Original Article

Two Independent Response Mechanisms to Auditory Stimuli Measured with Functional Near-Infrared Spectroscopy in Sleeping Infants

Trends in Hearing
Volume 28: 1–11

DOI: 10.1177/23312165241258056
journals.sagepub.com/home/tia
Sage

Onn Wah Lee[1,2,3], Darren Mao[1,2], Julia Wunderlich[1,2], Gautam Balasubramanian[1,2], Mica Haneman[1], Mikhail Korneev[1] and Colette M. McKay[1,2]

Abstract

This study investigated the morphology of the functional near-infrared spectroscopy (fNIRS) response to speech sounds measured from 16 sleeping infants and how it changes with repeated stimulus presentation. We observed a positive peak followed by a wide negative trough, with the latter being most evident in early epochs. We argue that the overall response morphology captures the effects of two simultaneous, but independent, response mechanisms that are both activated at the stimulus onset: one being the obligatory response to a sound stimulus by the auditory system, and the other being a neural suppression effect induced by the arousal system. Because the two effects behave differently with repeated epochs, it is possible to mathematically separate them and use fNIRS to study factors that affect the development and activation of the arousal system in infants. The results also imply that standard fNIRS analysis techniques need to be adjusted to take into account the possibilities of multiple simultaneous brain systems being activated and that the response to a stimulus is not necessarily stationary.

Keywords

infant, hearing, fNIRS, sleeping, arousal

Received 6 June 2023; Revised received 7 March 2024; accepted 7 May 2024

Introduction

Functional near-infrared spectroscopy (fNIRS) is an emerging neuroimaging technique to study functional hearing in the infant population because it is baby-friendly, low-cost, silent, portable, and less susceptible to motion and electrical artifacts (Pinti et al., 2020; Quaresima & Ferrari, 2019). fNIRS measures local changes in the concentration of oxy-(HbO) and deoxyhemoglobin (HbR) due to neurovascular coupling. Most fNIRS studies of speech perception in infants have reported a canonical-shaped response morphology, where the HbO increases from the baseline, the HbR decreases from the baseline, and both return to baseline within a predictable period post-stimulus onset (Bortfeld et al., 2007; Lloyd-Fox et al., 2019; Mao et al., 2021; Nakano et al., 2009; Wilcox et al., 2005). However, some studies have reported an inverted response shape (Cabrera & Gervain, 2020; Gervain et al., 2012; Zhang et al., 2022), and a review has suggested that the differences arise from the variation in experiment design and stimulus complexity (Issard & Gervain, 2018). The variable morphology of stimulus-evoked fNIRS responses in infants limits the use of the existing inference framework, which is based on adult data. The framework assumes that the canonical-shaped response to a stimulus is stationary throughout the experiment (Ye et al., 2009), an assumption that we found to be violated in our data from sleeping infants. In our pilot study, we observed that the speech-evoked responses were neither canonical nor reliably inverted, and they did not return to baseline within a relatively short experimental epoch as predicted from a

[1]Bionics Institute, Victoria, Australia
[2]Medical Bionics Department, University of Melbourne, Victoria, Australia
[3]Centre for Rehabilitation & Special Needs Studies, Faculty of Health Sciences, Universiti Kebangsaan Malaysia, Kuala Lumpur, Malaysia

Corresponding author:
Onn Wah Lee, Bionics Institute, East Melbourne, VIC 3002, Australia.
Email: leeonnwah@ukm.edu.my

canonical response. We, therefore, extended the inter-block silence interval to more than 22 s, allowing us to further investigate this unexpected morphology.

In this study, we describe the morphology of fNIRS response recorded from 16 naturally sleeping infants (infant shown in Figure 1A, recording montage shown in 1B). Each participant was presented with a 5.4 s speech stimulus block, consisting of 12 concatenated repetitions of a 450 ms long "ba" speech token; the stimulus block was repeated 20 times and separated by a silence period randomized between 22.0 and 32.0 s (Figure 1C). The morphology of fNIRS responses for groups of five sequential trials was first investigated, to see how the morphology might change over the course of the 20 epochs. We also evaluated the morphological differences in different regions of interest (ROIs) and between infants who listened to the stimuli through either the left or right ear. After discovering systematic changes in morphology with experiment duration, we hypothesized that the response morphology seen must be due to a sum of two independent and simultaneous responses evoked by the auditory stimulus. Using independent component analysis (ICA), we show that the data are consistent with the latter hypothesis and argue that the two independent responses reflect two independent mechanisms: an obligatory auditory response to auditory stimulation and a response related to activation of the arousal system during sleep. The findings of this study not only changes how the inference framework of fNIRS responses can be applied to sleeping infants, but also provides insight into multiple physiological brain processes that occur in response to speech stimuli during sleep.

Material and Methods

Participants

A total of 25 infants were recruited for this study, but 9 were not included in the final analysis. Two infants did not pass the tympanometry test on the test day, five infants did not sleep in the lab, and two infants were excluded due to poor cap placement by the end of the experiment. The final 16 infants (mean age = 7.4 months, standard deviation = 3.7 months; female = 8) passed the newborn hearing screening or had normal diagnostic audiology results, had aerated middle ears as measured by tympanometry on the test ear, and had no general medical condition as reported by the parent. This study was carried out in accordance with the Declaration of Helsinki, was approved by the Royal Children's Hospital human ethics committee (S/N 71941), and consent was obtained from a parent.

Stimuli

The stimulus used in this study was a natural recording of the "ba" speech token narrated by a native Australian female speaker. The recording was made in a sound-treated room, using an AT2020USB + microphone, sampled at 44.1 kHz, and 16-bit resolution. Multiple recordings of the "ba" syllable were made. Each of them was trimmed to 0.45 s, equalized to the root mean square of an 80 dB SPL 1000 Hz warble tone, and three audiologists agreed on one best-quality speech token. The intensity of 80 dB SPL was used for the stimulus equalization only and is not a presentation level in the experimental procedure. A pool of nine versions of the speech token was created by shifting the $F_0 \pm 1$ Hz using Praat (Boersma & Van Heuven, 2001) and jittering the intensity by ±2 dB SPL. A 5.4 s stimulus block of repeating "ba" sounds was then created by concatenating 12 speech tokens randomly drawn from the pool of speech tokens.

fNIRS Set-up

The fNIRS data were recorded using the NIRScout device manufactured by NIRx (NIRx Medical Technologies, LCC). This system emits near-infrared lights at wavelengths

