

Active versus passive listening to auditory streaming stimuli: A near-infrared spectroscopy study

Remijn, Gerard Bastiaan

Department of Psychiatry and Neurobiology, Graduate School of Medical Science, Kanazawa University

Kojima, Haruyuki

Department of Psychology, Faculty of Letters, Kanazawa University

<https://hdl.handle.net/2324/1912786>

出版情報 : Journal of Biomedical Optics. 15 (3), pp.037006-1-037006-9, 2010-05-01. Society of Photo-optical Instrumentation Engineers

バージョン :

権利関係 : ©2010 Society of Photo Optical Instrumentation Engineers (SPIE). One print or electronic copy may be made for personal use only. Systematic reproduction and distribution, duplication of any material in this publication for a fee or for commercial purposes, or modification of the contents of the publication are prohibited.



Active versus passive listening to auditory streaming stimuli: a near-infrared spectroscopy study

Gerard B. Remijn*,†

Kanazawa University
Graduate School of Medical Science
Department of Psychiatry and Neurobiology
Takara-machi 13-1
Kanazawa, 920-8641 Japan

Haruyuki Kojima

Kanazawa University
Department of Psychology
Faculty of Letters
Kanazawa, 920-1192 Japan

Abstract. We use near-infrared spectroscopy (NIRS) to assess listeners' cortical responses to a 10-s series of pure tones separated in frequency. Listeners are instructed to either judge the rhythm of these "streaming" stimuli (active-response listening) or to listen to the stimuli passively. Experiment 1 shows that active-response listening causes increases in oxygenated hemoglobin (oxy-Hb) in response to all stimuli, generally over the (pre)motor cortices. The oxy-Hb increases are significantly larger over the right hemisphere than over the left for the final 5 s of the stimulus. Hemodynamic levels do not vary with changes in the frequency separation between the tones and corresponding changes in perceived rhythm ("gallop," "streaming," or "ambiguous"). Experiment 2 shows that hemodynamic levels are strongly influenced by listening mode. For the majority of time windows, active-response listening causes significantly larger oxy-Hb increases than passive listening, significantly over the left hemisphere during the stimulus and over both hemispheres after the stimulus. This difference cannot be attributed to physical motor activity and preparation related to button pressing after stimulus end, because this is required in both listening modes. © 2010 Society of Photo-Optical Instrumentation Engineers. [DOI: 10.1117/1.3431104]

Keywords: near-infrared spectroscopy; hemodynamic activity; auditory streaming; attention; rhythm.

Paper 09430R received Sep. 23, 2009; revised manuscript received Mar. 5, 2010; accepted for publication Mar. 23, 2010; published online May 24, 2010.

1 Introduction

Near-infrared spectroscopy (NIRS) is a noninvasive optical imaging technique that measures changes in the oxygenation of brain tissue, i.e., changes in regional cerebral blood flow (CBF) and cerebral oxygenation. NIRS has been used to monitor changes in cerebral oxygenation in a variety of motor, cognitive, and perceptual tasks,¹⁻⁴ and there is growing consensus that NIRS data are consistent with those obtained with other brain imaging techniques, such as positron emission tomography (PET) and functional magnetic resonance imaging.⁵⁻⁷ Compared to these techniques, a number of factors add to NIRS' promise as a clinical tool: it is relatively low cost, the equipment is mobile, and measurements are fairly tolerable to bodily movements of participants. A number of NIRS studies have indeed been performed in a clinical setting with premature or medically at-risk infants and adults.⁸⁻¹⁰

In the present study we used NIRS to monitor cortical activation in response to auditory stimuli consisting of a series of pure tones alternating in frequency. In experiment 1, listeners judged the rhythm of the series while NIRS measurements were made. In experiment 2, they either judged the rhythm of

the stimuli or listened to them passively during measurements. The stimuli were so-called auditory streaming stimuli,¹¹⁻¹³ which typically consist of repetitive and alternating high (H) and low (L) pure tones of 100 (± 40) ms in duration (Fig. 1). When the frequency difference between the tones is smaller than about a quarter of an octave, a single, integrated "stream" of tones can be heard (e.g., H-L-H-L-H-L, etc.). When the frequency difference between H-L tones increases, however, the series splits into two streams (H-H-H and L-L-L-). This often occurs when the frequency difference exceeds half an octave.¹² Ambiguous percepts are often reported with frequency differences in between one quarter and half an octave between H and L.

To further explore future clinical applications of NIRS, the streaming stimuli were used for the following purposes. In the auditory domain, NIRS has been successful in assessing brain responses to speech, both in developing brains¹⁴⁻¹⁶ and adult brains.¹⁷ Only a few NIRS studies exist, however, on brain activity in response to less complex sounds.¹⁸ As a first aim, we therefore explored whether changes in the stimuli's frequency and perceived rhythm would be accompanied by spatiotemporal changes in cortical hemodynamics in areas that potentially respond to these features. If so, assessing cortical functioning and development of listeners who do not (yet) possess a sufficiently developed language system, such as

*Currently with Kyushu University, Fukuoka, Japan.

†Address all correspondence to: Gerard Remijn, Kyushu University, International Education Center, Minami-ku, Shiobaru 4-9-1, Fukuoka 815-8545, Japan. Tel: 81-92-553-9462; Fax: 81-92-553-9460; E-mail: remijn@design.kyushu-u.ac.jp

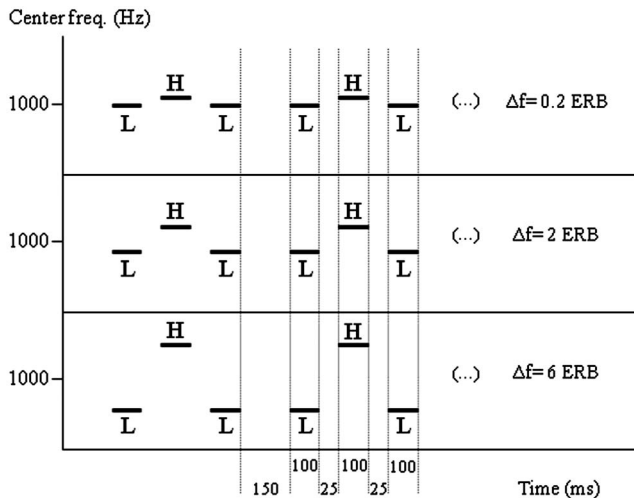


Fig. 1 Schematic impression of the stimuli used in experiment 1. Stimuli consisted of 20 triplets of 500 ms each. Triplets consisted of two tones with a relatively high frequency (H) flanking a tone with a relatively low frequency (H-L-H), or vice versa (L-H-L). Frequencies of H and L were calculated from a center frequency of 600, 800, 1000 (as depicted here), 1200, and 1400 Hz. The frequency separation between H and L was expressed as a measure of the equivalent rectangular bandwidth (ERB, see main text for details) of the center frequency. Frequency separations of 0.2, 2, and 6 ERB were used.

children, might be feasible with the use of nonverbal stimuli as well.

