Supporting Information

Mahmoudzadeh et al. 10.1073/pnas.1212220110

SI Text

Subjects. Table S1 shows the clinical features of the 12 included infants. Ten additional infants were tested. No useful functional data were obtained because of fussiness ($n = 7$); poor positioning of the probe ($n = 2$); and hair obstruction ($n = 1$).

Optical Imaging System (Functional Near-Infrared System) We used a multichannel frequency-domain based optical imaging system (Imagent; ISS) for acquiring oxygenated hemoglobin and de-oxygenated hemoglobin concentration changes during the auditory stimulation task. The Imagent is a frequency-domain tissue spectrometer comprising 32 intensity-modulated laser diodes at two wavelengths ($\lambda = 690$ and 830 nm) coupled to optical fibers and four gain-modulated photomultiplier tube (PMT) detectors to collect the signal at both wavelengths separately. The modulation frequency of the laser intensity was 110 MHz, and the cross-correlation frequency for heterodyne detection was 5 kHz. The reflected light was collected in PMTs and demodulated. Its mean intensity (DC), modulation amplitude (AC), and phase were determined. The average output power of the lasers was ~0.5 mW, and the acquisition rate of the optical system was 9.1912 Hz (approximately one sample each 110 ms).

A special probe made from soft and flexible foam (thickness = 5 mm) was designed to comfortably hold the source and detector fibers on the infant's head. A patch comprising two detectors and eight emitters (optodes) entering orthogonally to the head surface was placed on each side of the head. Each of the eight optodes contained two wavelength emitter glass fibers (690 and 830 nm). Light in these two wavelengths is differently absorbed by oxy- and deoxyhemoglobine and thus allows computing their respective concentration in the medium. The average distance between the temporal cortex and the scalp is ~7 mm in premature infants. Because the light penetration depth is about half of the distance between sources and detectors, we designed a round-grid layout, keeping a distance of 15 mm between them (Fig. S1). This layout allowed 10 points of measures (called channels thereafter), which are simultaneously sampled on each hemisphere.

Experimental Procedure. Stimuli were binaurally presented at a comfortable hearing level (~70 dB) via speakers placed at the infant's feet, at a distance of 30 cm from his/her head. The infants were tested asleep at night to avoid the day light and the intense day activity of a neonatal care unit. They laid in a supine position on a comfortable pad in a dark and quiet incubator. The incubator was further defended against ambient light by dark sheets. The probe was smoothly secured to the infant's head to cover the perisylvian areas on each side with straps and foam padding (Fig. S1). The probe covered about 15 cm^2 = 5 (length) × 3 cm (width). For information, the interparietal distance is on average 7.8 cm at 31-wk gestational age (wGA) for a cranial perimeter of 29 vs. 9.4 cm at 41 wGA for a cranial perimeter of 35 cm (chart provided by the Collège Français d'échographie Foetale).

Because the optodes were attached to the two pieces of foam, their relative positions were fixed from subject to subject. Two anatomical markers (left and right ear points) were embedded into the probes to facilitate their correct orientations and positions and to enable the recovery of the probe orientation for data analysis. Location was checked before functional data acquisition.

The geometric layout of the optical probe and its projection on the brain mesh of an individual 30 wGA preterm infant [courtesy of P. Hüppi and J. Dubois, Department of Pediatrics, Geneva University Hospitals, Geneva, Switzerland (1)] are shown in Fig. S1. In this figure, the numbers 1–16 correspond to the light sources, and the letters A, B, C, and D correspond to the detectors. Using this notation, we denote the source-detector light channels as CH1, CH2,...CH20. The Yakovlean torque, a human structural characteristic bending the right hemisphere dorsally and forward relative to the left, is already observed in preterm infants (2, 3). This torque leads to a higher and steeper sylvian scissure in the right hemisphere than in the left. We therefore checked that our optodes placed relative to external landmarks were not at a different distance from the left and right posterior temporal cortices. J. Dubois measured the distance between the tragus and the higher posterior point of the sylvian fissure on MRI in 10 preterm infants (range, 26–32 wGA; mean age, 29.8 ± 1.7 wGA). This point is a robust cerebral landmark, easily observed in all subjects and hemisphere and is localized between channels 17–18 and 8–9. The distance was not correlated with age and remained stable during this period (4.4–4.9 cm), with no significant left–right asymmetries. We also measured the skull thickness on CT scan in preterm infants. At this age, it is <2 mm, and any asymmetry would have been too small to affect the measures. Thus, we can estimate that the errors of location of our points of measure relatively to the underlying cortical regions were negligible across our age range and that structural asymmetries were unlikely to affect our measures (4).

