press). In this study, the subjects listened passively to the stimuli. Using magnetoencephalography, Näätänen et al. (1997) have shown that phonetic representation is based on a neural network predominantly located in the left-temporal lobe, distinct from that involved in acoustical processing which is bilateral. Here, the subjects read while the auditory stimuli were presented. Finally, Imaizumi et al. (1997) have recorded the magnetic fields elicited by a pitch change or a phoneme change in a series of repeated words while the subjects were either engaged in a deviant-detection task or reading a book. Hemispheric asymmetry, especially for phonetic processing, was significantly enhanced when the subjects were attending to the stimuli. In conclusion, a left advantage for phonetic processing is present in adults across different attentional levels, although the asym-

metry is greater when attention is actively oriented toward auditory stimuli. In our two experiments, higher voltages were recorded over the left hemisphere for both tones and syllables perception, as well as for tones and syllables discrimination, with no greater asymmetry for phonetic than for acoustic processing. It is possible that our use of a passive condition may have reduced a slight advantage of the left hemisphere in phonetic processing compared to acoustic processing to an undetectable level. It would be interesting to repeat these experiments to measure both electrophysiological and behavioral responses. However, the absence of a clear lateralization difference for phonetic and acoustic processing in our data contrasts with the high degree of intra-hemispheric specialization. This suggests that contrary to adults, the left advantage for phonetic processing is not present or still weak during the first months of life.

Finally, why were higher voltages found over the left hemisphere for all auditory stimuli? We cannot eliminate a difference in the orientation of the left- and right-temporal regions, generating a more vertical right than left dipole. In that case, the positivity maxima could be similar over the left- and right-frontal regions (as we have recorded it) while the negativity maximum of the right response could have been under the lowest position of the electrodes, around the cantho-meatal line. It is also possible that the strength of the right dipole is reduced either because the active neurons are less synchronized or because their orientation is less parallel in the right than in the left-temporal lobe. These explanations are related to limitations of the ERP technique but it is also possible that this asymmetry is a genuine functional advantage favoring the left hemisphere to process auditory stimuli. This could be related to the larger left-planum temporale. More active tissue on the left would result in higher voltages. Different maturation rates between both hemispheres could result in higher synaptic connectivity and, thus, in a more active network in the left-planum temporale. Even if we cannot separate these hypotheses at present, our data demonstrate a left-right electrophysiological asymmetry that needs to be further studied with other functional imaging techniques in order to understand whether it is a consequence of the ERP recording method or a genuine functional asymmetry that could contribute to the left-hemisphere specialization in language. Reggia, Goodall, and Skuro (1998) have recently presented a computational model demonstrating spontaneous lateralization to the side that reacts most, due to an inhibitory effect on the contra-lateral side via the corpus callosum. The intense exposure to speech by infants might explain how an initially overreacting left-auditory network could become increasingly more specialized for that type of stimuli in such a model.

In conclusion, the 4-month-old infant brain is already organized in functional networks. The different features of an auditory stimulus, such as the voice and the phonetic category of a syllable, are encoded by different neural networks in the sensory memory. This functional organization appears to be very similar to what is described in adults. The stronger left-hemispheric response to auditory stimuli should be further investigated in order to understand how this initial functional asymmetry to all auditory stimuli later turns into a left-hemispheric specialization for linguistic processing.

## METHODS

### Stimuli

Two syllables (/ba/ and /ga/), naturally produced by a French man and a French woman, were digitized at 16 kHz and matched for total duration (285 msec), pre-voiced and voiced formant transition duration (40 and 45 msec), and intensity (78 dB SPL). The fundamental frequency was 122 Hz for the male voice and 156 Hz for the female voice. Two sinewave tones were synthesized using a sound editing software (CoolEdit<sup>®</sup>) with the same base frequency of 330 Hz at a sampling rate of 16 kHz. These two tones differed by the number of harmonics: five for tone 1 and two for tone 2. This change in spectrum is perceived as a change of timbre. They were matched with the syllables for total duration and intensity.

### Procedure

We used the same experimental design that was already used by Dehaene-Lambertz and Dehaene (1994) and Dehaene-Lambertz and Baillet (1998). Each trial comprised four stimuli (stimulus onset asynchrony = 600 msec, intertrial interval = 4 sec) and there were 70 trials per condition. The first three stimuli were kept constant across the trial, the last one could change, or not, relatively to the previous stimuli. Contrary to classical oddball paradigm, this design allows better control and also study habituation and discrimination responses.

