

Mother and Stranger: An Electrophysiological Study of Voice Processing in Newborns

Maude Beauchemin^{1,2}, Berta González-Frankenberger^{1,2}, Julie Tremblay¹, Phetsamone Vannasing¹, Eduardo Martínez-Montes³, Pascal Belin^{4,5}, Renée Béland^{1,2}, Diane Francoeur¹, Ana-Maria Carceller¹, Fabrice Wallois⁶ and Maryse Lassonde^{1,2}

¹Centre de recherche, Centre Hospitalier Universitaire Sainte-Justine, Montréal, H3T 1C5, Canada, ²Centre de recherche en neuropsychologie et cognition (CERNEC), Département de Psychologie Université de Montréal, Montréal, H2V 2S9, Canada,

³Neurostatistics Department, Cuban Neuroscience Center, Havana, 6648, Cuba, ⁴Voice Neurocognition Laboratory, Institute of Neuroscience and Psychology, College of Medical, Veterinary and Life Sciences, University of Glasgow, Scotland, G12 8QB, UK,

⁵International Laboratories for Brain, Music and Sound (BRAMS), Université de Montréal & McGill University, Montréal, H2V 4P3, Canada and ⁶Groupe de Recherches sur l'Analyse Multimodale de la Fonction Cérébrale (GRAMFC, EA 4293), Université de Picardie Jules Verne, Centre Hospitalier Universitaire Nord, Amiens, 80054, France

Address correspondence to Dr Maryse Lassonde, Department of Psychology, Pavillon Marie-Victorin, Université de Montréal, 90, Ave Vincent d'Indy, Montréal, Quebec, H2V 2S9, Canada. Email: maryse.lassonde@umontreal.ca.

In the mature adult brain, there are voice selective regions that are especially tuned to familiar voices. Yet, little is known about how the infant's brain treats such information. Here, we investigated, using electrophysiology and source analyses, how newborns process their mother's voice compared with that of a stranger. Results suggest that, shortly after birth, newborns distinctly process their mother's voice at an early preattentive level and at a later presumably cognitive level. Activation sources revealed that exposure to the maternal voice elicited early language-relevant processing, whereas the stranger's voice elicited more voice-specific responses. A central probably motor response was also observed at a later time, which may reflect an innate auditory-articulatory loop. The singularity of left-dominant brain activation pattern together with its ensuing sustained greater central activation in response to the mother's voice may provide the first neurophysiologic index of the preferential mother's role in language acquisition.

Keywords: mismatch negativity, newborns, source analyses, voice processing

Introduction

Voices have proven to be special within the human auditory cortex (Belin et al. 2000). Using functional magnetic resonance imaging in human volunteers, voice selective regions can be found bilaterally along the upper bank of the superior temporal sulcus with predominant right hemisphere activation. This particularly salient stimulus in our auditory environment is also known to play a prominent role in most human interactions, even more so in the context of speaker identification (van Dommelen 1990; Belin et al. 2004). These specialized areas for voice processing are especially tuned to familiar voices as opposed to unfamiliar ones (Beauchemin et al. 2006). Behavioral evidence shows that this ability is present very early in life: Fetuses and newborns react preferentially to their mother's voice over that of a female stranger (DeCasper and Fifer 1980; Querleu et al. 1984; Ockleford et al. 1988; Kisilevsky et al. 2003). However, the means by which infants acquire this ability remain ambiguous: The neurophysiology of infant-mother interaction is still poorly understood as no study has ever looked at newborns' brain responses to their mother's voice. We therefore thoroughly investigated cortical activity in response to voices (mother and female stranger) as well as the brain regions that generate such activity in 16 newborn babies (mean age: 21 h). We found that not only do newborns process

their mother's voice more actively than that of a stranger but they also process it differently. The maternal voice (the sound /a/, 212 ms) preferentially activated language-relevant cortical areas, whereas the stranger's voice predominantly activated voice-specific areas. Additionally, the mother's and to a lesser extent the stranger's voice later activated central motor areas: a finding we interpret as reflecting an innate auditory-articulatory loop, more specifically tuned to the mother's voice.

Materials and Methods

Participants

Sixteen newborns (8–27 h, mean: 21 h) participated in the present study. They were full term (>37-week gestational age) and did not encounter any problem during pregnancy, labor, or delivery. Their birth record and neonatal exam indicated that they were healthy, and their physiological parameters at birth were considered normal (Apgar, weight, length, and head circumference). Mothers gave informed written consent, which was approved by the Ste-Justine Hospital's ethics committee, for their infant and were present during testing. All newborns explicitly showed an exogenous stimulus-related response (infantile N1 as labeled by Novak et al. 1989) that reflects normal auditory system functioning. Additionally, families of the infants did not have a history of hearing, language, or reading pathology.

Stimuli and Procedure

The stimuli and procedure were inspired from previous work in our laboratory (Beauchemin et al. 2006). Briefly, voice familiarity can objectively be measured with a modified oddball paradigm, using a frequent unfamiliar voice and 2 rare voices, only one of which is highly familiar (the mother's voice in the present case). A cerebral mismatch process (mismatch negativity [MMN]) is therefore triggered by the 2 rare stimuli in an automatic comparison process with the neuronal memory trace left by the repetitive unfamiliar frequent stimulus. The MMN has also been said to develop rather early in comparison with other event-related potential (ERP) waves. It has even been suggested to be the ontogenetically earliest discriminative response of the human brain (Cheour-Luhtanen et al. 1996).