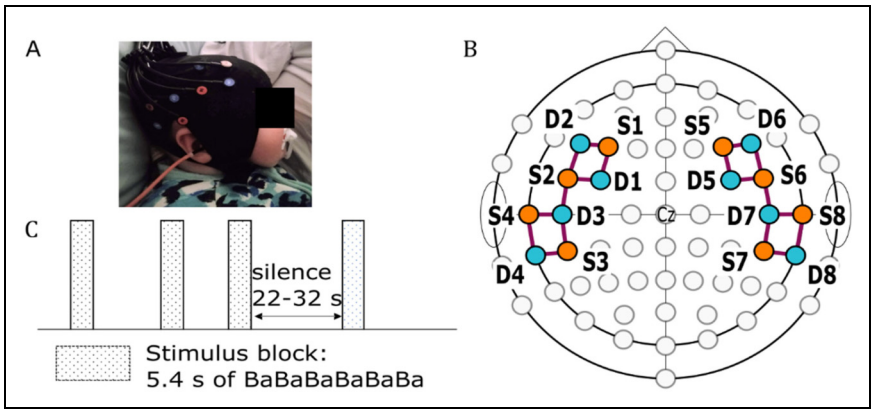

Figure 1. fNIRS optode arrangement and stimulus presentation protocol. (A) A sleeping infant with a mounted fNIRS cap. (B) The arrangement of sources (orange circles) and detectors (teal circles) positioned over F3/4, F5/6, FC3/4, FC5/6, C5/6, T7/8, CP5/6, and TP7/8 according to the International 10–10 system. The purple lines represent channels formed by the source-detector pairs. (C) The stimulus presentation protocol used in this study.

of 760 and 850 nm through LED optodes. We used flat optodes designed for use in young babies, which were arranged in two symmetrical square patterns on each hemisphere of the head (Figure 1B), forming a total of 18 channels. Each source and detector making up a channel were separated by a distance between 2 to 3 cm, which has been shown to be effective in recording the fNIRS response to auditory stimulus in infants (Taga et al., 2007). We did not use short channels because, for infants, an effective short channel length of 0.2 cm (that does not capture and remove response signals) is infeasible and it has been shown that the signal to noise in infant data does not generally require short channel analysis (Brigadoi & Cooper, 2015).

The channels were segmented into two ROI in each hemisphere, named prefrontal and temporal. We estimated the average specificity of each ROI in our montage to record activity from the inferior frontal gyrus and superior temporal gyrus, which play an important role in speech sound processing (van der Burght et al., 2023; Yi et al., 2019), using the devFOLD toolbox (Fu & Richards, 2021; Morais et al., 2018). According to the AAL3 atlas (Rolls et al., 2020) and 7.5 months old infant-specific channel information provided by the toolbox, the average specificity of each ROI for recording responses from either the inferior frontal gyrus or superior temporal gyrus was higher than 10%, as shown in Appendix A.

Procedures

The fNIRS recording started with 5 min of silence. Each participant listened to 20 trials of the stimulus block with inter-block silence intervals randomized between 22 and 32 s (Figure 1C). The stimuli were presented monaurally at 65 dB SPL through an ER3A insert tubephone placed either in the right ($n = 8$) or the left ear ($n = 8$). Participants with the tubephone in the right ear are referred to as the right test-ear group, while those with the tubephone in the left ear are referred to as the left test-ear group. All participants were tested while asleep in a dimly lit room, with only a table lamp on. If the infant woke up at any time during the experiment, the recording was terminated.

Data Preprocessing

The data were preprocessed offline using the NIRS Brain AnalyzIR toolbox (Santosa et al., 2018) and a custom script. The raw light intensity data was first converted to optical density. The motion artifacts were then corrected using the Temporal Derivative Distribution Repair (Fishburn et al., 2019) and channels with Scalp Coupling Index (SCI) threshold below 0.75 were removed (Pollonini et al., 2014). Across all participants, a total of 5.6% of channels were discarded. The optical data were converted to the hemodynamic concentration change, oxyhemoglobin (HbO) and de-oxyhemoglobin (HbR), using the modified Beer–Lambert Law with the default partial path length factor of 0.1. To remove systemic physiological noises in the data such as heart rate and respiratory rate, the data were bandpass filtered between 0.01 and 0.25 Hz using a Butterworth filter. The cleaned hemodynamic concentration change data were epoched between −3 and 27.4 s from stimulus onset and baseline corrected to the average of −3 and 0 s from stimulus onset.

Results

The Morphology of the fNIRS Response in Sleeping Infants Changes During the Experiment

Figure 2 shows the group ($n = 16$) average HbO response within each ROI averaged across every five sequential trials. It can be seen that, in each ROI, the morphology of the five trial averages changed over the course of the session. We observed a positive peak around 5.0–6.0 s, as expected from a standard canonical response, followed by a wide negative trough peaking around 10.0–20.0 s—well

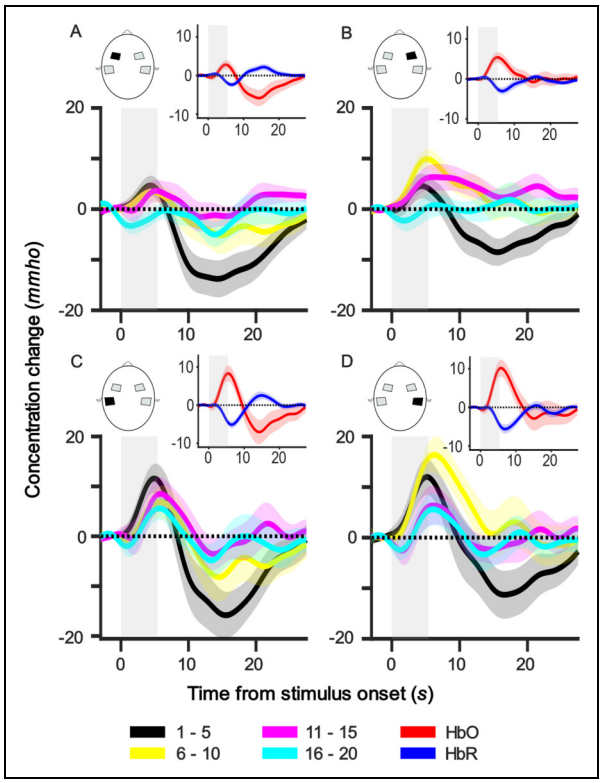

Figure 2. Group-level average HbO responses for every five trials in each ROI. Each panel shows the responses of one ROI, and the ROI is indicated in the head illustration. The subfigure at the top right corner of each panel illustrates the group-averaged HbO (red) and HbR (blue) responses of the respective ROI over the whole session. Grey bars represent the stimulus duration. Shaded regions represent one standard error of mean (SEM). mmho = mili mho.

after the stimulus offset at 5.4 s. However, the negative trough is dominant only in the first five trials (black line) and adapts over subsequent trials. Supplemental Figure 1 shows the group average HbO response per channel averaged across trials (Appendix B).

Figure 3 shows the average HbO responses for each test-ear group ($n = 8$) averaged across all trials and channels within each ROI. It can be seen that the morphology in each test-ear group is similar, but there are possible amplitude differences that are explored further in Effect of Test Ear, ROI, and Trial Number on the ICA Reconstructed Components section.

Often, a negative trough after a positive peak has been modeled as an undershoot (Friston et al., 1998) as the HbO returns to baseline: that is, it forms part of the same response to the stimulus as the positive peak. In such a model, the negative trough is assumed to be a passive recovery from the expansion of the blood capillary following the positive peak, as suggested by the balloon model (Buxton et al., 1998) in the blood oxygen level-dependent (BOLD) response of functional magnetic resonance imaging (fMRI). However, the pattern observed in our data is inconsistent with this explanation. If the negative trough is caused by an overshoot of recovery from capillary expansion, the magnitude of the trough would be correlated with the size of the positive peak. In contrast, the negative trough is initially much larger than the positive peak, has a longer latency than feasible for an overshoot, and reduces in size rapidly over trials independently of the size of the positive peak. The data are more consistent with the notion that two independent responses occur simultaneously: one being the positive canonical response to the stimulus and the other being a long-latency negative response that rapidly adapts over trials.