Second, we tested whether the listener's attentional engagement to the stimuli, i.e., active-response listening by requirement of rhythm judgment, would affect cortical hemodynamic activity differently than passive listening to the stimuli. The bulk of NIRS studies on auditory functioning described so far concerns the latter listening mode. With the use of other imaging techniques, it has been shown that active-response listening results in increased cortical activity compared to passive listening, also in areas outside the auditory cortex.¹⁹⁻²¹ With NIRS, enhancing effects of attentional engagement (alertness) have been shown so far in the visual domain, with tasks that required a psychophysical response such as a visual reaction task²² and a visual search task.²³ Whether attention modulates hemodynamic activity levels also in the auditory domain when a psychophysical response is required has not been investigated yet. NIRS' usefulness as a clinical tool, however, would be greatly enhanced if it could visualize whether a listener is actually attending to a stimulus rather than merely receiving it in an otherwise conscious state, especially when the listener is unable to verbalize what he/she perceives or is attending to.

2 Methods

2.1 Participants

Seven students and researchers of cognitive psychology, three females and four males, participated in experiment 1. Three additional participants (three males) joined experiment 2. The participants were 22 to 42 years of age, right-handed, and had normal hearing. Informed consent was obtained from all participants after an explanation about the workings of the NIRS equipment and the procedure of each experiment. The

latter was approved by the Ethics Committee of Kanazawa University Hospital and followed the Declaration of Helsinki.

2.2 Stimuli and Design

Figure 1 schematically shows the three stimulus types used in experiment 1. The stimuli consisted of pure tones of 100 ms in duration. The tones had a rise and fall time of 10 ms with cosine-shaped ramps. The tones were grouped in H-L-H or in L-H-L triplets, with tone H having a higher frequency than tone L. The pause in between the tones within a triplet was 25 ms, whereas the pause in between each triplet was 150 ms. Including the pauses, each triplet thus had a duration of 500 ms. A single stimulus condition lasted 10 s and consisted of 20 triplets presented in succession. Because of the difference between the “within” and “between” triplet pause duration, each stimulus condition could potentially have a galloping rhythm.

As described by Van Noorden,¹² whether a galloping rhythm in auditory streaming stimuli is actually perceived depends on the frequency separation between the H-L tones. The frequency separation between H-L in the present study was calculated by using a measure of the equivalent rectangular bandwidth (ERB) from the stimuli's center frequency. The ERB is an approximation of an auditory filter and is often used to express a frequency separation between two sounds [e.g., Ref. 24]. Here we used $ERB = 24.7(0.00437 \times cf + 1)$, following Glasberg and Moore,²⁵ with cf denoting the center frequency in between H-L.

Three different frequency separations were used in experiment 1: a gallop condition, a two-streams condition, and an ambiguous condition. In the gallop condition, the frequency separation between the H-L tones was 0.2 ERB of the center frequency. In the ambiguous and two-streams conditions, the frequency separation was 2 and 6 ERBs, respectively. The values of 0.2, 2, and 6 ERB correspond roughly to frequency separations of 0.5, 4.5, and 14.5 semitones, and should facilitate the perception of gallop, an ambiguous rhythm, and two streams, respectively, according to Van Noorden.¹²

The three ERB values were calculated from a total of five different center frequencies of 600, 800, 1000, 1200, and 1400 Hz. As an example, in a stimulus with a center frequency of 1000 Hz and a frequency separation of 2 ERB (2×132.7 Hz), the high tones H had a frequency of 1132.7 Hz, and the low tones L had a frequency of 867.3 Hz. The main reason for having five different center frequencies was to make the stimuli not too monotone for the participants and to counteract possible repetition effects on the oxy-Hb and deoxy-Hb changes, which are known to plateau or even become less pronounced when a sound is presented with a repetition rate over 8 Hz.²⁶ The main reason for having center frequencies in the 600- to 1400-Hz range is that tones within this frequency range are generally perceived as equally loud (see for example Equal Loudness Contours in Ref. 27).

2.3 Psychophysical Task and Procedure

In a darkened room, the listener was seated in front of the black screen of a computer (Epson Endeavor NJ1000), which was used to present the sound stimuli and for gathering the psychophysical data. The sound stimuli were presented through two speakers (Dell A215) placed behind the computer

screen. The level of the stimuli was 74.6 dBA on average, as measured with a Rion (Tokyo, Japan) NL-32 sound level meter. The background noise was 42.7 dBA on average and mainly came from the NIRS equipment. Before the start of the experiment the listener was given random examples of the stimuli and asked whether he/she could distinguish a galloping rhythm in some of the stimuli. All listeners responded that they could. For the actual experiment, the listener was instructed to judge the rhythm of the stimuli (active-response listening) by pressing one of two buttons on the computer's keyboard: one for stimuli heard as galloping and the other for stimuli with a different rhythm. Judgments had to be made within 2 s after the end of each stimulus. The listener was instructed to try to limit bodily movements to those necessary for key pressing, and to maintain a stable head position with the use of a head-and-chin rest.