Data Processing. Oxy- (HbO) and deoxy-hemoglobin (Hb) are chromophores that absorb light at different wavelengths. The modified Beer Lambert law was applied to the two wavelengths signals (690 and 830 nm) to convert the signal intensities into relative changes in (de)oxy-hemoglobin concentration.

To reject artifacted signal, we used a z-score–based algorithm. Because individual features, such as thickness of skull and color of hair, affect the signal strength, the signal was first homogenized in each participant by computing a z-score across blocks for each channel and for each time sample. An artifact was assumed if the z-score was exceeding 4 during a window time-locked to the onset of the block [−5 to 25 s] in any channel. In that case, the entire time window covering a block was excised from the data for all channels.

The remaining cleaned HbO and Hb signals were band-pass filtered (0.03–0.5 Hz) using a zero-phase filter (Butterworth, order: 6) to eliminate physiological noise (e.g., slow drifts, arterial pulse oscillations). Both signals were segmented relative to the onset of each block (−5 to +25 s). A linear detrend and then a baseline correction on the 5 s preceding the onset of the block were performed on segments. Finally, the segments were averaged in each subject by condition for Hb and HbO [on average, 19, 19, and 21 blocks for standard (ST), deviant voice (DV), and deviant phoneme (DP) conditions, respectively].

Data Visualization. To visualize the location of the brain activity measured with our optical system, we used a realistic head model of a 30 wGA premature neonate, provided by J. Dubois and P. Hüppi (Department of Pediatrics, Geneva University Hospitals, Geneva, Switzerland) (see ref. 1 for a precise description of the steps to obtain a head and brain mesh from T2w magnetic resonance images). A virtual layout was applied on the head model using the same ears preauricular biomarkers as reference positions than in real infants. This model allowed determining the position of each virtual near-infrared spectroscopy (NIRS) acquisition channel relative to the preterm's brain (Fig. S1). Each pair of NIRS emitter and detector was then used to compute the optical signal propagation in the brain according to a photon migration model (5). This

step results in brain volume filling with information following the banana shape of the light propagation between emitters and each of the neighboring detectors, using the photon path modeled by the prescribed hitting density function. The obtained individual photon migration probability distributions were then combined together using a tricubic spatial interpolation step to consider the interaction between the individual distributions. For visualization purposes, the volume-based information was projected on the infant's brain areas covered by the optical probe with a color scale proportional to the amplitude of the response. We estimate to have obtained responses from the superior temporal gyrus, the supramarginal gyrus, the inferior part of the central sulcus, and the inferior frontal region comprising Broca's area, but note that some of the regions reported in brain imaging experiments using linguistic stimuli in infants (6) are missing, such as a large part of the superior temporal sulcus, the dorsolateral prefrontal region that was not covered by the probe, and the insula, which is too deep.

The mean signal obtained in each subject and each condition was averaged over the 12 infants (grand average) for [Hb] and [HbO]. At each time step (every ~110 ms) of the sampling frequency (9.19 Hz) of these grand averages, we computed a surface-based topographic color map as described above to create a video animation of the brain activity during the duration of the stimulation.

Statistical Analysis. We downsampled the signal to 1 Hz in the studied time window (−5 +25 s) to decrease the number of measures. This downsampling does not affect the quality of the measures because the metabolic response is slow. We limited our analyses to the HbO signal because its signal-to-noise ratio is more robust than Hb (7, 8).

Habituation-recovery paradigms are based on the observation that stimulus repetition induces a decrease in neural activity (repetition suppression), whereas a sudden change causes a recovery of the neural activity in regions coding the parameter that changed (9, 10). This is observed at the macroscopic level using event related potentials (ERPs) (9) but also at the neuronal level (11, 12). Although the exact relation between neural activity and the metabolic response measured by fMRI is not completely understood, repetition affects the amplitude of the hemodynamic response and modifies its temporal characteristics (i.e., slope of the response, latency to the peak, duration of the plateau) in adults (13–16) and infants (17, 18). We thus expected a weaker amplitude and slower response in standard blocks relative to deviants blocks if preterm infants perceive the change of stimulus (18). To probe discrimination responses in our preterms, we thus considered the two parameters, amplitude and dynamics of the response, and used the following strategy.