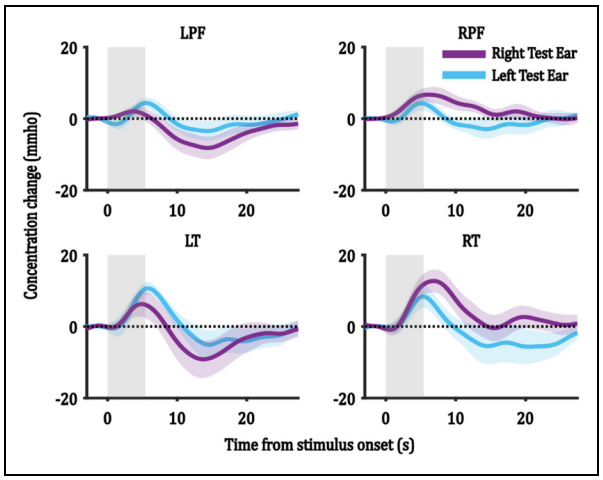

Figure 3. Group-level HbO responses for the right and the left test-ear groups in each ROI. Each panel shows the HbO responses of one ROI, and the ROI is indicated in the title of each subfigure. R/LPF = right/left prefrontal ROI and R/LT = right/left temporal ROI. Grey bars represent the stimulus duration. Shaded regions represent one SEM.

Evidence to support the notion of two simultaneous responses is provided by the fact that the latency of the positive peak in Trials 1–5 is shorter than in the later trials, which can be explained by the positive peak being partially canceled by its summation with the large negative response as they overlap in time.

In the next section, we perform further analysis to explore whether more than one independent neuronal response is associated with auditory processing in sleeping infants without imposing any prior assumptions on the characteristics of the independent responses. Specifically, we aimed to determine whether there were two independent responses consisting of one positive and one negative response, and how these two responses varied with test ear, ROI and the duration of the experiment.

Identifying Independent Responses in Auditory-Evoked fNIRS Responses

Independent component analysis (ICA) is a mathematical method used to identify independent signal sources common to a set of multiple recordings (Stone, 2002). In fNIRS studies, ICA has been used to separate noise from data and to extract neuronal activity-related sources (Kohno et al., 2007; Zhang et al., 2010; Zhao et al., 2021), by using each channel as an independent measurement. In this study, we applied ICA to identify independent signal components, or common signal sources, that contribute to auditory evoked fNIRS responses in sleeping infants. ICA imposes no a priori assumptions on the size, direction, or shape of response function components, making it an ideal nonbiased technique for identifying unknown independent components. Since we wanted to capture how the components may change in amplitude over time or ROIs, we used each epoch as well as each channel as the measurements input into ICA, assuming that the regions and epochs contain the same common components but with differing weights. By identifying these independent components, we can gain potential insight into the underlying neural processes associated with auditory processing in sleeping infants.

To perform ICA, we first detrended each epoch by subtracting a linear fit to the average of the first and last 3 s of the epoch. Subsequently, we averaged the epochs across participants in each test-ear group for each channel and trial. We then applied the default reconstruction ICA (*rica*) function from the Statistics and Machine Learning Toolbox in MATLAB (Version R2021a, MathWorks Inc.). Two independent components with large average weights were identified for each recorded epoch. The *rica* function outputs the component shapes and the weight of each independent component's contribution to each epoch/channel. We then reconstructed the independent components for each epoch/channel to their true size and sign by multiplying the components by their respective weights.

Figure 4 shows the average of the reconstructed ICA components for each test-ear group, averaged across all epochs/channels. Both test-ear groups yielded similar morphologies of the ICA-extracted components. These components consisted of a positive component (red lines in Figure 4) with early latencies ranging from 5 to 7 s from stimulus onset, and a negative component (blue lines in Figure 4) with late latencies ranging from 15 to 17 s from stimulus onset. In the following analysis, we use these two components and their weightings to analyze the effect of test ear, ROI, and trial number.

Effect of Test Ear, ROI, and Trial Number on the ICA Reconstructed Components

To investigate the effect of test ear, ROI, and trial number on the reconstructed ICA components, we extracted the peak amplitude (i.e., the weighting provided by ICA analysis) of the two reconstructed components for each epoch/channel. For each test-ear group and ICA component, these amplitudes were then averaged across channels within each ROI, and grouped every five trials. A three-way analysis of variance (ANOVA) was performed for each ICA component. The factors included test-ear group (left/right), ROI (left/right prefrontal ROIs and left/right temporal ROIs), and trial segment (Trials 1–5, 6–10, 11–15, and 16–20), with average amplitude as the dependent variable. Residual analysis was performed to test the assumptions of each three-way ANOVA. Normality was assessed using Shapiro–Wilk's normality test and homogeneity of variances was assessed by Levene's test. Residuals were normally distributed ($p > 0.05$) and variances were homogeneous ($p > 0.05$). Post hoc analysis was evaluated using the Tukey test with a family error rate of 0.05.

The Effect of ROI, Test Ear, and Trial Segment on the Positive HbO Component. Table 1 shows the result of the three-way ANOVA analysis for positive component, with post hoc analysis shown in Table 2. The analysis revealed a significant main effect of ROI with average amplitudes in the left/right temporal ROIs significantly larger than those in the left prefrontal ROI (Figure 5A), and a significant main effect of trial segment with average amplitudes in Trials 1–5 and 6–10 significantly larger than those in Trials 16–20 (Figure 5B). No significant main effect of test-ear group was observed, and no statistically significant interaction was observed among any combination of the three factors.

Given the absence of a significant main effect related to the test-ear group factor, the amplitudes of the positive component were subsequently pooled across the test-ear group to illustrate the effects of trial segment (Figure 5A) and ROIs (Figure 5B). It can be seen that the average amplitudes in the left/right temporal ROIs were larger than the left/right prefrontal ROIs (Figure 5B). Besides that, the average amplitudes in the right hemisphere were larger than those in the left hemisphere. To evaluate the effect of hemispheres (left/right) and functional regions (prefrontal/temporal), we computed a further two-way ANOVA with hemispheres and functional regions as factors and the average amplitudes as the dependent variable. This further analysis showed a significant main effect of functional region with average amplitudes in the temporal region larger than that of prefrontal region, with no significant main effect of hemispheres and no significant interaction between the two factors (Table 3).

The Effect of Tested Factors on the Negative Component. For the negative component, the three-way ANOVA analysis (Table 4) revealed a significant main effect of trial segment, a significant main effect of ROI and no significant main effect of test ear. Since there was, however, a significant interaction between test ear and ROI, the effect of ROI was analyzed further for each test-ear group separately, which showed that an effect of ROI was only present in the right test-ear group (Table 5).