In experiment 1, 12 blocks of stimuli were presented, with each block regarded as a single session. Each block consisted of 15 sound stimuli, with the three stimulus types (gallop, ambiguous, and two streams) represented in five stimuli, one for each center frequency (600, 800, 1000, 1200, and 1400 Hz). For practical reasons, the order of presentation within each block was pseudorandomized in that a certain stimulus type was never presented twice in succession. The stimuli's center frequency, though, was completely randomized. In total, 60 judgments were obtained for each of the three stimulus types. The many repetitions were given to obtain a relatively stable average of oxy- and deoxy-Hb changes in each individual brain.

In experiment 2, only the ambiguous stimuli were used. The listeners performed four blocks of 15 sound stimuli. In two blocks, similar to experiment 1, the listener was asked to judge the rhythm of each stimulus by button pressing (active-response listening condition). In the other two blocks, the listener was asked to passively listen to the sound stimuli and randomly press one of the two buttons after the end of each stimulus. During both active-response listening and passive listening, the listener was asked to keep his/her eyes open.

2.4 Near-Infrared Spectroscopy Measurements

NIRS measurements (for details, see Ref. 28) were made with a continuous wave system (ETG-4000, Hitachi Medical Company, Japan) with two 3×3 optode probe sets. Each set consisted of five light emitters and four photodetectors placed in a silicone rubber frame, comprising 12 channels in total. Oxy-, deoxy-, and total (oxy-deoxy) Hb values were obtained from channels 1 through 12 covering the left hemisphere and channels 13 through 24 covering the right hemisphere. The light emitted by the NIRS system had wavelengths of 695 and 830 nm (each ± 20 nm), and the frequency was modulated for wavelength and the number of channels. The unabsorbed light that left the brain was received by the photodetectors and amplified for each particular frequency. Because the optical path length cannot be measured by a continuous wave system, as used here, from here on the (de)oxy-Hb changes as measured are indicated by the scale unit $\text{mMol} \times \text{mm}$ (molar concentration times the unknown path length). The measurable depth with the interoptode distance was 15 to 25 mm beneath the scalp, following Hoshi et al.²⁹

Measurements were made over the right and left fronto-temporal areas of the listener's brain. Epochs of 25 s per stimulus were recorded, including 15 s for the rest period, with a sampling rate of 10 Hz. Each set of 12 channels was symmetrically placed on one side of the brain, with channels 1 through 12 covering the left hemisphere and channels 13 through 24 covering the right hemisphere. Following the international 10-20 system for EEG (Ref. 30; see Ref. 31, for correspondence with NIRS measurements), the penultimate posterior optode row was placed on the imaginary line connecting the electrode positions C3-T3 for the left hemisphere and C4-T4 for the right hemisphere (shown later).

Both sets subtended a fixed distance of 12 cm centered around electrode position Cz, with channels 2, 4, 5, and 7 approximately surrounding motor area C3 (left hemisphere), and channels 13, 15, 16, and 18 covering motor area C4 (right hemisphere). We opted to cover the motor areas, because studies have shown that the classical auditory system of the temporal lobe (i.e., areas T3 and T4) does not play a major role in rhythm perception.³² Rather, the areas involved in rhythm perception are those involved in stimulus prediction (the lateral and mesial premotor areas), as well as areas involved in motor activity and preparation, among others.³³⁻³⁶ Activity related to button pressing, required after stimulus end, was expected in (pre)motor areas as well.

2.5 Data Analysis

The raw oxy- and deoxy-Hb data were digitally low-pass filtered at 0.1 Hz to remove measurement artifacts, i.e., abrupt value changes that could have occurred because of bodily movements of the listener. After baseline correction, for each of the 24 channels, the oxy- and deoxy-Hb data were averaged over four time windows of 2.5 s (during stimulus) and over four time windows of 2.5 s after stimulus end (post-stimulus). A similar division of the time scale has been performed in other NIRS studies (e.g., Ref. 18).

As a first analysis, we compared the averaged data for each time window, both during and poststimulus, with the averaged (de)oxy-Hb values of a 2.5-s baseline before the start of each stimulus. For each participant, stimulus condition, center frequency, and channel, we subtracted the mean (de)oxy-Hb values for the baseline from those obtained during each of the four during-stimulus and the four poststimulus windows. We used the results to check for a significant change in (de)oxy-Hb by means of multiple t-tests against zero. The Bonferroni correction was applied to account for the number of comparisons per condition and window (24 being the number of channels). This lowered the alpha-level to $p=0.0021$.

Subsequently, the stimulus minus baseline (de)oxy-Hb values were subjected to three-way analyses of variance (ANOVAs) with repeated measures (participants and center frequency). The first two ANOVAs were on the oxy-Hb data and on the deoxy-Hb data obtained during stimulus presentation, with stimulus condition (3), time window (4), and hemisphere (2) as main factors. Values for the latter were obtained by averaging the (de)oxy-Hb values relative to the baseline for the 12 channels on the left hemisphere (LH) and the 12 channels on the right hemisphere (RH). The poststimulus (de)oxy-Hb data were subjected to the same three-way

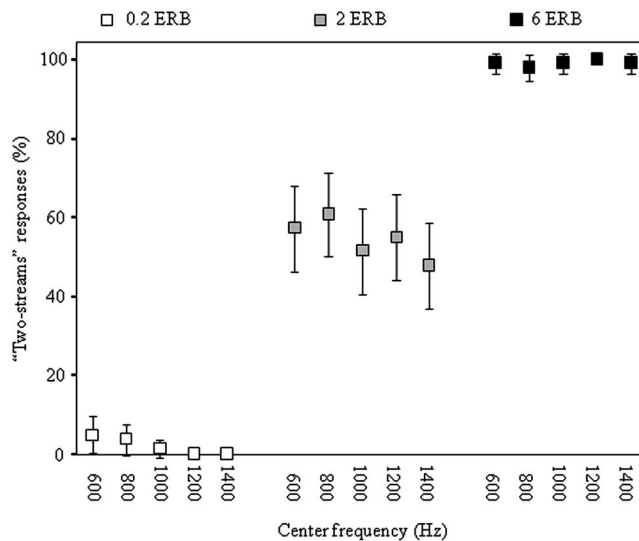


Fig. 2 Rhythm judgments obtained in experiment 1. The error bars show 95% confidence intervals.

ANOVAs. Posthoc tests were performed with Tukey honestly significant difference (HSD) tests ($p < 0.05$).