In experiment 1, the four syllables were used. For each infant, one of the four syllables, counterbalanced across subjects, was designated as the standard for the entire experiment (for example /ba/ produced by the male voice /ba/m). Three trial types were randomly presented. In the control trials, the standard was repeated four times (in our example /ba/m /ba/m /ba/m /ba/m). In the phoneme trials, a change of phoneme was introduced in the fourth position, preceded by three repetitions of the standard (/ba/m /ba/m /ba/m /ga/m). The voice was kept constant within each phoneme trial. Finally, in the voice trials, a change of voice was encountered in the fourth position, preceded by three repetitions of the

standard (/ba/m /ba/m /ba/m /ba/f). The phoneme was kept constant within each voice trial.

In experiment 2, both syllables were produced by the male speaker and both tones were used. Experiment 2 was divided into four blocks: two blocks with tones and two with syllables in alternation. The standard was different in each block; thus, the four stimuli successively served as standard and deviant within the same subject. The order of the blocks was counterbalanced across subjects. Two trial types were randomly presented in each block. In the control trials, the standard was repeated four times. In the deviant trials, a change was introduced in the fourth position, preceded by three repetitions of the standard.

Stimuli randomization, presentation, and synchronization with ERP recording system were done using the EXPE software package (Pallier & Dupoux, 1997), on a PC-compatible with a Proaudio Spectrum 16 D/A Board.

## Subjects

After the parents gave their written informed consent, 16 (4 female and 12 male) French 4-month-old infants (mean age = 112 days, range 102–123 days) were tested in experiment 1, and 15 (8 female and 7 male) in experiment 2 (mean age = 116 days, range 103–148 days). In experiment 2, four more babies were tested but were rejected because they had less than 30 artifact-free trials either in tones or in syllables blocks.

## Scalp Electrical Recording

ERPs were collected using a 64-channel geodesic electrode net referred to the vertex, and digitized at 125 Hz over a 3072-msec epoch including a 150-msec baseline. Channels contaminated by eye or motion artifacts were automatically rejected. Trials with more than 25 contaminated channels were rejected. In order to keep the infants quiet during the experimental run, attractive visual stimuli were presented on a computer screen independently from the auditory stream. For each subject and each condition, the remaining trials (average 89 trials per infant in experiment 1 and 142 trials per infant in experiment 2, comprising on average 67 trials for tones and 74 trials for syllables) were averaged, baseline corrected, transformed to an average reference and digitally filtered (band pass 0.5–20 Hz). 2-D reconstructions of scalp voltage at each time step were computed using spherical spline interpolation (Perrin, Pernier, Bertrand, & Echallier, 1989).

## Statistical Analyses

With this experimental design, three responses could be analyzed within a trial: the evoked response to the first

stimulus of the trials (S1), the response to stimulus repetition across the first three stimuli of the trials (S1 to S3) and the response to a stimulus change at the fourth position (S4). Because our aim was to study the brain reactions to different types of auditory stimuli, we have only reported the analyses in which a two-level factor (acoustic vs. phonetic) could be computed. Therefore, in the first experiment, we will only present the analyses of S4, where a phoneme change is opposed to a voice change. On another hand, we will present the analyses of the three responses in the second experiment: S1, repetition from S1 to S3, and S4 in which evoked responses to tones and syllables could be compared.

Four clusters of electrodes were selected for the statistical analyses: a frontal site including six electrodes on each side, a central site around the vertex, a supero-temporal site, and an occipital site. These last three sites included four electrodes on each side. The midline electrodes were discarded in order to study hemispheric differences. These locations were selected in agreement with previous published results (Dehaene-Lambertz & Dehaene, 1994; Dehaene-Lambertz & Baillet, 1998) and because significant statistical differences were observed at these locations when the time-course of 2-D maps of *t* test value calculated in the comparison of deviant versus control condition was inspected. In infants, the evoked response to a sound comprised two peaks, thus, successive time-windows centered on peak 1 and peak 2 maxima have been examined. ANOVAs were performed on voltage averaged over the four sites during each of the selected time-windows, with location (frontal, central, temporal, and occipital) and hemisphere (right and left) as obligatory factors and with other factors depending on the experiment (1 or 2) and on the part of the trial (S1, S1 to S3, or S4) analyzed. Because of the voltage inversion between the anterior and posterior regions (see figures), main effects might have been spurious, we therefore only reported the interactions of location with the other factors of the analysis. When this interaction was significant, post hoc analyses were performed at each site. Only significant effects were reported.

## Acknowledgments

This study was supported by the Ministère Français de la Santé et de la Recherche PHRC 1995 No. AOM95011, the Groupe-ment d'Intérêt Scientifique Sciences de la Cognition No. PO 9004, SESEP, the Fondation pour la Recherche Médicale, the Fondation Mustela, and the McDonnell Pew Foundation.

Reprint requests should be sent to: Ghislaine Dehaene-Lambertz, Laboratoire de Sciences Cognitives et Psycholinguistique, 54 boulevard Raspail, 75270 Paris cedex 06, France. Tel.: +33-1-49-54-22-62; fax: +33-1-45-44-98-35; e-mail: ghis@escp.ehess.fr.