Stimuli were samples of the French vowel /a/ pronounced by the natural voice of different speakers, all females. People were instructed to pronounce /a/ as in the word "allô," the French word for hello, to control for undesired variation in pitch when pronouncing the vowel. The amplitude envelope of the beginning of the recorded vowel was kept intact in order to keep the natural rising envelope of the sound, thus maintaining the voice as natural and identifiable as possible. All voices were segmented using Cool Edit 2000. Stimuli were 212 ms in duration (including 10 ms rise and fall time) and were presented at a constant 75 dB sound pressure level at the newborn's head so that loudness would not be responsible for any difference in brain

activation. The stimulus onset asynchrony (time interval between the onsets of 2 consecutive stimuli) was 800 ms, thus keeping the interstimulus interval at 588 ms. Stimuli were presented in a pseudo-random oddball sequence composed of 3 different stimuli: 1) a frequent unfamiliar voice (85% of the total number of presented stimuli), 2) an infrequent unfamiliar voice (7.5% of the total number of presented stimuli), and 3) a second infrequent stimulus (7.5% of the total number of presented stimuli) being the mother's voice. Any infrequent stimulus was always preceded by at least 3 frequent stimuli to increase the likelihood that a neural trace for the frequent stimulus had formed. The familiar voice was recorded when mothers visited our laboratory during their pregnancy. The unfamiliar infrequent voice was that of the attending nurse who accompanied both the mother and the newborn to our laboratory as it has been demonstrated that newborns' brain's response is larger to a novel rare stimulus than to a known rare stimulus (Sambeth et al. 2006). Therefore, as we did not want to get novelty detection confused with familiar voice "recognition," we recorded the voice of the babies' attending nurse with whom the mother had met at her monthly appointments and then every 2 weeks near the end of her pregnancy. The unfamiliar infrequent voice (the attending nurse) was the same for all newborns. It is worth mentioning that the attending nurse is herself a mother, which allowed us to control for any "motherese" effect. Caretakers in most cultures use a special speech register when talking to infants that has a unique acoustic signature called motherese. This infant-directed speech is slower, has a higher average pitch, and contains exaggerated pitch contours (Kuhl et al. 1997). Finally, the frequent unfamiliar voice was selected among a pool of available voices according to pitch similarity to both rare voices.

For the nurse's voice, fundamental frequency (F0) and center frequencies formants for F1 and F2 were 200, 865, 1413 Hz, respectively. For mothers' voices, mean F0 and mean center frequencies formants were F0 = 202 Hz (standard deviation [SD] = 17 Hz), F1 = 944 Hz (SD = 85 Hz), and F2 = 1608 Hz (SD = 133 Hz). For unfamiliar frequent voices, mean F0 and mean center frequencies formants were F0 = 210 Hz (SD = 24 Hz), F1 = 975 Hz (SD = 53 Hz), and F2 = 1580 Hz (SD = 134 Hz). To ensure that voices were acoustically comparable, we calculated the mean F0 difference between the nurse's voice and both mothers' voices and unfamiliar frequent voices prior to statistically comparing those differences. A paired *t*-test on F0 differences did not reveal any significant difference ($t_{15} = 0.87, P = 0.39$). Similarly, we computed and statistically contrasted mean differences between F2/F1 ratios of the nurse's voice relative to both mothers' voices and unfamiliar frequent voices. Again, a paired *t*-test did not show any significant difference ($t_{15} = -1.57, P = 0.13$). Both analyses confirmed that the mean intervowel space's distance between utterances of mothers and the nurse was not larger than that between utterances of unfamiliar frequent voices and the nurse's.

Four blocks of 390 stimuli were presented for a total of 1560 stimuli (117 infrequent mother's voice stimuli, 117 infrequent stranger's voice stimuli, and 1326 frequent unfamiliar voice stimuli) and for a total duration of 24 min. The sequence of stimuli was generated by the E-Prime Psychology Software on a DELL computer located in an adjacent room. Stimuli were binaurally delivered through speakers, positioned 20 cm left and right from the newborn's head, at auricular height.

ERPs are small voltage fluctuations reflecting, with high temporal resolution, patterns of neuronal activity evoked by a stimulus. Brain responses were acquired with a high-density recording system, the Geodesic 128-Sensor Net (Electrical Geodesics Inc.) at a sampling rate of 250 Hz, a bandpass filter of 0.1–100 Hz, and Cz referenced. Mothers were seated in a comfortable chair located in a semiobscure Faraday room with their newborn lying on their laps. All babies were in active sleep state during recording as, even when asleep, neonates are able to process external information actively (Cheour, Ceponiene, et al. 2002; Fifer et al. 2010) and as the MMN can be observed in infants throughout all sleep stages as well as when they are awake (Cheour, Alho, et al. 1998; Cheour, Ceponiene, et al. 1998; Cheour et al. 2000). Sleep stage classification was scored off-line according to the electroencephalogram (EEG) signal, electromyogram (EMG) and electrooculogram (EOG) (Anders 1979), and during all recordings, newborns were in

active state. It is noteworthy that all mismatch responses (mother/frequent and stranger/frequent) were simultaneously recorded, thereby excluding differential sleep patterns as being responsible for any differences observed in mismatch responses. Electrode impedance was kept below 50 k Ω , which is within recommended range when using a powerful amplifier such as the one used in this study (Net Amps 200). Time-locked evoked potentials (EPs) were recorded with the Net Station program on a Macintosh G4 computer. Mothers and newborns were monitored through an infrared camera equipped with an integrated speaker, allowing them to communicate at all times with the experimenters located in an adjacent room. The attending nurse was also always present in the Faraday room, together with mothers and newborns. All infants were fed immediately before testing to increase the likelihood that they would sleep through the entire procedure. Pauses were provided between blocks to monitor the baby's temperature and respiratory state, thereby preventing them from reaching the quiet sleep stage.