Figure 6 illustrates the average amplitude of the negative component in each trial segment, averaged across ROIs and test-ear groups. The average amplitude of the negative

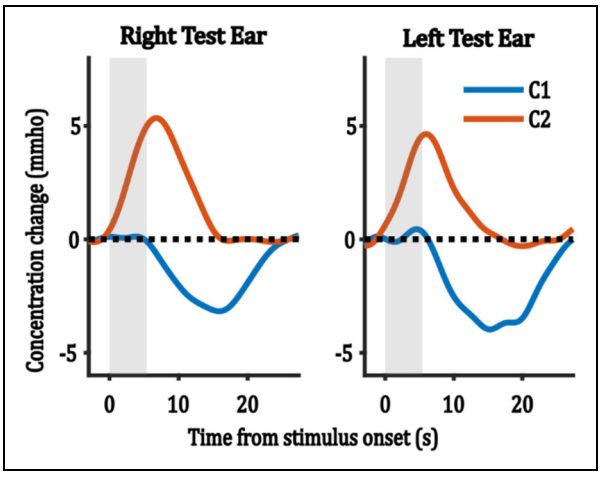

Figure 4. The average reconstructed ICA components across all epochs/channels for each test-ear group. C denotes component and grey bars represent stimulus duration.

Table 1. The Result of Three-way ANOVA for the Positive HbO Component, with Test ear, ROI, and Trial Segment Factors, and Average Amplitude as the Dependent variable.

	df	F-value	*p*-value
Positive component			
Test ear	1	0.25	0.616
ROI	**3**	**5.57**	**0.001****
Trial segment	**3**	**5.60**	**0.001****
Test ear: ROI	3	2.32	0.078
Test ear: Trial segment	3	1.19	0.317
ROI: Trial segment	9	0.96	0.473
Test ear: ROI: Trial segment	9	1.24	0.276
Residuals	128		

Note. The boldfaced numbers indicate statistically significant results.
*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

Table 2. Post Hoc Analysis for the Main Effect of ROI and Trial Segment on the Average Amplitude of the Positive Component.

Group A	Group B	A-B	Adjusted *p*-value
Main effect of Trial segment			
Trial 6–10	Trial 1–5	1.732	0.815
Trial 11–15	Trial 1–5	−1.833	0.787
Trial 16–20	**Trial 1–5**	**−5.931**	**0.016***
Trial 11–15	Trial 6–10	−3.565	0.271
Trial 16–20	**Trial 6–10**	**−7.663**	**<0.001*****
Trial 16–20	Trial 11–15	−4.097	0.164
Main effect of ROI			
LT	**LPF**	**5.931**	**0.016***
RPF	LPF	3.114	0.391
RT	**LPF**	**7.439**	**0.001****
RPF	LT	−2.818	0.481
RT	LT	1.508	0.869
RT	RPF	4.325	0.129

Note. The boldfaced numbers indicate statistically significant results. The analysis was multiple comparisons based on Tukey's procedure with a family error rate of 0.05.
*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

component was largest (most negative) in Trials 1–5, decreased to close to zero in Trials 6–10, and remained stable thereafter. Multiple comparisons analysis confirmed that the average amplitude in Trials 1–5 was significantly larger than those in Trials 6–10, 11–15, and 16–20 (Table 6).

Figure 7 shows the average amplitude of the negative component in each ROI, averaged across trials, for each of the right and the left test-ear groups. In the right test-ear group (Figure 7A), post hoc analyses showed that the average amplitudes in the left prefrontal ROI were significantly larger than those in the right prefrontal ROI, and the average amplitudes in the left temporal ROI were significantly larger than those in both the right prefrontal and right temporal ROIs (Table 7). Conversely, in the left test-ear group, no significant differences in average amplitude were observed between the ROIs (Figure 7B).

Similar to the positive component, for each test-ear group, we further evaluated the effect of hemispheres and functional region on the average amplitudes of the negative component using the two-way ANOVA (Table 8). For the right test-ear group, results showed a significant main effect of hemispheres with the average amplitudes in the left hemisphere larger than (more negative) those in the right hemisphere. No significant main effect of functional region and no significant interaction between factors were observed. For the left test-ear group, no significant effect was observed.

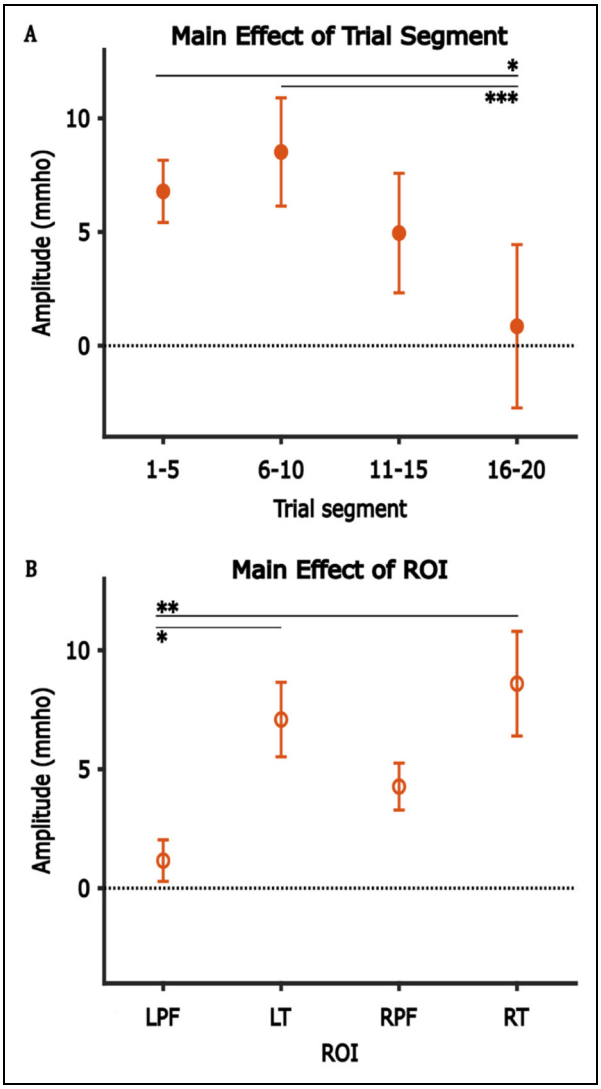

Figure 5. The main effect of trial segment (A) and ROI (B) on the amplitude of the positive component. Error bars represent one SEM. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

Table 3. The Results of Two-Way ANOVA for the Positive Component, with Hemisphere (Hem) and Functional Region (FR) Factors, and Average Amplitude as the Dependent Variable.

	df	F-value	*p*-value
Hem	1	2.458	0.119
FR	**1**	**12.109**	**<0.001*****
Hem:FR	1	0.297	0.587
Residuals	156		

Note. The boldfaced numbers indicate statistically significant results.
***$p < 0.001$.

Discussion

Two Simultaneous Speech-Evoked Brain Responses in Sleeping Infants

The ICA analysis showed that our measured overall HbO responses are consistent with the hypothesized sum of two simultaneous and independent responses: a positive response that peaked around 5–7 s from stimulus onset; and a wide negative response that peaked around 15–17 s from stimulus onset and rapidly adapted in amplitude across trials.