Besides the global time course of (de)oxy-Hb throughout stimulus presentation and judgment by means of time windows, the temporal aspects of the oxy-Hb data were further analyzed by specifying the average point in time where oxy-Hb reached its peak for each stimulus condition over both hemispheres. Both a during-stimulus and a poststimulus two-way ANOVA was performed for the data of experiments 1 and 2.

3 Results

3.1 Psychophysical data

Figure 2 shows the results of the two alternative forced choice (2AFC) task of experiment 1. As can be derived from the 95%

confidence intervals, the 0.2-ERB conditions resulted in significantly more galloping percepts than the 2-ERB and 6-ERB conditions. The 6-ERB conditions caused significantly more two-streams percepts than the ambiguous 2-ERB conditions. A two-way ANOVA with repeated measures (7 participants \times 12 sessions) and posthoc Tukey tests ($p < 0.05$) confirmed the existence of significant differences between all the conditions [$F(2,664)=98.82$, $p < 0.01$]. In experiment 2, only 2-ERB stimuli were used. The average proportion of two-streams percepts for these stimuli was $0.62 (\pm 0.24)$, showing that the listeners heard the stimuli as having an ambiguous rhythm in this experiment as well.

3.2 (De)oxygenated Hemoglobin—Temporal Characteristics

Figure 3 shows the average time courses of the oxy-Hb changes obtained in experiment 1 in response to the three stimulus types. The figure shows that oxy-Hb slowly increased after the start of a stimulus, reached a first peak, on average at $5.14 \text{ s} (\pm 0.25 \text{ s})$ over LH and at $5.69 \text{ s} (\pm 0.32 \text{ s})$ over RH after stimulus onset, and gradually decreased as the stimulus reached its end. Two-way ANOVA showed that the difference in the temporal peak of oxy-Hb between LH and RH was significant [$F(1,68)=12.67$, $p < 0.01$]. The oxy-Hb peak during stimulus presentation, however, did not differ in time of occurrence for the 0.2-, 2-, and 6-ERB conditions ($p=0.38$).

With regard to the overall time course of the oxy-Hb changes, three-way ANOVA showed a significant main effect of time window for oxy-Hb during stimulus [$F(3,204)=8.93$, $p < 0.01$]. Posthoc tests ($p < 0.05$) showed that oxy-Hb was significantly higher during windows 2 and 3 (2.5 to 7.4 s after stimulus onset) as compared to window 1 (0 to 2.4 s). Oxy-Hb during window 3 was also significantly higher than during window 4 (7.5 to 9.9 s).

A second increase in oxy-Hb occurred after stimulus end, when the listeners made the rhythm judgment. This oxy-Hb increase peaked on average at $6.09 \text{ s} \pm 0.13 \text{ s}$ over LH and at

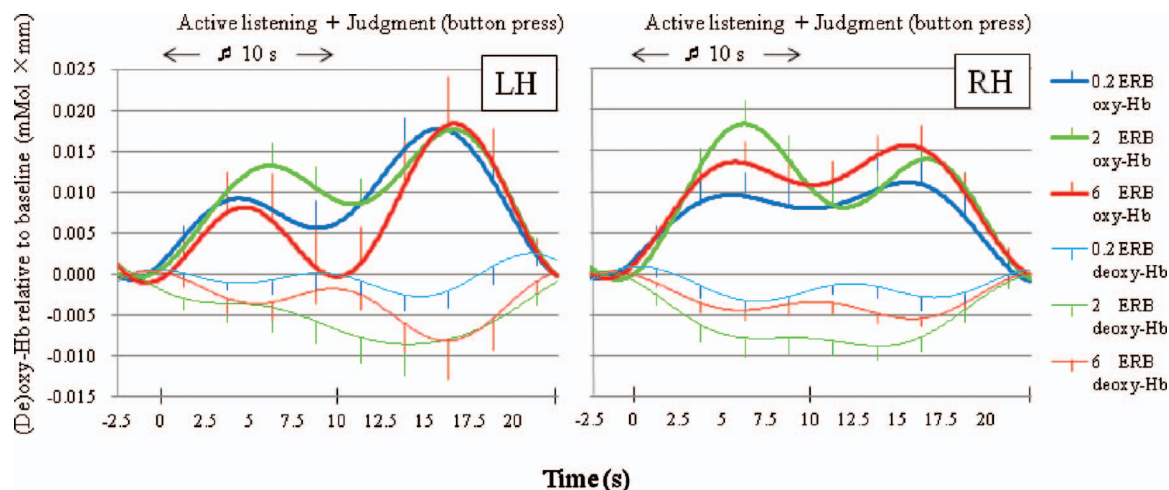


Fig. 3 Time course of the (de)oxy-Hb group means in response to the three stimulus conditions used in experiment 1. Values are relative to (de)oxy-Hb group means obtained during the period of 2.5 s before the start of the stimulus. The left panel shows the oxy-Hb group means averaged over the channels covering the left hemisphere (LH); the right panel shows the oxy-Hb group means averaged over the channels covering the right hemisphere (RH). Error bars indicate standard error of the mean.