Data Analysis

EEG time series of 600 ms with a 100 ms prestimulus interval were edited off-line by visual inspection using the BrainVision Analyzer program (Brain products). Baseline corrections were performed together with several preprocessing filters and ocular corrections (EOG). Artifacts were also manually removed. EEG time series with amplitudes over 100 μ V were withdrawn from the analysis. Electrodes were referenced to both linked mastoids with a frequency bandpass filter of 1–30 Hz at 24 dB/octave. For each subject, EEG time series were edited in response to the 3 types of stimuli: 1) mother's voice (mean = 97.2 stimuli remaining from the 117 presented ones), 2) stranger's voice (mean = 99.5 stimuli remaining from the 117 presented ones), and 3) frequent voice (mean = 839.8 stimuli remaining from the 1326 presented ones). The edited EEG time series for each subject and for each condition were exported to Matlab v7.0.4 (The MathWorks, Inc.) in order to apply the statistical analyses.

Statistical Analyses

MMN responses of the newborns to the mother's voice and to the stranger's voice were computed by subtracting the EPs obtained in the frequent condition from those in response to the mother's and stranger's voices, respectively. A nonparametric permutation test was applied to find the time points with significant differences in each of the following comparisons: 1) mother's voice against frequent, 2) stranger's voice against frequent, and 3) mother's MMN against stranger's MMN. A total of 500 permutations under the null hypothesis of no-difference between conditions were carried out separately on the EEG time series of 3 midline electrodes (Fz, FCz, and Cz), where differences in the electrophysiological response (i.e., MMN) were expected (Riera and Fuentes 1998; Kazemi, Grebe, et al. 2007; Kazemi, Moghaddam, et al. 2007). The results of the permutation test were shown as plots of probabilities for accepting the null hypothesis for each sampled time point and for each electrode, defining as significant those time points with probability below the significance level of 0.05 (Fig. 1C).

Source Analyses

In order to identify the generators of the MMN responses, source analyses on the 117 electrodes (i.e., excluding the 11 EOG and EMG electrodes from the 128-electrode set) were computed with low resolution electromagnetic tomography analysis (LORETA; Pascual-Marqui et al. 1994) in the 600 ms EEG time series, rendering a solution every 4 ms, and using the NEURONIC Source Localizer and Tomographic Viewer programs (Neuronic Inc.). The newborn magnetic resonance image used as template for the source analysis was created by Kazemi and collaborators (Kazemi, Grebe, et al. 2007; Kazemi, Moghaddam, et al. 2007) and preprocessed with the NEURONIC iMagic Pro software (Neuronic Inc.) for extracting the surface of the head and fitting the electrodes to this surface. An isotropic and piecewise homogeneous 3-sphere head model with 6,368 (generators) voxels inside the newborn's brain with a resolution of

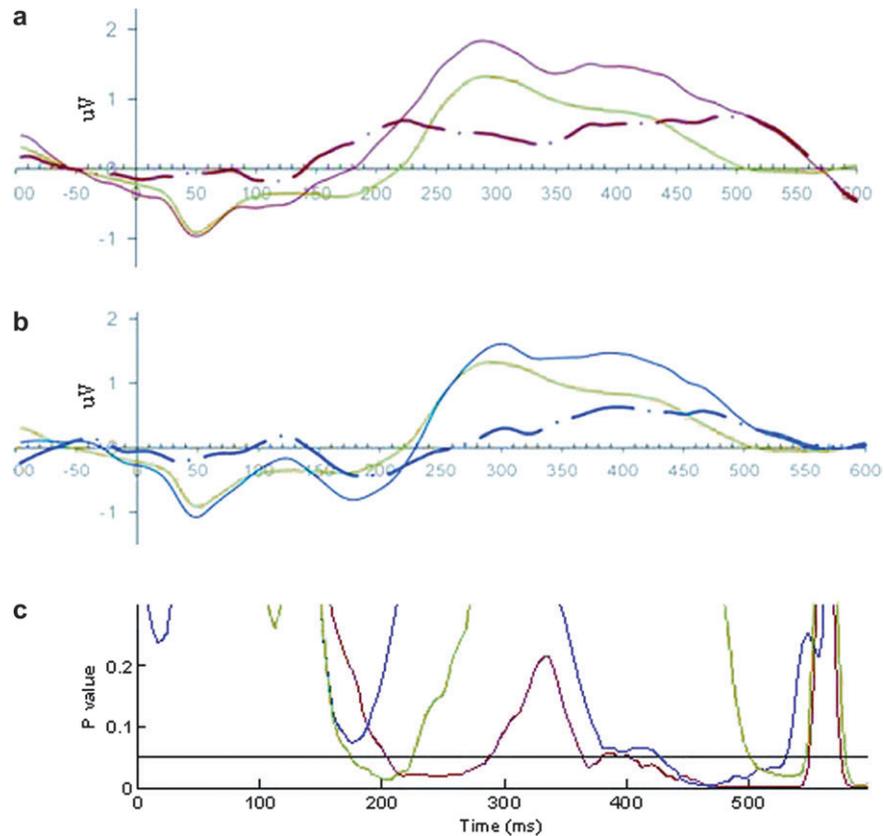


Figure 1. Event-related brain potentials and their correspondent permutations analyses. (A) Grand-average ERPs in response to the frequent unfamiliar voice (green), the rare mother's voice (red), and the resulting subtraction (MMN, dashed). For purposes of clarity, only the response recorded over FCz is represented here, but similar results were obtained on all electrodes that are usually activated during MMN processing (Fz, FCz, and Cz). (B) Grand-average ERPs in response to the frequent unfamiliar voice (green), the rare stranger's voice (blue), and the resulting subtraction (MMN, dashed). Again, we report only the result obtained over FCz. (C) Permutations analyses showing significant time windows (below significance level of 0.05, horizontal line on the figure) for the mother's voice response (red), the stranger's voice response (blue), and for the comparison between both MMNs (green) on electrode FCz.