Table 4. The Results of Three-way ANOVA Showing the Effect of ROI, Trial Segment, and Test Ear on the Negative HbO Component.

	df	F-value	*p*-value
Negative component			
Test ear	1	0.48	0.489
ROI	**3**	**3.70**	**0.014***
Trial segment	**3**	**15.78**	**< 0.001*****
Test ear: ROI	**3**	**3.31**	**0.022***
Test ear: Trial segment	3	0.36	0.784
ROI: Trial segment	9	0.53	0.849
Test ear: ROI: Trial segment	9	0.97	0.472
Residuals	128		

Note. The boldfaced numbers indicate statistically significant results.
*$p < 0.05$, ***$p < 0.001$.

Table 5. The Results of Two-Way ANOVA for the Negative Component in Each Test Ear, with ROI, and Trial Segment Factors, and Average Amplitude as the Dependent Variable.

	df	F-value	*p*-value
Right test ear			
ROI	**3**	**5.582**	**0.002****
Trial segment	**3**	**8.303**	**<0.001*****
ROI: Trial segment	9	0.942	0.496
Residuals	64		
Left test ear			
ROI	3	0.252	0.860
Trial segment	**3**	**7.703**	**< 0.001*****
ROI: Trial segment	9	0.446	0.905
Residuals	64		

Note. The boldfaced numbers indicate statistically significant results.
**$p < 0.01$, ***$p < 0.001$.

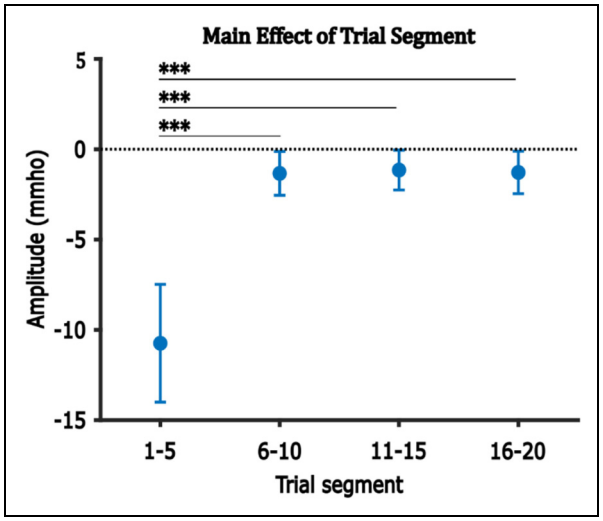

Figure 6. The main effect of Trial segment on the amplitude of the negative component. Error bars represent one SEM. ***$p < 0.001$.

Table 6. Post hoc Analysis for the Main Effect of Trial Segment on the Average Amplitude of the Negative Component.

Group A	Group B	A-B	Adjusted *p*-value
Trial 6–10	**Trial 1–5**	**9.407**	**<0.001*****
Trial 11–15	**Trial 1–5**	**9.591**	**<0.001*****
Trial 16–20	**Trial 1–5**	**9.460**	**<0.001*****
Trial 11–15	Trial 6–10	0.184	1.00
Trial 16–20	Trial 6–10	0.053	1.00
Trial 16–20	Trial 11–15	−0.131	1.00

Note. The boldfaced numbers indicate statistically significant results.
The Analysis was Multiple Comparisons Based on Tukey's Procedure with a Family Error Rate of 0.05.
***$p < 0.001$.

The ICA extracted positive response (C2 in Figure 4) is consistent with an auditory response mechanism. This response, with a typical canonical shape, has been reported in fNIRS studies investigating the hemodynamic response to sound, no matter whether the participants were tested while asleep (Mao et al., 2021; Nakano et al., 2009) or awake (Lloyd-Fox et al., 2019; Taga et al., 2018). A similar response shape has also been reported in an fMRI study of speech perception in awake and asleep infants (Dehaene-Lambertz et al., 2002). These pieces of evidence demonstrate that an infant's brain is activated by sound even during sleep, reflecting the need to stay alert to the environment.

Another finding in this study which supports the notion that the positive response is an auditory response mechanism is the average amplitudes of the positive response in the bilateral temporal ROIs being larger than those in the left prefrontal ROIs. The temporal region is known to be an area involved in processing auditory stimuli (Peelle et al., 2010; Zatorre et al., 2002). Numerous studies reported similar findings where fNIRS responses to speech sounds were larger in the temporal regions than those in the prefrontal regions (Mao et al., 2021; Shader et al., 2021). Our study also found that the average responses in Trials 1–5 and 6–10 were significantly larger than average responses in Trials 16–20, which suggests gradual habituation of responses with repeating stimulation of the same stimuli.

Mechanisms of the Negative Response

The ICA extracted negative response illustrated in Figure 4 is unlike a standard canonical response, because the response peaked around 9.6–11.6 s after the stimulus offset. This characteristic indicates that the mechanism evoking this reduction in neural activity remains active long after the stimulus ends.

The distinctly different speed of habituation and the very different effects of ROI on the positive and negative responses argue strongly that these two responses have

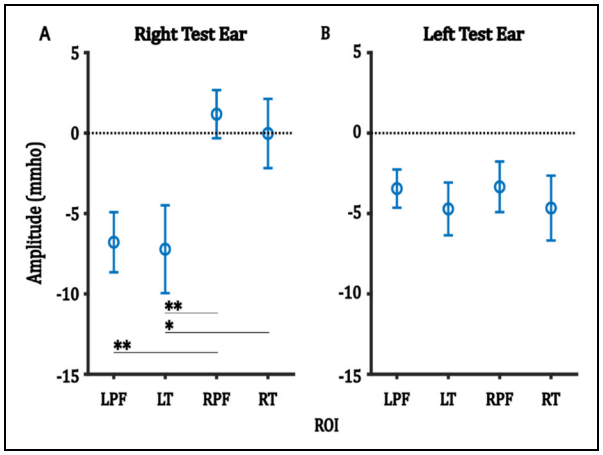

Figure 7. For each test-ear group, the main effect of ROI on the average amplitude of the negative component. Error bars represent one SEM. *$p < 0.05$, **$p < 0.0$.

Table 8. Results of Two-Way ANOVA for the Negative Component of Each Test-Ear Group, with Hemisphere (Hem) and Functional Region (FR) Factors, and Average Amplitude as the Dependent Variable.

	df	F-value	*p*-value
Negative component			
Test ear right			
Hem	**1**	**12.885**	**<0.001*****
FR	1	0.150	0.700
Hem:FR	1	0.033	0.856
Residuals	76		
Test ear left			
Hem	1	0.002	0.960
FR	1	0.627	0.431
Hem:FR	1	0.000	0.987
Residuals	76		

Note. The boldfaced numbers indicate statistically significant results.
***$p < 0.001$.

Table 7. Post Hoc Analysis for the Main Effect of ROI for the Negative Component in Right Test-Ear Group.