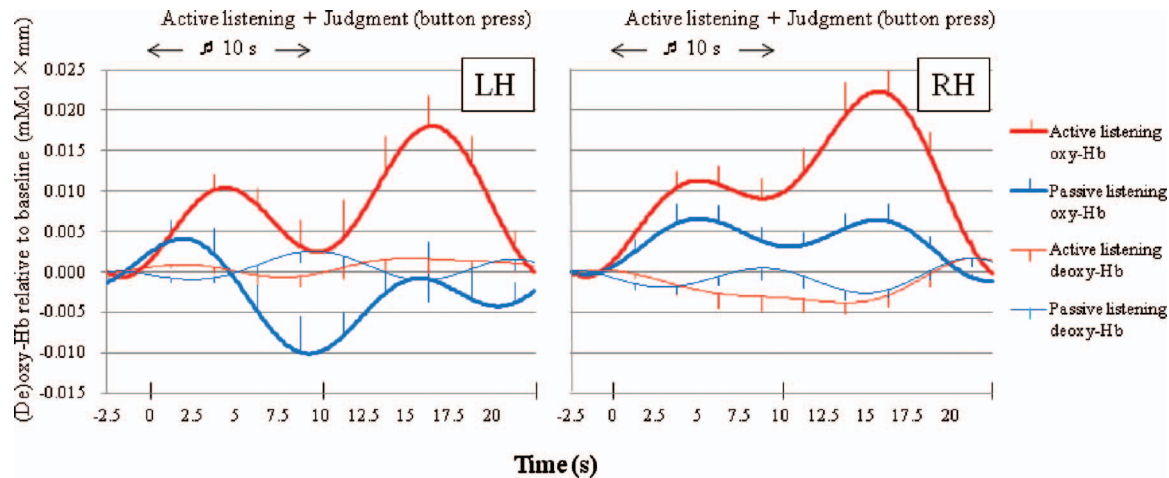


Fig. 4 Time course of the (de)oxy-Hb group means during active-response listening (red) and passive listening (blue) obtained in experiment 2. Values are relative to (de)oxy-Hb group means obtained during the period of 2.5 s before the start of the stimulus. The left panel shows the oxy-Hb group means averaged over the channels covering the left hemisphere (LH); the right panel shows the oxy-Hb group means averaged over the channels covering the right hemisphere (RH). Error bars indicate standard error of the mean.

5.21 s \pm 0.26 s over RH after stimulus end. Again, this difference was significant, with oxy-Hb peaking faster over RH than LH this time [$F(1,68)=13.17$, $p<0.01$]. Here too, stimulus condition had no influence on the speed with which oxy-Hb peaked ($p=0.78$). Three-way ANOVA showed a significant main effect of window [$F(3,204)=7.38$, $p<0.01$]. Posthoc tests ($p<0.05$) showed that oxy-Hb was significantly higher in window 7 (5 to 7.4 s after stimulus offset) as compared to windows 5 and 6 (0 to 4.9 s after stimulus offset) and window 8 (7.5 to 9.9 s after stimulus offset). Oxy-Hb in window 6 (2.5 to 4.9 s after stimulus offset) was higher than that in window 5 (0 to 2.5 s after stimulus offset).

In experiment 1, the effect of window on deoxy-Hb was not significant during stimulus ($p=0.08$), but was significant for the poststimulus period [$F(3,204)=4.62$, $p<0.01$]. Posthoc tests ($p<0.05$) showed that average deoxy-Hb was significantly higher during window 8 (7.5 to 9.9 s after stimulus offset) compared to window 7 (5 to 7.4 s after stimulus offset).

Figure 4 shows the average time courses of the oxy-Hb changes obtained in experiment 2 in the active-response and passive listening conditions. In experiment 2, oxy-Hb during stimulus peaked on average at 4.95 (\pm 0.27 s) over LH and at 4.93 (\pm 0.32 s) over RH. This difference was not significant ($p=0.89$). With regard to the overall time course of oxy-Hb, three-way ANOVA revealed a significant main effect of window [$F(3,147)=3.34$, $p<0.05$], with posthoc tests showing that oxy-Hb was higher during window 2 (2.5 to 4.9 s after stimulus onset) than during window 1 (0 to 2.4 s). After stimulus end, oxy-Hb peaked again on average at 5.97 (\pm 0.45 s) over LH and at 5.47 (\pm 0.59 s) over RH after stimulus end, which was significant [$F(1,44)=4.12$, $p<0.05$]. The effect of window was also significant for the poststimulus three-way ANOVA of experiment 2 [$F(3,147)=7.12$, $p<0.01$]. Here window 7 (5 to 7.4 s after stimulus offset) showed higher oxy-Hb than windows 5 and 8 (0 to 2.4 and 7.5 to 9.9 s after stimulus offset). Window 6 (2.5 to 4.9 s after stimulus offset) also showed higher oxy-Hb than win-

dow 5 (0 to 2.4 s after stimulus offset). For deoxy-Hb, no significant effect of window appeared in experiment 2.

3.3 (De)oxygenated Hemoglobin—Spatial Characteristics

Figure 5 shows the results of the t-tests ($p<0.0021$, with Bonferroni correction) against zero for both experiments 1 and 2. The upper panel shows that significant oxy-Hb changes in experiment 1 occurred most often over channel 4, followed by channel 7 and 10 on LH. Channels 4 and 7, along with channels 2 and 5, approximately surrounded motor area C3; channel 10 is located inferior to channel 5. Most significant oxy-Hb changes on RH were found over channels 13, 16, 18, and 20. The first three channels, along with channel 15 approximately surrounded motor area C4; channel 20 is located inferior to channel 15. Significant deoxy-Hb changes occurred less frequently, but could be found generally over the same areas as those over which significant oxy-Hb changes appeared.

Figure 5 suggests that generally more channels signaled a significant hemodynamic change over RH as compared to LH. In experiment 1, a significant effect of hemisphere was found in the during-stimulus ANOVA for oxy-Hb [$F(1,204)=10.71$, $p<0.01$] and a significant hemisphere by window interaction [$F(3,204)=7.48$, $p<0.01$]. Posthoc tests ($p<0.05$) showed that oxy-Hb was larger in RH than LH for time windows 3 and 4 (the last 5 s of the stimulus). No effect of hemisphere on oxy-Hb was found in experiment 2. Hemispheric differences with regard to deoxy-Hb were also not found in experiments 1 and 2.