4 mm was used for obtaining the electric lead field (Riera and Fuentes 1998). LORETA solutions were first calculated for each condition in every individual subject, and these were averaged across subjects to find the mean LORETA solution for each condition.

Results

Figure 1A illustrates the grand-average ERPs (ERPs) in response to the frequent unfamiliar voice, the rare mother's voice, and the resulting subtraction (mismatch negativity or MMN). Permutations (Lage-Castellanos et al. 2010; Fig. 1C) showed 2 statistically significant time windows (an early preattentive component and a later presumably cognitive component) in which the response to the mother's voice differed from the frequent unfamiliar voice. The mother's voice elicited a significantly greater response than the frequent unfamiliar voice from 204 ms ($t_{1,15} = 2.050$, $P < 0.05$) to 284 ms ($t_{1,15} = 1.994$, $P < 0.05$) and then again from 364 ms ($t_{1,15} = 2.039$, $P < 0.05$) until 548 ms ($t_{1,15} = 2.880$, $P < 0.05$), peaking at 524 ms ($t_{1,15} = 5.261$, $P < 0.05$). These results were significant on all 3 midline electrodes (Fz, FCz, and Cz) where MMN is found to be of maximum amplitude (Ilvonen et al. 2004; Naatanen et al. 2004; Beauchemin et al. 2006).

Figure 1B illustrates the grand-average ERPs in response to the frequent unfamiliar voice, the rare unfamiliar voice (the research nurse) and the resulting MMN. As shown in Figure 1C, the rare unfamiliar voice also tended to be different from the

frequent unfamiliar voice at around 176 ms ($t_{1,15} = 1.587$, $P > 0.05$), but this difference never reached significance. Permutations analyses did reveal a significant late component highly similar to that found for the mother's voice: the rare unfamiliar voice was significantly different from the frequent unfamiliar voice between 428 ms ($t_{1,15} = 1.991$, $P < 0.05$) and 528 ms ($t_{1,15} = 2.236$, $P < 0.05$), peaking at 464 ms ($t_{1,15} = 2.657$, $P < 0.05$). At the maximum peaking amplitude, a much larger t -value obtained in the mother's voice condition ($t = 5.261$ vs. $t = 2.657$ for the stranger's voice condition) implies that this late component is more robust in response to the mother's (0.8 μ V) than to the stranger's voice (0.3 μ V).

Additional permutations analyses were performed to compare the mother's voice MMN with the stranger's (Fig. 1C). Statistically significant time windows were found to closely match those of previous permutations computed on both rare voice conditions relative to the frequent one. The mismatch evoked by the mother's voice differed from that of the stranger's starting at 176 ms ($t_{1,15} = 1.935$, $P < 0.05$) until 224 ms ($t_{1,15} = 2.075$, $P < 0.05$), with a maximal differentiation at 204 ms ($t_{1,15} = 2.369$, $P < 0.05$). Permutations yielded a second significant time window from 504 ms ($t_{1,15} = 2.310$, $P < 0.05$) to 544 ms ($t_{1,15} = 1.921$, $P < 0.05$), with a maximal peak at 528 ms ($t_{1,15} = 2.573$, $P < 0.05$). All infants showed this specific pattern as the amplitude of their response to their

mother's voice and to the nurse's voice remained within 2 SDs from the mean according to z -score calculations.

We then compared the total amplitude of brain activity elicited by both MMNs. BrainVision-assisted t -tests were computed on amplitude topographies during the 2 time windows showing maximal differences between the responses evoked by the mother's and the stranger's voice (Fig. 2). At the early latency (i.e., at 204 ms), the mother's voice elicited greater activation over several fronto-temporal electrode sites, $t_{1, 15} > 2.1$, $P = 0.05$, that correspond to the MMN

topographical distribution reported in adults (Beauchemin et al. 2006). At the late latency (i.e., at 528 ms), the mother's voice also produced greater activity over several brain regions, particularly in the right hemisphere, $t_{1, 15} > 2.1$, $P = 0.05$.

For the very first time, source analyses were computed on the newborn brain using a neonatal atlas template provided by Kazemi and collaborators (Kazemi, Grebe, et al. 2007; Kazemi, Moghaddam, et al. 2007). The mean LORETA solutions corresponding to the newborns' response to the mother's voice and to the stranger's voice are shown in Figure 3. We

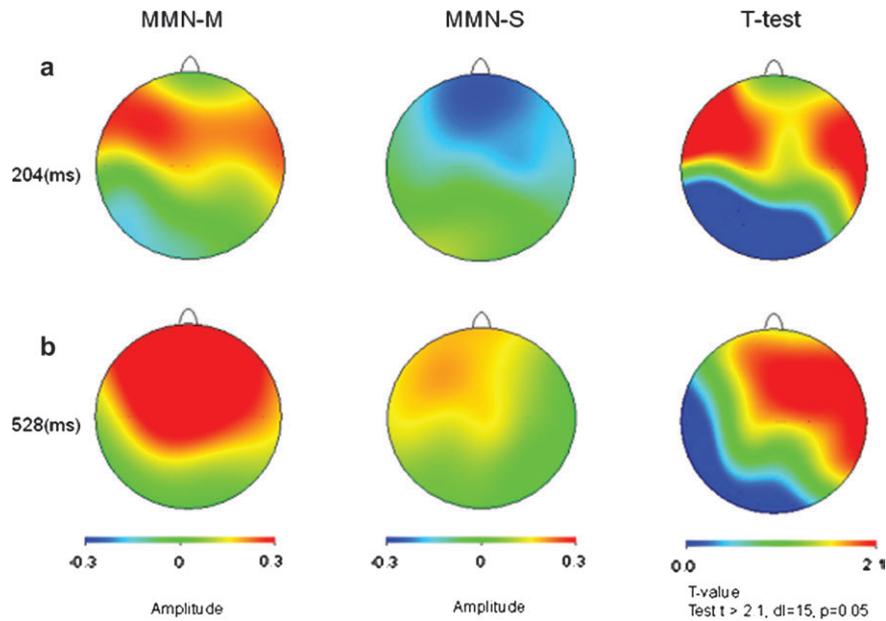


Figure 2. Brain topographies. Time course of brain activity topographies elicited when comparing both MMNs (MMN-M: mother's and MMN-S: stranger's) in the first preattentive time window (A) and in the later time window (B). Note that the mother's voice elicited a much greater amplitude of activation over both time points in several electrode sites.