Group A	Group B	A-B	Adjusted *p*-value
LT	LPF	−0.434	0.998
RPF	**LPF**	**7.966**	**0.019****
RT	LPF	6.764	0.060
RPF	**LT**	**8.400**	**0.012****
RT	**LT**	**7.198**	**0.040***
RT	RPF	−1.201	0.968

Note. The boldfaced numbers indicate statistically significant results.
The Analysis was Multiple Comparisons Based on Tukey's Procedure with a Family Error Rate of 0.05.
*$p < 0.05$, **$p < 0.01$.

different mechanisms. For the effect of trial, positive responses exhibited a slower reduction over the subsequent trials compared to negative responses, which sharply decreased within the first five trials (Figures 5B and 6). For the effect of ROI, positive HbO responses exhibited a larger response in the temporal regions than the prefrontal regions, but no hemispheric difference. Conversely, negative responses were found to be larger in the left hemispheres than the right hemispheres in the right test-ear group, with no statistically significant response difference observed between ROIs in the left test-ear group.

Previous studies in infants have induced brain arousal response using tactile stimulation, confirmed using EEG, and these responses rapidly habituate with repeated stimulation (Kisilevsky & Muir, 1984; McNamara et al., 1999). Similarly, a study in adults has characterized fNIRS responses associated with spontaneous arousals during sleep, identified through EEG techniques (Nasi et al., 2012). The authors reported a negative trough in the average fNIRS HbO response with a peak latency of 20 s following the onset of arousal, with similar morphology to the negative response observed in our study. Furthermore, Bangash et al. (2008) found a reduction in the cerebral blood flow velocity of the middle cerebral artery, which supplies various regions of the cerebral cortex, including the prefrontal and temporal lobes, when sleeping adults were aroused by acoustic stimulation (confirmed by EEG). This cerebrovascular response peaked around 9 s after the stimulus offset, similar to the peak latency of ICA extracted negative response in our study.

The mechanism of the arousal response during sleep has been described as arising from the reticular activating system (RAS). RAS regulates the sleep-arousal-wake cycle (Moruzzi & Magoun, 1949), and is also responsible for the *fight-or-flight* response (Garcia-Rill, 2015). It reacts to salient or novel stimuli and controls habituation (Garcia-Rill et al., 2007). An important characteristic of the RAS that is consistent with our data is that the arousal response in RAS habituates following the repetition of the same stimulus (Glickman & Feldman, 1961; Sharpless & Jasper, 1956).

In fMRI studies, several potential general mechanisms have been proposed to explain negative hemodynamic responses. One possibility is the active suppression of neural activity in the observed cortical area. An animal study has shown a decrease in the BOLD response of fMRI when inhibitory neurons in the supplementary somatosensory area were stimulated (Moon et al., 2021). The authors also reported a biphasic BOLD response, with a small positive peak followed by a large negative trough in the somatosensory area. In humans, a concurrent fMRI and EEG study reported that regions showing a large negative BOLD response also exhibited large inhibitory cortical potentials (Mullinger et al., 2014). The precise control of RAS on the arousal response remains unclear. However, the diffuse

nature of the negative response in this data (it being seen across all ROIs) could be a result of the nonspecific, widespread projection of neurons from the RAS and other neuromodulatory systems to the cortex (Gent et al., 2018; Starzl & Magoun, 1951). Some studies proposed that active suppression serves as a mechanism to counteract arousal effects (Czisch et al., 2002; Jahnke et al., 2012).

Another suggested mechanism for a negative HbO response is a reduction in activity within the default mode network (DMN). The DMN consists of brain regions that exhibit high activity during the resting state (Raichle et al., 2001). The suppression of DMN activity enhances the efficiency of stimulus processing and is associated with higher performance in cognitive tasks (Anticevic et al., 2012). In the current study, the negative responses were measured from bilateral inferior frontal and temporal regions, which have been identified as part of the DMN in sleeping infants (Fransson et al., 2007). While the negative response in this study was larger in the left hemisphere when the right ear was stimulated, Fransson and colleagues did not observe any specific localization of the DMN activity in sleeping infants.

In fNIRS studies, negative responses have been reported during breath-holding tasks (Emir et al., 2008). Even though we did not measure the breathing rate, we think it is unlikely the infant would hold their breath when the stimuli were presented because they were asleep and not actively paying attention to them. Additionally, there is no significant change expected in the BOLD response if breath holding lasts only for a few seconds (Liu et al., 2002). Furthermore, a study investigating the arousal response in sleeping adults mentioned earlier reported an increase, not a decrease, in breathing rate when participants were aroused (Bangash et al., 2008).

Conclusion

With the evidence presented above, we hypothesize that auditory stimulation simultaneously evokes two independent responses in sleeping infants. The positive HbO response is hypothesized to be an obligatory auditory response and the negative HbO response is hypothesized to be a process related to the brain arousal caused by sensory stimulation during sleep. The interpretation of stimulus-evoked fNIRS response in sleeping infants should not be restricted to the positive peak response post-stimulus onset. A better model to describe the hemodynamic response function of infants would incorporate the negative response and the differing adaptation rate of the two fNIRS responses with repeated stimulus presentations.

Acknowledgements

The authors would like to thank the Victorian Infant Hearing Screening Program (VIHSP) for its support in participant recruitment. The Bionics Institute acknowledges the support it receives from the Victoria State Government through its Operational Infrastructure Support Program.

Data Availability

The data that support the findings of this study are available from the corresponding author, Onn Wah Lee, upon reasonable request.

Declaration of Conflicting Interests

The authors declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: The authors declare that this research received no external financial or nonfinancial support other than those already stated. Onn Wah Lee, Colette M. McKay, Darren Mao, Gautam Balasubramanian and Julia Wunderlich have patent #2022901011 pending with IP Australia. The research is relevant to a spinoff company, NIRGenie Pty Ltd, which was established after this research was completed. None of the authors are employed by NIRGenie Pty Ltd at this present time.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the University of Melbourne (Melbourne Research Scholarship), Garnett Passe and Rodney Williams Memorial Foundation, Australian Government (Medical Research Future Fund), State Government of Victoria (Victorian Medical Research Acceleration Fund), MTPConnect (EarGenie BMTH 03) and National Health and Medical Research Council (GNT1154233).

ORCID iDs

Onn Wah Lee https://orcid.org/0000-0002-7370-9470
Colette M. McKay https://orcid.org/0000-0002-1659-9789

Supplemental Material

Supplemental material for this article is available online.

References

Anticevic, A., Cole, M. W., Murray, J. D., Corlett, P. R., Wang, X. J., & Krystal, J. H. (2012). The role of default network deactivation in cognition and disease. *Trends in Cognitive Sciences*, *16*(12), 584–592. https://doi.org/10.1016/j.tics.2012.10.008

Bangash, M. F., Xie, A., Skatrud, J. B., Reichmuth, K. J., Barczi, S. R., & Morgan, B. J. (2008). Cerebrovascular response to arousal from NREM and REM sleep. *Sleep*, *31*(3), 321–327. https://doi.org/10.1093/sleep/31.3.321

Boersma, P., & Van Heuven, V. (2001). Speak and unSpeak with PRAAT. *Glot International*, *5*(9/10), 341–347.