To deal with the problem of repeated measures (20 channels × 30 time bins = 600), we had to reduce the number of points for each analysis. An elegant solution proposed by Maris and Oostenveld (19), initially for electrophysiological recordings with dense arrays of captors, is to use cluster-based statistics. However, if this type of statistics can rigorously prove that a statistical effect is observed, it may blur the location and latency of this effect because a single common point between close effects in space of time is sufficient to create a single cluster. We have thus completed this analysis by the analysis of the area under the curve (AUC) to locate the channels showing the maxima of the effects while reducing the number of measures to 20 (one by channel) in each condition.

The dynamics of the response in each condition was evaluated by determining its onset relative to baseline in each channel and its global slope on each hemisphere. Because these analyses suggested a transient response for the deviant voice condition, not captured by the previous analyses, we reexamined the amplitude differences focused on the first 6 s of stimulation. We chose this window because it is the classical latency to reach plateau in adults, and it was compatible with our preterms' hemodynamics (Fig. S1). In deviant

blocks, at least one deviant trial was presented during this period and thus was expected to increase the metabolic needs.

Finally, we examined whether the younger infants already presented discrimination responses by dividing the group in two (younger and older than 31 wGA).

Below, we present our approach in more detail.

Cluster-based statistics on the amplitude of the response. To control the risk of false-positive or familywise error, due to multiple comparisons, we used nonparametric statistical tests as implemented in Fieldtrip (20), a toolbox developed in Matlab (available at http://fieldtrip.fcdonders.nl). This approach involved the following steps:

(*i*) Paired *t* tests between two conditions of interest, ST against DV condition, ST against DP condition, DV against DP condition, and left against right hemisphere, were performed at each channel (10 × 2 hemispheres) and time point (25 time points), with a threshold set at $P = 0.05$, two-tailed.

(*ii*) Channel-time points that exceeded the threshold were then grouped into connected clusters on the basis of their temporal and spatial proximity. Only the most adjacent channels were considered as neighbors, providing two neighbors by channel, except for the two most frontal channels (channels 1 and 11), which had only one posterior neighbor. Supra- and infra-sylvian channels were not considered as possible neighbors because the NIRS spatial resolution is sufficient to disentangle frontal and temporal responses, which are not expected to behave similarly. Each cluster was assigned a cluster value equal to the sum of each channel-time point *t*-value.

(*iii*) The type I error rate for the complete spatiotemporal set of channels and time points was controlled by evaluating the cluster-level statistics under the randomization null distribution of the maximum cluster value, which is obtained by the random permutations of the original conditions. Here we used the complete set of possible permutations: 4,095 corresponding to 2^{12} subjects − 1 (the original partition). The Monte-Carlo *P* value is estimated as the proportion of random partitions that result in larger cluster statistics than the observed one.

(*iv*) We report all clusters whose corrected *P* value was <0.025 (two-tailed test). For each cluster, we give the extremes of the time window and all channels present in the connected set of channels/time points, but note that it does not imply that a particular channel displayed significant difference during the entire window. For example, see channel 7 in Fig. 1: a single star is present at 10 s, but this sample is connected in time and space with channels 6, 3, 2, and 1.

AUC. The hemodynamic response can be characterized by two parameters: the amplitude and the duration of the response. Computing the AUC provides information reflecting the accumulative variation of HbO concentration during the whole stimulation period. Individual AUCs of the hemodynamic response were thus evaluated for each channel and each condition using only the positive part of the signals. For each comparison between two conditions, we evaluated the probability of a larger *t*-statistic in our observed comparison relative to the distribution of the *t*-values obtained when the complete set of random partitions of our data was used (here 4096-1). We used the Holm's method to correct our *P* value for false positives due to multiple comparisons (20 channels). In this method, the *P* values are ordered and examined in a stepwise fashion until the null hypothesis can no longer be rejected. In our case, the first significant *P* value had to be inferior to 0.05/20 (0.0025) to be accepted, and then 0.05/19 (0.0026), etc.