Color versions of the figures can be found at <http://www.lhess.fr/centres/lscp>.

## Notes

1. Analyses of (stimulus 1) S1 and stimulus repetition (S1 to S3) were calculated for the first experiment but are not reported here: They showed the same significant effects that were already reported in Dehaene-Lambertz and Dehaene (1994) and Dehaene-Lambertz and Baillet (1998) studies. They were also similar to the results in experiment 2 of the present paper.

2. Cheour et al. (1998) have described that the mismatch response for a vowel change is negative at Cz and consider it as the mismatch negativity described in adults. Because they present data from a single electrode, however, it is difficult to figure out the topography of this event. Their reference electrode was placed at the left mastoid while it is at the nose in comparative adults' studies (Näätänen, 1990; Näätänen et al., 1997). This choice is surprising because the left mastoid is an active electrode in the mismatch response and this choice could modify the response polarity. In our data, the use of many electrodes permits one to calculate an average reference and to obtain reference-free event-potentials. When our data are recalculated with a left-mastoid reference as in Cheour et al.'s study, almost no positivity is visible but a diffuse negativity is present. For a phonetic change, the negativity is more posterior and extends over both temporal regions. For a timbre change, it is a more medial response, from the back to the front of the head. Contrary to adults, there is no polarity inversion over the temporal regions. Although we agree with these authors that they have recorded a mismatch response, we, however, think that the neural substrate of the mismatch responses is different because of brain maturation in infants and adults.

## REFERENCES

- Ahoniska, J., Cantell, M., Tolvanen, A., & Lyytinen, H. (1993). Speech perception and brain laterality: The effect of ear advantage on auditory event-related potential. *Brain and Language*, *45*, 127–146.
- Bertoncini, J., Morais, J., Bijeljac-Babic, R., MacAdams, S., Peretz, I., & Mehler, J. (1989). Dichotic perception and laterality in neonates. *Brain and Cognition*, *37*, 591–605.
- Best, C. T., Hoffman, H., & Glanville, B. B. (1982). Development of infant ear asymmetries for speech and music. *Perception and Psychophysics*, *31*, 75–85.
- Blumstein, S. E., Baker, E., & Goodglass, H. (1977). Phonological factors in auditory comprehension in aphasia. *Neuropsychologia*, *15*, 19–30.
- Broca, P. (1861). Remarques sur le siège de la faculté du langage articulé suivie d'une observation d'aphémie. *Bulletin de la Société anatomique de Paris*, *6*, 330.
- Celsis, P., Boulanouar, K., Doyon, B., Ranjeva, J. P., Berry, I., & Chollet, F. (in press). Differential fMRI responses in the left posterior superior temporal gyrus and left supramarginal gyrus to habituation and change detection in syllables and tones. *NeuroImage*.
- Cheour, M., Ceponiene, R., Lehtokoski, A., Luuk, A., Allik, J., Alho, K., & Näätänen, R. (1998). Development of language-specific phoneme representations in the infant brain. *Nature Neuroscience*, *1*, 351–353.
- Chi, J. G., Dooling, E. C., & Gilles, F. H. (1977). Gyrus development of the human brain. *Annals of Neurology*, *1*, 86–93.
- Chobor, K. L., & Brown, J. W. (1987). Phoneme and timbre monitoring in left and right cerebrovascular accident patient. *Brain and Language*, *30*, 278–284.
- Dehaene, S., Dupoux, E., Mehler, J., Cohen, L., Paulesu, E., Perani, D., Moortele, P. F. v. d., Lehiricy, S., & Bihan, D. L. (1997). Anatomical variability in the cortical representation of first and second languages. *NeuroReport*, *8*, 3809–3815.
- Dehaene-Lambertz, G. & (1997). Electrophysiological correlates of categorical phoneme perception in adults. *NeuroReport*, *8*, 919–924.
- Dehaene-Lambertz, G., & Baillet, S. (1998). A phonological representation in the infant brain. *NeuroReport*, *9*.
- Dehaene-Lambertz, G., & Dehaene, S. (1994). Speed and cerebral correlates of syllable discrimination in infants. *Nature*, *370*, 292–295.
- Entus, A. K. (1977). Hemispheric asymmetry in processing of dichotically presented—speech and nonspeech stimuli by infants. In S. J. Segalowitz & F. A. Gruber (Eds.), *Language development and neurological theory* (pp. 63–73). New York: Academic Press.
- Giard, M. H., Lavikainen, J., Reinikainen, R., Perrin, F., Bertrand, O., Pernier, J., & Näätänen, R. (1995). Separate representations of stimulus frequency, intensity, and duration in auditory sensory memory: An event-related potential and dipole-model analysis. *Journal of Cognitive Neuroscience*, *7*, 133–143.
- Glanville, B. B., Best, C. T., & Levenson, R. (1977). A cardiac measure of cerebral asymmetries in infant auditory perception. *Developmental Psychology*, *13*, 54–59.
- Imazumi, S., Mori, K., Kiritani, S., Kawashima, R., Sugiura, M., Fukuda, H., Itoh, K., Kato, T., Nakamura, A., Hatano, K., Kojima, S., & Nakamura, K. (1997). Vocal identification of speaker and emotion activates different brain regions. *NeuroReport*, *8*, 2809–2812.
- Johnsrude, I. S., Zatorre, R. J., Milner, B. A., & Evans, A. C. (1997). Left-hemisphere specialization for the processing of acoustic transients. *NeuroReport*, *8*, 1761–1765.
- Kim, K. H. S., Relkin, N. R., Lee, K.-M., & Hirsch, J. (1997). Distinct cortical areas associated with native and second languages. *Nature*, *338*, 171–174.
- Kreiman, J., & Lancker, D. V. (1988). Hemispheric specialization for voice recognition: Evidence from dichotic listening. *Brain and Language*, *34*, 246–252.
- Lenneberg, E. H. (1967). *Biological foundations of language*. New York: Wiley.
- Mazoyer, B. M., Dehaene, S., Tzourio, N., Frak, G., Murayama, N., Cohen, L., Salamon, G., Syrota, A., & Mehler, J. (1993). The cortical representation of speech. *Journal of Cognitive Neuroscience*, *5*, 467–479.
- Molfese, D. L., & Burger-Judisch, L. M. (1991). Dynamic temporal-spatial allocation of resources in the human brain. An alternative to the static view of hemispheric differences. In F. L. Kitterle (Ed.), *Cerebral laterality: Theory and research* (pp. 71–102). Hillsdale, NJ: Erlbaum.
- Muter, V., Taylor, S., & Vargha-Khadem, F. (1997). A longitudinal study of early intellectual development in hemiplegic children. *Neuropsychologia*, *35*, 289–298.
- Näätänen, R. (1990). The role of attention in auditory information processing as revealed by event-related potentials and other brain measures of cognitive function. *Behavioral and Brain Sciences*, *13*, 201–288.
- Näätänen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huotilainen, M., Iivonen, A., Vainio, M., Alku, P., Ilmonemi, 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.
- Novak, G. P., Kurtzberg, D., Kreuzer, J. A., & Vaughan, J. H. G. (1989). Cortical responses to speech sounds and their formants in normal infants: Maturation sequences and spatiotemporal analysis. *Electroencephalography and Clinical Neurophysiology*, *73*, 295–305.