Bortfeld, H., Wruck, E., & Boas, D. A. (2007). Assessing infants' cortical response to speech using near-infrared spectroscopy. *Neuroimage*, *34*(1), 407–415. https://doi.org/10.1016/j.neuroimage.2006.08.010

Brigadoi, S., & Cooper, R. J. (2015). How short is short? Optimum source-detector distance for short-separation channels in functional near-infrared spectroscopy. *Neurophotonics*, *2*(2), 025005. https://doi.org/10.1117/1.NPh.2.2.025005

Buxton, R. B., Wong, E. C., & Frank, L. R. (1998). Dynamics of blood flow and oxygenation changes during brain activation: The balloon model. *Magnetic Resonance in Medicine*, *39*(6), 855–864. https://doi.org/10.1002/mrm.1910390602

Cabrera, L., & Gervain, J. (2020). Speech perception at birth: The brain encodes fast and slow temporal information. *Science Advances*, *6*(30), eaba7830. https://doi.org/10.1126/sciadv.aba7830

Czisch, M., Wetter, T. C., Kaufmann, C., Pollmacher, T., Holsboer, F., & Auer, D. P. (2002). Altered processing of acoustic stimuli during sleep: Reduced auditory activation and visual deactivation detected by a combined fMRI/EEG study. *Neuroimage*, *16*(1), 251–258. https://doi.org/10.1006/nimg.2002.1071

Dehaene-Lambertz, G., Dehaene, S., & Hertz-Pannier, L. (2002). Functional neuroimaging of speech perception in infants. *Science*, *298*(5600), 2013–2015. https://doi.org/10.1126/science.1077066

Emir, U. E., Ozturk, C., & Akin, A. (2008). Multimodal investigation of fMRI and fNIRS derived breath hold BOLD signals with an expanded balloon model. *Physiological Measurement*, *29*(1), 49–63. https://doi.org/10.1088/0967-3334/29/1/004

Fishburn, F. A., Ludlum, R. S., Vaidya, C. J., & Medvedev, A. V. (2019). Temporal derivative distribution repair (TDDR): A motion correction method for fNIRS. *Neuroimage*, *184*, 171–179. https://doi.org/10.1016/j.neuroimage.2018.09.025

Fransson, P., Skiold, B., Horsch, S., Nordell, A., Blennow, M., Lagercrantz, H., & Aden, U. (2007). Resting-state networks in the infant brain. *Proceedings of the National Academy of Sciences*, *104*(39), 15531–15536. https://doi.org/10.1073/pnas.0704380104

Friston, K. J., Fletcher, P., Josephs, O., Holmes, A., Rugg, M. D., & Turner, R. (1998). Event-related fMRI: Characterizing differential responses. *Neuroimage*, *7*(1), 30–40. https://doi.org/10.1006/nimg.1997.0306

Fu, X. X., & Richards, J. E. (2021). devfOLD: A toolbox for designing age-specific fNIRS channel placement. *Neurophotonics*, *8*(4), 045003. https://doi.org/10.1117/1.NPh.8.4.045003

Garcia-Rill, E. (2015). Chapter 1 - governing principles of brain activity. In E. Garcia-Rill (Ed.), *Waking and the reticular activating system in health and disease* (pp. 1–16). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-801385-4.00001-X

Garcia-Rill, E., Buchanan, R., McKeon, K., Skinner, R. D., & Wallace, T. (2007). Smoking during pregnancy: Postnatal effects on arousal and attentional brain systems. *Neurotoxicology*, *28*(5), 915–923. https://doi.org/10.1016/j.neuro.2007.01.007

Gent, T. C., Bandarabadi, M., Herrera, C. G., & Adamantidis, A. R. (2018). Thalamic dual control of sleep and wakefulness. *Nature Neuroscience*, *21*(7), 974–984. https://doi.org/10.1038/s41593-018-0164-7

Gervain, J., Berent, I., & Werker, J. F. (2012). Binding at birth: The newborn brain detects identity relations and sequential position in speech. *Journal of Cognitive Neuroscience*, *24*(3), 564–574. https://doi.org/10.1162/jocn_a_00157

Glickman, S. E., & Feldman, S. M. (1961). Habituation of the arousal response to direct stimulation of the brainstem. *Electroencephalography and Clinical Neurophysiology*, *13*(5), 703–709. https://doi.org/10.1016/0013-4694(61)90102-X

Issard, C., & Gervain, J. (2018). Variability of the hemodynamic response in infants: Influence of experimental design and stimulus complexity. *Developmental Cognitive Neuroscience*, *33*, 182–193. https://doi.org/10.1016/j.dcn.2018.01.009

Jahnke, K., von Wegner, F., Morzelewski, A., Borisov, S., Maischein, M., Steinmetz, H., & Laufs, H. (2012). To wake or not to wake? The two-sided nature of the human K-complex. *Neuroimage*, *59*(2), 1631–1638. https://doi.org/10.1016/j.neuroimage.2011.09.013

Kisilevsky, B. S., & Muir, D. W. (1984). Neonatal habituation and dishabituation to tactile stimulation during sleep. *Developmental Psychology*, *20*(3), 367–373. https://doi.org/10.1037/0012-1649.20.3.367

Kohno, S., Miyai, I., Seiyama, A., Oda, I., Ishikawa, A., Tsuneishi, S., Amita, T., & Shimizu, K. (2007). Removal of the skin blood flow artifact in functional near-infrared spectroscopic imaging data through independent component analysis. *Journal of Biomedical Optics*, *12*(6), 062111. https://doi.org/10.1117/1.2814249

Liu, H. L., Huang, J., Wu, C. T., & Hsu, Y. Y. (2002). Detectability of blood oxygenation level-dependent signal changes during short breath hold duration. *Magnetic Resonance Imaging*, *20*(9), 643–648. https://doi.org/10.1016/s0730-725x(02)00595-7

Lloyd-Fox, S., Blasi, A., McCann, S., Rozhko, M., Katus, L., Mason, L., Austin, T., Moore, S. E., Elwell, C. E., & Team, B. P. (2019). Habituation and novelty detection fNIRS brain responses in 5- and 8-month-old infants: The Gambia and UK. *Developmental Science*, *22*(5), e12817. https://doi.org/10.1111/desc.12817

Mao, D., Wunderlich, J., Savkovic, B., Jeffreys, E., Nicholls, N., Lee, O. W., Eager, M., & McKay, C. M. (2021). Speech token detection and discrimination in individual infants using functional near-infrared spectroscopy. *Scientific Reports*, *11*(1), 24006. https://doi.org/10.1038/s41598-021-03595-z

McNamara, F., W, H., & Thach, B. T. (1999). Habituation of the infant arousal response. *Sleep*, *22*(3), 320–326. https://doi.org/10.1093/sleep/22.3.320

Moon, H. S., Jiang, H., Vo, T. T., Jung, W. B., Vazquez, A. L., & Kim, S. G. (2021). Contribution of excitatory and inhibitory neuronal activity to BOLD fMRI. *Cerebral Cortex*, *31*(9), 4053–4067. https://doi.org/10.1093/cercor/bhab068