Evaluation of the activation dynamics. In each channel and condition, we evaluated the latency of the first sample showing a significant amplitude increase relative to the 5-s baseline noise. This eval-

uation was performed using individual signals and a statistical *t* test to compare the amplitude at each time point after the beginning of the auditory stimulation against the baseline mean amplitude. To determine each channel onset time significance, we used a corrected significance level at $P_{cor} < 0.0017$ (i.e., 0.05/30 time points) and report the first of at least two significant samples. We completed this analysis by an analysis of the slope of the HbO signal from onset of the auditory stimulation to maximum peak. For each subject, the slope of the HbO signal of all NIRS channels over one hemisphere was computed based on the following formula:

$$\text{Slope} = \Delta\text{HbO}/\Delta\text{time} = (\text{HbO}_{peak} - \text{HbO}_{stim.onset})/(\text{time}_{peak} - \text{time}_{stim.onset})$$

t test paired comparisons between the three conditions taken two by two were performed over the left and right channels averaged together on each side.

Analyses restricted to block onsets. Because the previous analysis showed a dynamic initially similar for both deviant conditions at least over the right hemisphere, we restricted our analysis of the deviant voice condition to the first 6 s of the blocks. We thus performed a paired *t* test between standard and deviant voice conditions on the mean signal averaged across the first six samples of the stimulation period using permutations and the Holm's method to correct for multiple comparisons across the 20 channels. As a follow-up, we examined the [DP–ST] and [DP–DV] contrasts in the channels showing a deviant voice effect in the previous analysis and in the homologous channels on the controlateral hemisphere, because voice and phonetic processing are differently lateralized in adults (21).

Comparison between youngest and oldest infants. Finally, we evaluated whether there was maturational change during our time window and separated our group in two subsets of six infants: younger and older than 31 wGA at birth (mean age at test, 29w4d vs. 31w6d GA). We used the same cluster-based statistics approach than above on the [DP–ST] difference and [DV–ST] difference. We also performed an ANOVA on the mean of the first 6 seconds of the block with age as a between-subject factor and conditions (3 levels), channels (10 levels), and hemisphere (2 levels) as within-subject factors. We finally performed analyses of the correlation between age and the discrimination responses.

1. Dubois J, et al. (2008) Mapping the early cortical folding process in the preterm newborn brain. *Cereb Cortex* 18(6):1444–1454.
2. Chi JG, Dooling EC, Gilles FH (1977) Gyral development of the human brain. *Ann Neurol* 1(1):86–93.
3. Dubois J, et al. (2008) Primary cortical folding in the human newborn: An early marker of later functional development. *Brain* 131(Pt 8):2028–2041.
4. Patil AV, Safaie J, Moghaddam HA, Wallois F, Grebe R (2011) Experimental investigation of NIRS spatial sensitivity. *Biomed Opt Express* 2(6):1478–1493.
5. Sassaroli A, Tong Y, Frederick BB, Renshaw PF, Fantini S (2006) Calculations of BOLD signals by use of NIRS photon migration hitting density functions. *Biomedical Optics, Technical Digest* (Optical Society of America, Fort Lauderdale, FL), paper ME30.
6. Dehaene-Lambertz G, Dehaene S, Hertz-Pannier L (2002) Functional neuroimaging of speech perception in infants. *Science* 298(5600):2013–2015.
7. Bortfeld H, Wruck E, Boas DA (2007) Assessing infants' cortical response to speech using near-infrared spectroscopy. *Neuroimage* 34(1):407–415.
8. Hoshi Y, Kobayashi N, Tamura M (2001) Interpretation of near-infrared spectroscopy signals: A study with a newly developed perfused rat brain model. *J Appl Physiol* 90(5):1657–1662.
9. Dehaene-Lambertz G, Dehaene S (1994) Speed and cerebral correlates of syllable discrimination in infants. *Nature* 370(6487):292–295.
10. Wacongne C, Changeux JP, Dehaene S (2012) A neuronal model of predictive coding accounting for the mismatch negativity. *J Neurosci* 32(11):3665–3678.
11. Miller EK, Gochin PM, Gross CG (1991) Habituation-like decrease in the responses of neurons in inferior temporal cortex of the macaque. *Vis Neurosci* 7(4):357–362.
12. Ulanovsky N, Las L, Nelken I (2003) Processing of low-probability sounds by cortical neurons. *Nat Neurosci* 6(4):391–398.
13. Kruggel F, von Cramon DY (1999) Temporal properties of the hemodynamic response in functional MRI. *Hum Brain Mapp* 8(4):259–271.
14. Thierry G, Ibarrola D, Démonet JF, Cardebat D (2003) Demand on verbal working memory delays haemodynamic response in the inferior prefrontal cortex. *Hum Brain Mapp* 19(1):37–46.
15. Dehaene-Lambertz G, et al. (2006) Functional segregation of cortical language areas by sentence repetition. *Hum Brain Mapp* 27(5):360–371.
16. Sigman M, Jobert A, Lebihan D, Dehaene S (2007) Parsing a sequence of brain activations at psychological times using fMRI. *Neuroimage* 35(2):655–668.
17. Dehaene-Lambertz G, et al. (2006) Functional organization of perisylvian activation during presentation of sentences in preverbal infants. *Proc Natl Acad Sci USA* 103(38): 14240–14245.
18. Dehaene-Lambertz G, et al. (2010) Language or music, mother or Mozart? Structural and environmental influences on infants' language networks. *Brain Lang* 114(2): 53–65.
19. Maris E, Oostenveld R (2007) Nonparametric statistical testing of EEG- and MEG-data. *J Neurosci Methods* 164(1):177–190.
20. Oostenveld R, Fries P, Maris E, Schoffelen JM (2011) FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput Intell Neurosci* 2011:156869.
21. Zatorre RJ, Belin P (2001) Spectral and temporal processing in human auditory cortex. *Cereb Cortex* 11(10):946–953.