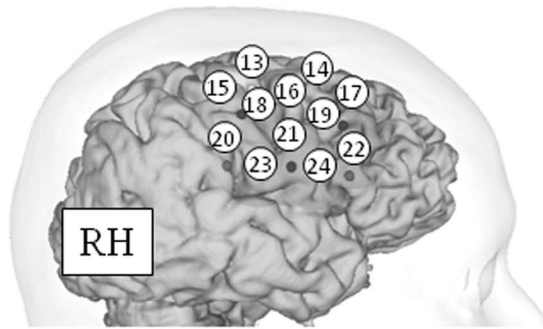
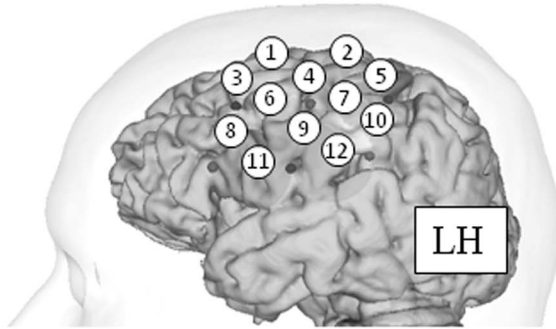
3.4 (De)oxygenated Hemoglobin—Perceived Rhythm

Neither the during-stimulus ($p=0.71$) nor the poststimulus ANOVA ($p=0.98$) showed a significant effect of frequency separation on oxy-Hb in experiment 1. Deoxy-Hb also did not vary with stimulus condition ($p=0.20$ and $p=0.09$ for the during-stimulus and poststimulus ANOVA, respectively). The significant influence of frequency separation on perceived

Experiment 1: channels with a significant oxy-Hb change

Channels with a significant deoxy-Hb change

Stimulus (Δf)		0.2 ERB		2 ERB		6 ERB		0.2 ERB		2 ERB		6 ERB	
#	Window (s)	LH	RH	LH	RH	LH	RH	LH	RH	LH	RH	LH	RH
1	0.0 - 2.4	4	13, 16	4, 10	13, 20	3, 4, 11	13	-	-	-	-	-	-
2	2.5 - 4.9	4	13, 16	4, 7, 10	13, 16, 18, 20	3, 4, 7	11, 13, 16, 17, 18, 20	-	-	-	16, 19, 20, 21, 23	-	-
3	5.0 - 7.4	4	13, 16, 18	4, 7, 10	13, 16, 18, 19, 20, 23	-	11, 16, 20	-	-	6	16, 19, 21, 23	-	-
4	7.5 - 9.9	4	14, 16, 18	4	13, 16, 18, 19, 20, 23	-	16, 18, 19	-	-	6	16, 19, 21, 22, 23	16	-
5	10.0 - 12.4	4	13, 16, 18	4	16, 18, 20	-	16, 18, 19, 20	-	-	4, 6	16, 18, 19, 20, 21, 22, 23	16	-
6	12.5 - 14.9	1, 4	13, 16, 18	4	16, 17, 20	4	13, 16, 17, 18, 20	-	13, 16	4	16, 18, 19, 20, 23	16	-
7	15.0 - 17.4	3, 4, 5	13, 16, 18, 20	1, 4, 7, 10	13, 16, 17, 18, 20	4	13, 16, 18, 20	-	13, 16	4, 7	16, 19, 20, 23	16	-
8	17.5 - 19.9	4, 6, 7	13, 16	4, 6, 7, 10	13, 16, 17, 18, 20	4	13, 16, 17, 20	8	-	7	16, 19, 20	-	-



Experiment 2: channels with a significant oxy-Hb change

Channels with a significant deoxy-Hb change

$(\Delta f) = 2$ ERB		Active - response		Passive - button press		Active - response		Passive - button press	
#	Window (s)	LH	RH	LH	RH	LH	RH	LH	RH
1	0.0 - 2.4	-	16, 20	5	13	-	-	-	23
2	2.5 - 4.9	4, 5, 7, 10	13, 15, 16, 20	5	13	-	-	-	-
3	5.0 - 7.4	5, 7	15, 16, 20	-	-	-	-	-	-
4	7.5 - 9.9	5, 7	15, 16, 20	-	-	-	16	8	-
5	10.0 - 12.4	5, 7	13, 14, 15, 16, 20	-	-	-	16	8	16
6	12.5 - 14.9	4, 5, 6, 7	13, 14, 15, 16, 17, 18, 19, 20	-	16	8	-	5	-
7	15.0 - 17.4	3, 4, 5, 6, 7, 10	13, 15, 16, 17, 18, 19, 20	5	-	8	-	5	-
8	17.5 - 19.9	4, 5, 7	13, 15, 16, 17, 18, 19, 20	5	-	-	-	5, 8	-

Fig. 5 Results of t-tests against zero (with Bonferroni correction, $p < 0.0021$) for (de)oxy-Hb means for the four time windows during stimulus (windows 1 through 4) and the four windows after stimulus (windows 5 through 8). The upper panel shows the results of experiment 1. The channel numbers following optode placement for LH (left hemisphere) and RH (right hemisphere) are shown in the middle. The lower panel shows the results of experiment 2.

rhythm, as found in the behavioral data of experiment 1 (Fig. 2), thus was not accompanied by significant changes in (de)oxy-Hb values.

3.5 (De)oxygenated Hemoglobin—Active Versus Passive Listening

Figure 6 show the oxy-Hb levels as obtained in experiment 2 during active-response listening and passive listening. The figure was made through MRI-fusion software (Shimadzu, Kyoto, Japan), with the oxy-Hb data plotted over the brain of an average-sized male head. NIRS optode positions (both emitters and receivers) were as during measurements and, along with nasion, Cz and ear references, obtained with a 3-D digitizer (Fastrak, Polhemus, Incorporated, Colchester, Vermont) after the experiment. The during-stimulus ANOVA showed a significant main effect of listening mode [$F(1,147) = 5.19$, $p < 0.05$], as well as a significant three-way interaction of listening mode by hemisphere by window [$F(3,147) = 3.49$, $p < 0.05$]. Posthoc tests ($p < 0.05$) revealed that active listening caused significantly larger oxy-Hb changes over the left hemisphere during the final 7.5 s of the stimulus. The poststimulus ANOVA also showed a significant

effect of listening mode [$F(1,147) = 10.15$, $p < 0.01$]. Here too, the three-way interaction was significant [$F(3,147) = 4.99$, $p < 0.01$]. A significant difference in oxy-Hb for active and passive listening was not found over the right hemisphere during the final 2.5 s of the stimulus—other windows showed significantly higher oxy-Hb during active listening. An effect of listening mode was not found for deoxy-Hb values, both during ($p = 0.18$) and after stimulus presentation ($p = 0.46$) of experiment 2.