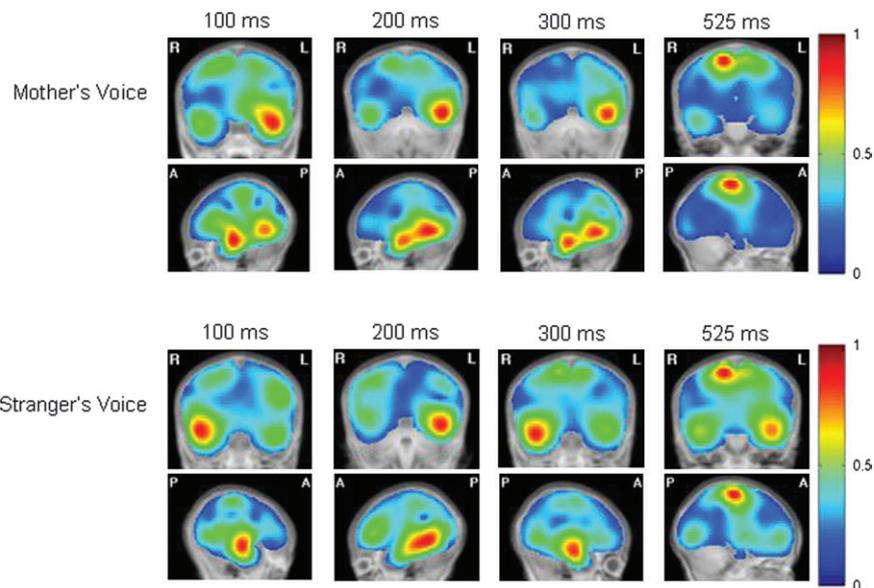


Figure 3. Source distribution. Mean LORETA solutions (represented in coronal and sagittal planes) for the MMNs induced by the mother's and the stranger's voices at 4 different latencies. These images are normalized in order to compare just the source distribution and not the amplitudes of activation (which are presented in Fig. 2). Note that the orientation of the sagittal planes—(A)nterior/(P)osterior—might switch in order to illustrate the lateralization of the maximal source distribution (in the corresponding hemisphere).

illustrated source distribution at 4 specific time points: 1) 100 ms poststimulus presentation (just prior to significant differentiation of mother's vs. stranger's voice), 2) 200 ms (during the time window that revealed a maximal significant difference between the mother's vs. stranger's voice), 3) 300 ms (time point that just follows the end of the significant time window), and 4) 525 ms (time point where the mother's voice is most different from the stranger's in this late component).

Looking at average source distributions (Fig. 3), we found that mother's voice processing initially activated the left posterior temporal lobe (100 ms), an area particularly involved in language comprehension in the adult brain. This source generation was maintained in the left temporal areas during the time window that revealed a significant difference (200 ms) and even subsequently (300 ms). Source generators then shifted to right central regions at the second time point where a significant difference was found (525 ms). As for the processing of the stranger's voice, right temporal lobe processing was rather found in its early stage (100 ms), followed by a brief left temporal processing (200 ms). Source generation then switched back to the right temporal lobe (300 ms), and finally, in a similar way as the response to the mother's voice, the newborn's response to the stranger's voice was later found over right central sites (525 ms).

Table 1 indicates the percentage of newborns having their maximal source distribution in different regions, namely temporal right, temporal left, central right, and others for each time point illustrated in Figure 3. In response to the mother's voice, the most frequent maximal activations were temporal left (100 ms, as in Fig. 3), temporal left (200 ms, as in Fig. 3), central right (300 ms, vs. temporal left in Fig. 3), and central right (525 ms, as in Fig. 3). In response to the stranger's voice, the most frequent maximal activations were temporal right (100 ms, as in Fig. 3), temporal right (200 ms, vs. temporal left in Fig. 3), temporal right (300 ms, as in Fig. 3), and temporal left (525 ms, vs. central right in Fig. 3). Hence, when the rare voices were compared with the standard voice, the mother's voice elicited, in general, a maximal left temporal activation until 525 ms where a right central activation was observed. In contrast, the stranger's voice elicited right temporal activations until 525 ms where a left temporal activation was observed. Although differing slightly from grand averages provided in Figure 3, these percentages further confirm the distinct processing of both, mother's and stranger's, voices in terms of left/right hemisphere activation. Differences between these percentages (Table 1) and grand averages (Fig. 3) are observed

Table 1
Percentage of newborns with maximal LORETA solution in different regions

	Time	Region			
		Temp left	Temp right	Cent right	Other
Mother	100	43	18	25	12
	200	50	18	25	6
	300	25	25	43	6
	525	18	25	50	6
Stranger	100	37	43	0	18
	200	25	37	18	18
	300	31	31	18	18
	525	56	25	18	0

Note: The highest percentage for the 4 different latencies illustrated in Figure 3 is indicated in bold.

since most babies show a source distribution as illustrated in Figure 3, albeit not a maximal one.