Fig. S1. (*Upper*) Schematic representation of the location of the optical probe on an infant's head relatively to anatomical landmarks. (*Lower*) Estimated projection of the optodes on the brain of a 30 wGA preterm infant [courtesy of Petra Huppi and Jessica Dubois (1)]. The eight emitters (red circles) are arranged on two 1.5-cm-diameter circles centered by the two detectors (black circles), creating 10 points of measure (channels) over each hemisphere (blue circles).

Fig. S2. Time course of the grand average of Hb (blue line) and HbO (red line) across all conditions. The *y* axis represents the concentration changes in arbitrary unit (a.u.). The *x* axis displays the time in seconds. The black rectangle bar corresponds to the duration of a block. (*A*) RH, right hemisphere; (*B*) LH, left hemisphere.

Fig. S3. AUC for the different conditions. Comparisons between conditions are performed two by two in each channel. *P < 0.05, **P_{cor} < 0.05. Channels 1–10 are over the right hemisphere (RH) and 11–20 are over the left hemisphere (LH).

Fig. S4. Individual responses in each condition at three different locations in each hemisphere. The black rectangle corresponds to the period of stimulation (20 s). a.u., arbitrary unit.

Fig. S5. Comparison of the deviant phoneme and deviant voice conditions vs. standard. The time courses of the grand average of Hb0 in deviant phoneme (DP), deviant voice (DV), and standard (ST) blocks are plotted for each channel. The black rectangle indicates the duration of the stimulation block and the stars the *t*-value of the samples (right scale) included in significant clusters as determined by a cluster-based analysis.

Fig. S6. Comparison of the DP condition vs. ST and DV: same legend as Fig. S4. The pink stars indicate the *t*-value of the samples (right scale) included in significant clusters as determined by a cluster-based analysis in the DP vs. DV comparison ($P_{cor} < 0.025$).

Fig. S7. Time course of the grand average of Hb0 in deviant phoneme (DP), deviant voice (DV), and standard (ST) blocks, split in function of age (mean age, 29w4d GA for the six youngest and 31w6d for the six oldest). There is no significant difference between the two groups (cluster based statistic on the whole segment). An ANOVA performed on the mean of the first 6 seconds of the block with age as a between-subject factor and conditions (3 levels), channels (10 levels), and hemisphere (2 levels) as within-subject factors did not reveal any main effect of age [$F(1,10) < 1$] nor any significant interaction of age with any combination of the other factors [$P > 0.5$, except for the triple interaction hemisphere × channel × age: $F(9,90) = 2.96$, $P = 0.004$]. The youngest group already reacted to a change of phoneme [main effect of condition: $F(1,5) = 14.6$, $P = 0.012$, condition × channel $F(9,45) = 6.57$, $P < 0.0001$] and a change of voice [main effect of condition: $F(1,5) < 1$ but condition × channel $F(9,45) = 3.72$, $P = 0.001$].