- Pallier, C., & Dupoux, E. (1997). EXPE: An expandable programming language for on-line psychological experiments. *Behavior Research Methods, Instruments, & Computers*, *29*, 322–327.
- Perrin, F., Pernier, J., Bertrand, D., & Echallier, J. F. (1989). Spherical splines for scalp potential and current density mapping. *Electroencephalography and Clinical Neurophysiology*, *72*, 184–187.
- Reggia, J. A., Goodall, S., & Skuro, Y. (1998). Computational studies of lateralization of phoneme sequence generation. *Neural Computation*, *10*, 1277–1297.
- Shankweiler, D., & Studdert-Kennedy, M. (1967). Identification of consonants and vowels presented to left and right ears. *Quarterly Journal of Experimental Psychology*, *19*, 59–63.
- Sherg, M., Vajsar, J., & Picton, T. W. (1989). A source analysis of the late human auditory evoked potentials. *Journal of Cognitive Neuroscience*, *1*, 336–355.
- Shucard, J., Shucard, D., Cummings, K., & Campos, J. (1981). Auditory evoked potentials and sex-related differences in brain development. *Brain and Language*, *13*, 91–102.
- Sidtis, J. J., & Volpe, B. T. (1988). Selective loss of complex-pitch or speech discrimination after unilateral lesion. *Brain and Language*, *34*, 225–245.
- Tiitinen, H., May, P., Reinikainen, K., & Näätänen, R. (1994). Attentive novelty detection is governed by pre-attentive sensory memory. *Nature*, *372*, 90–92.
- Vargha-Khadem, F., & Corballis, M. (1979). Cerebral asymmetry in infants. *Brain and Language*, *8*, 1–9.
- Witelson, S. F., & Pallie, W. (1973). Left hemisphere specialization for language in the newborn: Neuroanatomical evidence for asymmetry. *Brain*, *96*, 641–646.
- Woods, M. T., & Carey, S. (1979). Language deficits after apparent clinical recovery from childhood aphasia. *Annals of Neurology*, *3*, 273–280.
- Zatorre, R. J., Evans, A. C., Meyer, E., & Gjedde, A. (1992). Lateralization of phonetic and pitch discrimination in speech processing. *Science*, *256*, 846–849.