To better support the singularity of left-dominant brain activation pattern in response to the mother's voice, we have looked at the maximal LORETA solutions at every sample point (every 4 ms) for a total of 149 samples. When looking at the lateralization of those solutions, we found that the mother's voice activated the left hemisphere in 116 samples (77.85%) and the right hemisphere in 33 samples (22.14%). The stranger's voice activated the left hemisphere in 61 samples (40.93%) and the right hemisphere in 87 samples (59.06%).

Discussion

Our findings suggest that, shortly after birth, newborns process their mother's voice differently and more actively than that of strangers at both an early and a late stage of processing. The time course of activation sources revealed mother's voice-specific brain activation patterns. Indeed, the mother's voice was found to be preferentially processed in the left temporal lobe at early latencies before activating central right regions. Since the voice stimulus was the pronounced vowel /a/, activation of the left temporal lobe could suggest that exposure to the maternal voice elicited language-relevant stimulus processing. In contrast, the stranger's voice elicited predominant right temporal lobe voice-specific response (Belin et al. 2002) prior to stimulating central right brain areas. In addition to providing further support to previously demonstrated tuning of specialized voice processing brain areas for familiar voices (Beauchemin et al. 2006), these results also suggest that this tuning is functional from birth or at least within the first 24 h following birth.

General perceptual and cognitive abilities may account for infants' distinctive mother's voice processing. The first distinction we found in the ERPs waveform is thought to reflect the mismatch process, commonly known as a preattentive sensory memory detection of changes in a sound stream, arguing in favor that sleeping neonates are forming representations of specific stimuli and distinguishing between them during sleep (Alho et al. 1998; Cheour, Alho, et al. 1998; Cheour, Ceponiene, et al. 1998; Cheour, Martynova, et al. 2002; Sambeth et al. 2008).

The MMN amplitude usually increases with an increasing acoustic difference between deviant and standard stimuli (Tiitinen et al. 1994). However, Cheour, Ceponiene, et al. (1998) have demonstrated that brain memory traces for speech sounds override a greater acoustical difference between 2 stimuli: MMN amplitudes are larger in response to a native vowel than they are for a nonnative one, even if that nonnative vowel is acoustically more different from the standard stimulus. This assumption is also supported by previous work done in our laboratory showing greater cortical activity in response to a familiar voice when compared to that of an unfamiliar voice (Beauchemin et al. 2006). According to the contention that long-term memory traces exert a marked impact on auditory short-term memory, and therefore on MMN elicitation (Naatanen et al. 1997; Huotilainen et al. 2001; Beauchemin et al. 2006), our results also provide preliminary evidence that long-term memory is efficient from birth.

In terms of polarity, the MMN obtained in response to the mother's voice tends to be positive when compared to that obtained in response to the nurse's voice, which in the contrary tends to be negative. Kushnerenko et al. (2002) have

proposed that, in infants, a response of positive polarity might be an analog of the adult P3a component, indexing an involuntary attention switch to the deviant stimuli. These authors suggest that, because it emerges at the same latency as the MMN in response to a deviant stimulus, the large amplitude P3a can mask the MMN. Earlier findings have also demonstrated that the P3a amplitude increases as a function of the magnitude of stimulus change (Yago et al. 2001), and we have shown that it is particularly modulated by voice familiarity in adults (Beauchemin et al. 2006). We therefore suggest that the MMN in response to the deviant mother's voice might have been partly overlapped by the subsequent positivity, making it positive in polarity. This also suggests that the familiar mother's voice elicit an involuntary attention deployment but that the less familiar nurse's voice fails to do so, arguing further in favor of the specificity of the response to the mother's voice.

Even if monolingual French speakers were used in the current study, it is well known that infants are able to discriminate almost all phonetic contrasts (Eimas et al. 1971; Streeter 1976; Werker et al. 1981). Many acoustical features can differentiate between 2 voices, even pronouncing the same vowel: the fundamental frequency of phonation, for example. But even when F0 are similar between 2 voices, other acoustical cues such as formant frequencies (related to the vowel being spoken but also to the speaker's vocal tract) or aspects of voice quality such as harmonicity (that can be measured by the harmonics-to-noise ratio) and time/frequency irregularities (captured by measures such as jitter or shimmer) contribute to the perceived differentiation of the 2 voices.

This being said, because of the many potential candidate acoustical features for the distinction between the mother's voice and other voices, it is unclear at present which one(s) contribute to the effect we report. This important question can be addressed in the future by using acoustical manipulations of the different voices used such that they are equated along one or more acoustical features (e.g., equating voices in both F0 and first formant frequencies using morphing). Then, it will be possible to examine which of the different potential acoustical cues to voice identity are being used by the newborn's brain. But this question is beyond the scope of the present study that focuses on the first report of cerebral response to voice identity differences in the newborn.

A second later difference was also found which, we postulate, reveals a higher-level cognitive process. Sleeping neonates are able to process external information actively (Fifer et al. 2010), and authors have demonstrated their ability to learn, even when asleep (Cheour, Martynova, et al. 2002; Fifer et al. 2010). In fact, Sambeth et al. (2008) have shown that newborns process structural aspects of language while sleeping. The distinct left-dominant brain activation pattern that was found in the present study for the vowel /a/ when pronounced by the newborn's mother as opposed to right-dominant brain response for that of strangers may shed some light on the uniqueness of the mother's voice for language acquisition in infants. Regardless of culture, language and speech are acquired quickly and seemingly without effort. Social interaction assists language learning as revealed by Kuhl et al. (2003) in their study comparing live social interactions with televised foreign language learning. Social interaction might play a more significant role in early language development than previously thought (Kuhl 2007). In both speech perception and speech production, the presence of

a human being interacting with a child has a strong influence on learning (Kuhl 2004). Cases of children raised in social isolation have proven the severe and negative impact of social deprivation on language development to the extent that normal language skills are never acquired (Kuhl 2004; Kuhl et al. 2005).