Table S1. Clinical features of the tested infants

Infant no.	Sex	GA at birth (wk)	GA at test (wk)	Birth weight (g)	EEG	Apgar (1 min)	Apgar (5 min)	Brain US	Delivery	Presentation	Clinical conditions (etiology)
1	M	32	32 2/7	1,530	N	9	7	N	Vaginal	Transverse	Twin
2	F	28 1/7	28 5/7	1,490	N	9	10	N	Cesarian	Cephalic	RPH
3	M	32	32 2/7	1,550	N	0	6	N	Cesarian	Cephalic	Preeclampsia
4	M	32	32 3/7	2,300	N	10	10	N	Vaginal	Cephalic	PROM
5	M	30 6/7	31 1/7	990	N	8	9	N	Cesarian	Cephalic	Preeclampsia
6	F	31	31 4/7	1,260	N	10	10	N	Cesarian	Cephalic	Preeclampsia
7	F	31	31 2/7	1,560	N	8	10	N	Cesarian	Breech	Anoxo-ischemia
8	M	29	29 6/7	1,610	N	10	10	N	Vaginal	Cephalic	Twin, Preeclampsia, PROM
9	M	31 1/7	31 4/7	2,300	N	10	10	N	Vaginal	Breech	Twin
10	F	28	28 3/7	995	N	8	8	N	Vaginal	Cephalic	Twin, Preeclamsia
11	M	28	28 4/7	760	N	7	10	N	Cesarian	Cephalic	Preeclampsia
12	M	30 2/7	30 5/7	1,440	N	9	10	N	Vaginal	Cephalic	PROM, Chorioamniotis

Brain US, brain ultrasonography; F, female; M, male; N, normal; PROM, premature rupture of membranes; RPH, retroplacental hematoma.

Table S2. Post hoc analyses of frontal responses during the first 6 s of stimulation in preterms divided relative to their birth age (inferior or superior to 31 wGA)

Channel	<31 wGA		≥31 wGA	
Right Ch 5 (DV vs. ST)	$F(1,5) = 10.8$	$P = 0.022$	$F(1,5) = 10.44$	$P = 0.023$
Left Ch 12 (DP vs. ST)	$F(1,5) = 8.5$	$P = 0.036$	$F(1,5) = 6.9$	$P = 0.047$
Ch 5 vs. Ch 12 * DP vs. DV	$F(1,5) = 7.9$	$P = 0.037$	$F(1,5) = 5.79$	$P = 0.061$
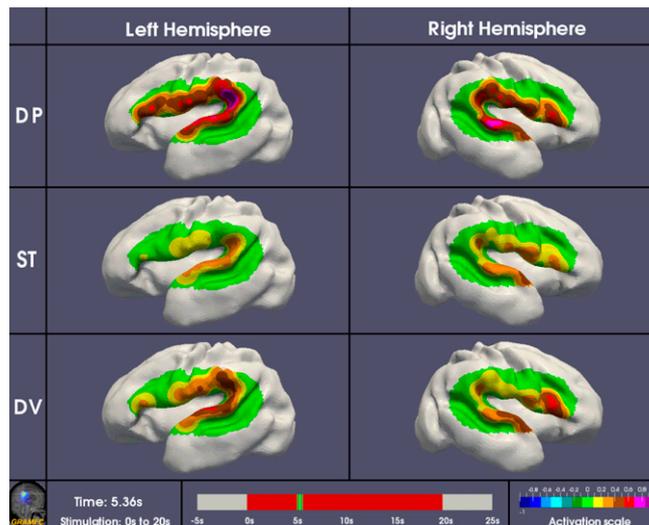
Movie S1. Infant's hemodynamic response in the three types of blocks. The red rectangle indicates the time window during which auditory stimuli were presented to infants.

Movie S1