4 Discussion and Conclusions

In the present study, we explored whether cortical hemodynamics related to the perception and judgment of sound stimuli consisting of simple, rhythmic tones could be visualized with NIRS. With our (potentially) rhythmic stimuli we mainly targeted the (pre)motor areas of the brain, because these are known to be more involved in rhythm perception than the primary auditory areas³² and are likely to mediate the actual act of button pressing, through which listeners were asked to judge the rhythm of the stimuli. The present study showed that hemodynamic response patterns were typical to those in response to sensory stimulation in general with re-

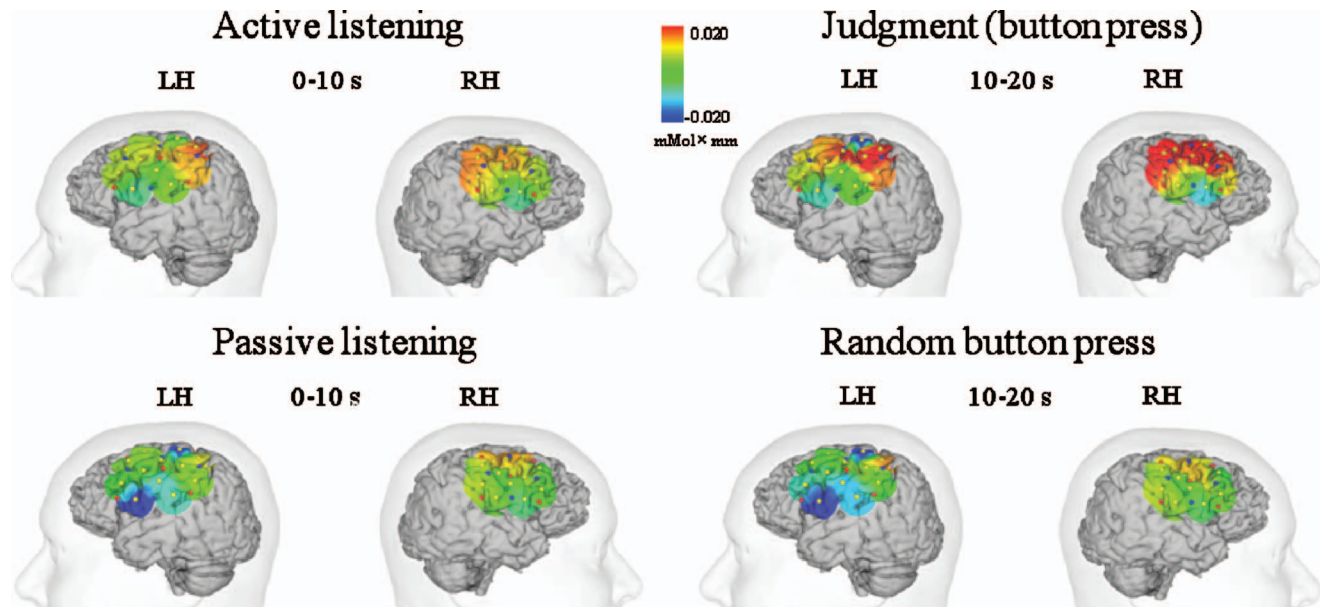


Fig. 6 Oxy-Hb group means during active-response listening (upper panels) and passive listening (lower panels) obtained in experiment 2, plotted over an average-sized male brain (LH, is left hemisphere, and RH is right hemisphere).

gand to their time course.²⁹ Oxy-Hb levels increased gradually, peaked during stimulus, decreased, and picked up again after the stimulus and decreased again. Deoxy-Hb levels in general dipped slightly below zero and stayed relatively flat after stimulus onset.

Oxy-Hb levels during stimulus presentation were partly lateralized, with increased activity mainly over (pre)motor areas of RH during the last 5 s of the 10-s stimulus in experiment 1. This is in line with earlier research showing that these areas respond to rhythmic, unstressed sounds,³⁷ which the stimuli basically are. It has to be noted, though, that the LH-RH difference did not appear in experiment 2 in the active-listening condition as used in experiment 1. Furthermore, the temporal peak of the oxy-Hb levels occurred faster over LH than RH in experiment 1. We further expected lateralized hemodynamic changes in response to the button pressing after stimulus end. Contralateral hemispheric dominance has been reported for a number of simple motor tasks and handedness.^{38,39} Because listeners made their judgment by pressing a button with their right hand, one thus can expect relatively larger hemodynamic changes over the (pre)motor areas of LH. Rather surprisingly, though, these were not found. Oxy-Hb in the active-response listening mode of experiment 2 in particular was unexpectedly high over RH after stimulus end. We have no plausible explanation for this.

We also did not observe an effect of frequency separation on hemodynamics, neither during nor after stimulus presentation. Although the behavioral data showed that changes in the sounds' frequency separation had significant effects on perceived rhythm, different rhythm percepts were not accompanied by significant differences in (de)oxy-Hb. It is possible that such differences become apparent when primary auditory areas (T3 and T4) are targeted as well. Studies have shown that the actual process of stream segregation is reflected in primary cortical activity,^{40,41} although nonprimary auditory areas also seem involved.⁴² More specifically, the perception of

two segregated streams is accompanied by activity that is more sustained and larger in magnitude than that reflecting the perception of a single stream.⁴³ Future NIRS research is necessary to remedy the present study's limitations with regard to cortical areas covered. The oxy-Hb levels also did not differ between the three stimulus conditions with regard to their temporal peak after stimulus onset. Perceptually, the two-streams percept takes several seconds to build up, whereas the galloping rhythm is more quickly established.¹³ Clearly though, hemodynamic response patterns do not reflect this buildup, if theoretically possible.