Morais, G. A. Z., Balardin, J. B., & Sato, J. R. (2018). fNIRS Optodes' location decider (fOLD): A toolbox for probe arrangement guided by brain regions-of-interest. *Scientific Reports*, *8*(1), 3341. https://doi.org/10.1038/s41598-018-21716-z

Moruzzi, G., & Magoun, H. W. (1949). Brain stem reticular formation and activation of the EEG. *Electroencephalography and Clinical Neurophysiology*, *1*(4), 455–473. https://doi.org/10.1016/0013-4694(49)90219-9

Mullinger, K. J., Mayhew, S. D., Bagshaw, A. P., Bowtell, R., & Francis, S. T. (2014). Evidence that the negative BOLD response is neuronal in origin: A simultaneous EEG-BOLD-CBF study in humans. *Neuroimage*, *94*, 263–274. https://doi.org/10.1016/j.neuroimage.2014.02.029

Nakano, T., Watanabe, H., Homae, F., & Taga, G. (2009). Prefrontal cortical involvement in young infants' analysis of novelty. *Cerebral Cortex*, *19*(2), 455–463. https://doi.org/10.1093/cercor/bhn096

Nasi, T., Virtanen, J., Toppila, J., Salmi, T., & Ilmoniemi, R. J. (2012). Cyclic alternating pattern is associated with cerebral hemodynamic variation: A near-infrared spectroscopy study of sleep in healthy humans. *PLoS One*, *7*(10), e46899. https://doi.org/10.1371/journal.pone.0046899

Peelle, J. E., Johnsrude, I. S., & Davis, M. H. (2010). Hierarchical processing for speech in human auditory cortex and beyond. *Frontiers in Human Neuroscience*, *4*, 51. https://doi.org/10.3389/fnhum.2010.00051

Pinti, P., Tachtsidis, I., Hamilton, A., Hirsch, J., Aichelburg, C., Gilbert, S., & Burgess, P. W. (2020). The present and future

use of functional near-infrared spectroscopy (fNIRS) for cognitive neuroscience. *Annals of the New York Academy of Sciences*, *1464*(1), 5–29. https://doi.org/10.1111/nyas.13948

Pollonini, L., Olds, C., Abaya, H., Bortfeld, H., Beauchamp, M. S., & Oghalai, J. S. (2014). Auditory cortex activation to natural speech and simulated cochlear implant speech measured with functional near-infrared spectroscopy. *Hearing Research*, *309*, 84–93. https://doi.org/10.1016/j.heares.2013.11.007

Quaresima, V., & Ferrari, M. (2019). Functional near-infrared spectroscopy (fNIRS) for assessing cerebral Cortex function during human behavior in natural/social situations: A concise review. *Organizational Research Methods*, *22*(1), 46–68. https://doi.org/10.1177/1094428116658959

Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., & Shulman, G. L. (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences*, *98*(2), 676–682. https://doi.org/10.1073/pnas.98.2.676

Rolls, E. T., Huang, C. C., Lin, C. P., Feng, J., & Joliot, M. (2020). Automated anatomical labelling Atlas 3. *Neuroimage*, *206*, 116189. https://doi.org/10.1016/j.neuroimage.2019.116189

Santosa, H., Zhai, X., Fishburn, F., & Huppert, T. (2018). The NIRS Brain AnalyzIR Toolbox [article]. *Algorithms*, *11*(5), 73–73. https://doi.org/10.3390/a11050073

Shader, M. J., Luke, R., Gouailhardou, N., & McKay, C. M. (2021). The use of broad vs restricted regions of interest in functional near-infrared spectroscopy for measuring cortical activation to auditory-only and visual-only speech. *Hearing Research*, *406*, 108256. https://doi.org/10.1016/j.heares.2021.108256

Sharpless, S., & Jasper, H. (1956). Habituation of the arousal reaction. *Brain*, *79*(4), 655–680. https://doi.org/10.1093/brain/79.4.655

Starzl, T. E., & Magoun, H. W. (1951). Organization of the diffuse thalamic projection system. *Journal of Neurophysiology*, *14*(2), 133–146. https://doi.org/10.1152/jn.1951.14.2.133

Stone, J. V. (2002). Independent component analysis: An introduction. *Trends in Cognitive Sciences*, *6*(2), 59–64. https://doi.org/10.1016/s1364-6613(00)01813-1

Taga, G., Homae, F., & Watanabe, H. (2007). Effects of source-detector distance of near infrared spectroscopy on the measurement of the cortical hemodynamic response in infants. *Neuroimage*, *38*(3), 452–460. https://doi.org/10.1016/j.neuroimage.2007.07.050

Taga, G., Watanabe, H., & Homae, F. (2018). Developmental changes in cortical sensory processing during wakefulness and sleep. *Neuroimage*, *178*, 519–530. https://doi.org/10.1016/j.neuroimage.2018.05.075

van der Burght, C. L., Numssen, O., Schlaak, B., Goucha, T., & Hartwigsen, G. (2023). Differential contributions of inferior frontal gyrus subregions to sentence processing guided by intonation. *Human Brain Mapping*, *44*(2), 585–598. https://doi.org/10.1002/hbm.26086

Wilcox, T., Bortfeld, H., Woods, R., Wruck, E., & Boas, D. A. (2005). Using near-infrared spectroscopy to assess neural activation during object processing in infants. *Journal of Biomedical Optics*, *10*(1), 011010. https://doi.org/10.1117/1.1852551

Ye, J. C., Tak, S., Jang, K. E., Jung, J., & Jang, J. (2009). NIRS-SPM: Statistical parametric mapping for near-infrared spectroscopy. *Neuroimage*, *44*(2), 428–447. https://doi.org/10.1016/j.neuroimage.2008.08.036

Yi, H. G., Leonard, M. K., & Chang, E. F. (2019). The encoding of speech sounds in the superior temporal gyrus. *Neuron*, *102*(6), 1096–1110. https://doi.org/10.1016/j.neuron.2019.04.023

Zatorre, R. J., Belin, P., & Penhune, V. B. (2002). Structure and function of auditory cortex: Music and speech. *Trends in Cognitive Sciences*, *6*(1), 37–46. https://doi.org/10.1016/s1364-6613(00)01816-7

Zhang, F., Gervain, J., & Roeyers, H. (2022). Developmental changes in the brain response to speech during the first year of life: A near-infrared spectroscopy study of Dutch-learning infants. *Infant Behavior & Development*, *67*, 101724. https://doi.org/10.1016/j.infbeh.2022.101724

Zhang, H., Zhang, Y. J., Lu, C. M., Ma, S. Y., Zang, Y. F., & Zhu, C. Z. (2010). Functional connectivity as revealed by independent component analysis of resting-state fNIRS measurements. *Neuroimage*, *51*(3), 1150–1161. https://doi.org/10.1016/j.neuroimage.2010.02.080

Zhao, Y., Sun, P. P., Tan, F. L., Hou, X., & Zhu, C. Z. (2021). NIRS-ICA: A MATLAB toolbox for independent component analysis applied in fNIRS studies. *Frontiers in Neuroinformatics*, *15*, 683735. https://doi.org/10.3389/fninf.2021.683735