If language and speech acquisition is learning and environment dependent, then there must be shared neural systems that oversee perception and action, "mirror systems" (Kuhl and Meltzoff 1996; Kuhl and Rivera-Gaxiola 2008). Although it was known that prenatal exposure to native language prosody influences newborns' perception (Mehler et al. 1988; Moon et al. 1993), Mampe et al. (2009) have recently shown a tendency for infants to utter melody contours similar to those perceived prenatally. These data are suggesting that not only did infants memorize the main intonation patterns of their respective surrounding language but also that they were able to reproduce these patterns in their own production. Similarly, Chen et al. (2004) have demonstrated that newborns can perform corresponding mouth movements to both vowel and consonant vocal models. These authors believe that it would be more plausible and parsimonious to account for their findings in terms of a unified underlying intermodal mapping, especially considering that there was no difference in the performance of matching mouth movements between infants who closed their eyes and those who had their eyes open. Overall, their data are suggesting that newborns can map perceived sounds onto corresponding mouth movements, even if they have not seen these mouth movements in others, arguing in favor that infants possess some kind of auditory-articulatory map from birth (Kuhl and Meltzoff 1996).

Our finding of subsequent significant central right brain activation to the mother's voice, and to a lesser extent to the stranger's voice, is compatible with the implication of premotor/supplementary motor areas as underlying neurobiological substrates of the late motor production component of the previously proposed innate auditory-articulatory map (Kuhl and Meltzoff 1996). In sum, the singularity of left-dominant brain activation pattern to the mother's voice together with its ensuing sustained greater right central activation could very well provide the first neurophysiologic index of language acquisition occurring through imitation, a process that would be particularly linked to the special mother-infant interaction.

Funding

Natural Sciences and Engineering Research Council of Canada grant (M.L.) and graduate scholarship (M.B.), Canada Foundation for Innovation (M.L.), and Canada Research Chair (M.L.).

Notes

Conflict of Interest: None declared.

References

- Alho K, Connolly JF, Cheour M, Lehtokoski A, Huotilainen M, Virtanen J, Aulanko R, Ilmoniemi RJ. 1998. Hemispheric lateralization in preattentive processing of speech sounds. *Neurosci Lett*. 258:9-12.
- Anders TF. 1979. Night-waking in infants during the first year of life. *Pediatrics*. 63:860-864.

- Beauchemin M, De Beaumont L, Vannasing P, Turcotte A, Arcand C, Belin P, Lassonde M. 2006. Electrophysiological markers of voice familiarity. *Eur J Neurosci*. 23:3081-3086.
- Belin P, Fecteau S, Bedard C. 2004. Thinking the voice: neural correlates of voice perception. *Trends Cogn Sci*. 8:129-135.
- Belin P, Zatorre RJ, Ahad P. 2002. Human temporal-lobe response to vocal sounds. *Brain Res Cogn Brain Res*. 13:17-26.
- Belin P, Zatorre RJ, Lafaille P, Ahad P, Pike B. 2000. Voice-selective areas in human auditory cortex. *Nature*. 403:309-312.
- Chen X, Striano T, Rakoczy H. 2004. Auditory-oral matching behavior in newborns. *Dev Sci*. 7:42-47.
- Cheour M, Alho K, Ceponiene R, Reinikainen K, Sainio K, Pohjavuori M, Aaltonen O, Naatanen R. 1998. Maturation of mismatch negativity in infants. *Int J Psychophysiol*. 29:217-226.
- Cheour M, Ceponiene R, Lehtokoski A, Luuk A, Allik J, Alho K, Naatanen R. 1998. Development of language-specific phoneme representations in the infant brain. *Nat Neurosci*. 1:351-353.
- Cheour M, Ceponiene R, Leppanen P, Alho K, Kujala T, Renlund M, Fellman V, Naatanen R. 2002. The auditory sensory memory trace decays rapidly in newborns. *Scand J Psychol*. 43:33-39.
- Cheour M, Leppanen PH, Kraus N. 2000. Mismatch negativity (MMN) as a tool for investigating auditory discrimination and sensory memory in infants and children. *Clin Neurophysiol*. 111:4-16.
- Cheour M, Martynova O, Naatanen R, Erkkola R, Sillanpaa M, Kero P, Raz A, Kaipio ML, Hiltunen J, Aaltonen O, et al. 2002. Speech sounds learned by sleeping newborns. *Nature*. 415:599-600.
- Cheour-Luhtanen M, Alho K, Sainio K, Rinne T, Reinikainen K, Pohjavuori M, Renlund M, Aaltonen O, Eerola O, Naatanen R. 1996. The ontogenetically earliest discriminative response of the human brain. *Psychophysiology*. 33:478-481.
- DeCasper AJ, Fifer WP. 1980. Of human bonding: newborns prefer their mothers' voices. *Science*. 208:1174-1176.
- van Dommelen WA. 1990. Acoustic parameters in human speaker recognition. *Lang Speech*. 33(Pt 3):259-272.
- Eimas PD, Siqueland ER, Jusczyk P, Vigorito J. 1971. Speech perception in infants. *Science*. 171:303-306.
- Fifer WP, Byrd DL, Kaku M, Eigsti IM, Isler JR, Grose-Fifer J, Tarullo AR, Balsam PD. 2010. Newborn infants learn during sleep. *Proc Natl Acad Sci U S A*. 107:10320-10323.
- Huotilainen M, Kujala A, Alku P. 2001. Long-term memory traces facilitate short-term memory trace formation in audition in humans. *Neurosci Lett*. 310:133-136.
- Iivonen T, Kujala T, Kozou H, Kiesilainen A, Salonen O, Alku P, Naatanen R. 2004. The processing of speech and non-speech sounds in aphasic patients as reflected by the mismatch negativity (MMN). *Neurosci Lett*. 366:235-240.
- Kazemi K, Grebe R, Moghaddam AH, Lagadec P, Gondry-Jouet C, Wallois F. 2007. Design of a digital phantom of the neonatal brain. *Conf Proc IEEE Eng Med Biol Soc*. 2007: 5509-5512.
- Kazemi K, Moghaddam HA, Grebe R, Gondry-Jouet C, Wallois F. 2007. A neonatal atlas template for spatial normalization of whole-brain magnetic resonance images of newborns: preliminary results. *Neuroimage*. 37:463-473.
- Kisilevsky BS, Hains SM, Lee K, Xie X, Huang H, Ye HH, Zhang K, Wang Z. 2003. Effects of experience on fetal voice recognition. *Psychol Sci*. 14:220-224.
- Kuhl P, Rivera-Gaxiola M. 2008. Neural substrates of language acquisition. *Annu Rev Neurosci*. 31:511-534.
- Kuhl PK. 2004. Early language acquisition: cracking the speech code. *Nat Rev Neurosci*. 5:831-843.
- Kuhl PK. 2007. Is speech learning 'gated' by the social brain? *Dev Sci*. 10:110-120.
- Kuhl PK, Andruski JE, Chistovich IA, Chistovich LA, Kozhevnikova EV, Ryskina VL, Stolyarova EI, Sundberg U, Lacerda F. 1997. Cross-language analysis of phonetic units in language addressed to infants. *Science*. 277:684-686.
- Kuhl PK, Coffey-Corina S, Padden D, Dawson G. 2005. Links between social and linguistic processing of speech in preschool children with autism: behavioral and electrophysiological measures. *Dev Sci*. 8:F1-F12.
- Kuhl PK, Meltzoff AN. 1996. Infant vocalizations in response to speech: vocal imitation and developmental change. *J Acoust Soc Am*. 100:2425-2438.
- Kuhl PK, Tsao FM, Liu HM. 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.
- Kushnerenko E, Ceponiene R, Balan P, Fellman V, Naatanen R. 2002. Maturation of the auditory change detection response in infants: a longitudinal ERP study. *Neuroreport*. 13:1843-1848.
- Lage-Castellanos A, Martinez-Montes E, Hernandez-Cabrera JA, Galan L. 2010. False discovery rate and permutation test: an evaluation in ERP data analysis. *Stat Med*. 29:63-74.
- Mampe B, Friederici AD, Christophe A, Wermke K. 2009. Newborns' cry melody is shaped by their native language. *Curr Biol*. 19:1994-1997.
- Mehler J, Jusczyk P, Lambertz G, Halsted N, Bertocini J, Amiel-Tison C. 1988. A precursor of language acquisition in young infants. *Cognition*. 29:143-178.
- Moon C, Cooper R, Fifer W. 1993. Two-day-olds prefer their native language. *Infant Behav Dev*. 16:495-500.
- Naatanen R, Lehtokoski A, Lennes M, Cheour M, Huotilainen M, Iivonen A, Vainio M, Alku P, Ilmoniemi RJ, Luuk A, et al. 1997. Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature*. 385:432-434.
- Naatanen R, Syssoeva O, Takegata R. 2004. Automatic time perception in the human brain for intervals ranging from milliseconds to seconds. *Psychophysiology*. 41:660-663.
- Novak GP, Kurtzberg D, Kreuzer JA, Vaughan HG, Jr. 1989. Cortical responses to speech sounds and their formants in normal infants: maturational sequence and spatiotemporal analysis. *Electroencephalogr Clin Neurophysiol*. 73:295-305.
- Ockleford EM, Vince MA, Layton C, Reader MR. 1988. Responses of neonates to parents' and others' voices. *Early Hum Dev*. 18:27-36.
- Pascual-Marqui RD, Michel CM, Lehmann D. 1994. Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain. *Int J Psychophysiol*. 18:49-65.
- Querleu D, Lefebvre C, Titran M, Renard X, Morillon M, Crepin G. 1984. [Reaction of the newborn infant less than 2 hours after birth to the maternal voice]. *J Gynecol Obstet Biol Reprod (Paris)*. 13:125-134.
- Riera JJ, Fuentes ME. 1998. Electric lead field for a piecewise homogeneous volume conductor model of the head. *IEEE Trans Biomed Eng*. 45:746-753.
- Sambeth A, Huotilainen M, Kushnerenko E, Fellman V, Pihko E. 2006. Newborns discriminate novel from harmonic sounds: a study using magnetoencephalography. *Clin Neurophysiol*. 117:496-503.
- Sambeth A, Ruohio K, Alku P, Fellman V, Huotilainen M. 2008. Sleeping newborns extract prosody from continuous speech. *Clin Neurophysiol*. 119:332-341.
- Streeter LA. 1976. Language perception of 2-month-old infants shows effects of both innate mechanisms and experience. *Nature*. 259:39-41.
- Tiitinen H, May P, Reinikainen K, Naatanen R. 1994. Attentive novelty detection in humans is governed by pre-attentive sensory memory. *Nature*. 372:90-92.
- Werker JF, Gilbert JH, Humphrey K, Tees RC. 1981. Developmental aspects of cross-language speech perception. *Child Dev*. 52:349-355.
- Yago E, Corral MJ, Escera C. 2001. Activation of brain mechanisms of attention switching as a function of auditory frequency change. *Neuroreport*. 12:4093-4097.