Experiment 2 showed that hemodynamic activity significantly increased during active-response listening as compared to passive listening. The main effect of listening mode occurred both during and after stimulus presentation. The enhancing effect of active-response listening was significant over LH, except for the first 2.5 s of the stimulus. After the stimulus, the enhancing effect occurred over both hemispheres. The increase was not merely due to motor activity connected to button pressing, since such a response was also required in the passive listening condition. Other studies have also shown that sustained and/or selective attention necessary during active listening has a widespread influence on neural activity. Auditory attention is known to cause increased activity not only over the auditory cortex,^{44,45} but also over a broad range of other cortical areas, including frontal, prefrontal, parietal and supplementary motor areas.^{46,47}

In view of the enhancing effect of active-response listening, future NIRS research might address possible effects of attention on auditory streaming with more specific listening instructions. It is assumed that the frequency separation between the tones in auditory streaming stimuli may not be the only catalyst of different rhythm percepts. A number of studies have suggested that focused attention to either the low or high tones can facilitate the buildup of the two-streams percept and the integration of successive tones within the

streams.^{48,49} Within the active listening mode, one could ask listeners to adopt such analytical listening. Experiment 1 indicated that the perception of different rhythms does not cause significant differences between hemodynamic response patterns, nor in oxy-Hb peak latency. However, the effort of focused attention, i.e., analytical listening to either the high or low tones, might bring out different hemodynamic response patterns. Neuromagnetic data obtained during analytical listening have suggested that focused attention enhanced cortical responses in addition to physical manipulations of frequency separation.⁴² Furthermore, in view of the fact that listeners can often exert control over percepts in the ambiguous stimulus,¹² asking listeners to engage in active switching between rhythm percepts might cause different hemodynamic patterns as well.

Acknowledgments

This study was supported by the COE program Innovative Brain Science for Development, Learning and Memory of Kanazawa University and grants from the Ishikawa High-Tech Sensing Cluster (Knowledge Cluster Initiative from the Japanese Ministry of Education, Culture, Sports, Science and Technology). We thank Koichiro Miyaji and Shuichiro Taya for their technical assistance, and two anonymous reviewers for their help with the manuscript.

References

1. F. F. Jöbsis, "Noninvasive, infrared monitoring of cerebral and myocardial oxygen sufficiency and circulatory parameters," *Science* **198**, 1264–1267 (1977).
2. T. Kato, A. Kamei, S. Takashima, and T. Ozaki, "Human visual cortical function during photic stimulation monitoring by means of near-infrared spectroscopy," *J. Cereb. Blood Flow Metab.* **13**, 516–520 (1993).
3. C. Hirth, H. Obrig, K. Villringer, A. Thiel, J. Bernarding, W. Muhlnickel, H. Flor, U. Dimagi, and A. Villringer, "Non-invasive functional mapping of the human motor cortex using near-infrared spectroscopy," *NeuroReport* **7**, 1977–1981 (1996).
4. Y. Hoshi and M. Tamura, "Near-infrared optical detection of sequential brain activation in the prefrontal cortex during mental tasks," *Neuroimage* **5**, 292–297 (1997).
5. A. Kleinschmidt, H. Obrig, M. Requardt, K. D. Merboldt, U. Dirnagl, A. Villringer, and J. Frahm "Simultaneous recording of cerebral blood oxygenation changes during human brain activation by magnetic resonance imaging and near-infrared spectroscopy," *J. Cereb. Blood Flow Metab.* **16**, 817–826 (1996).
6. K. Villringer, S. Minoshima, C. Hock, H. Obrig, S. Ziegler, U. Dirnagl, M. Schwaiger, and A. Villringer, "Assessment of local brain activation. A simultaneous PET and near-infrared spectroscopy study," *Adv. Exp. Med. Biol.* **413**, 149–153 (1997).
7. G. Strangman, J. P. Culver, J. H. Thompson, and D. A. Boas, "A quantitative comparison of simultaneous BOLD fMRI and NIRS recordings during functional brain activation," *Neuroimage* **17**, 719–731 (2002).
8. C. Hock, K. Villringer, H. Heekeren, M. Hofmann, R. Wenzel, A. Villringer, and F. Müller-Spahn, "A role for near infrared spectroscopy in psychiatry?," *Adv. Exp. Med. Biol.* **413**, 105–123 (1997).
9. J. H. Meek, M. Noone, C. E. Elwell, and J. S. Wyatt, "Visually evoked haemodynamic responses in infants with cerebral pathology," *Pediatr. Res.* **45**, 909 (1999).
10. A. Gallagher, D. Bastien, I. Pelletier, P. Vannasing, A. D. Legatt, S. L. Moshé, R. Jehle, L. Carmant, F. Lepore, R. Béland, and M. Lassonde, "A noninvasive, presurgical expressive and receptive language investigation in a 9-year-old epileptic boy using near-infrared spectroscopy," *Epilepsy Behav.* **12**, 340–346 (2008).
11. G. A. Miller and G. A. Heise, "The trill threshold," *J. Acoust. Soc. Am.* **22**, 720–725 (1950).
12. L. P. A. S. Van Noorden, "Temporal coherence in the perception of tone sequences," Unpublished Doctoral Dissertation, Technische Hogeschool Eindhoven, Eindhoven, the Netherlands (1975).
13. A. S. Bregman, "Auditory Scene Analysis: The Perceptual Organization of Sound," MIT Press, Cambridge, MA (1990).
14. K. Sakatani, S. Chen, W. Lichty, H. Zuo, and Y. Wang, "Cerebral blood oxygenation changes induced by auditory stimulation in newborn infants measured by near infrared spectroscopy," *Early Hum. Dev.* **55**, 229–236 (1999).
15. H. Bortfeld, E. Wruck, and D. A. Boas, "Assessing infants' cortical response to speech using near-infrared spectroscopy," *Neuroimage* **34**, 407–415 (2007).
16. Y. Saito, T. Kondo, S. Aoyama, R. Fukumoto, N. Konishi, K. Nakamura, M. Kobayashi, and T. Toshima, "The function of the frontal lobe in neonates for response to a prosodic voice," *Early Hum. Dev.* **83**, 225–230 (2007).
17. M. J. Herrmann, A. C. Ehlis, and A. J. Fallgatter, "Frontal activation during a verbal-fluency task as measured by near-infrared spectroscopy," *Brain Res. Bull.* **61**, 51–56 (2003).
18. D. Abla and K. Okanoya, "Statistical segmentation of tone sequences activates the left inferior frontal cortex: a near-infrared spectroscopy study," *Neuropsychologia* **46**, 2787–2795 (2008).
19. J. J. Vannest, P. R. Karunanayaka, M. Altaye, V. J. Schmidhorst, E. M. Plante, K. J. Eaton, J. M. Rasmussen, and S. K. Holland, "Comparison of fMRI data from passive listening and active-response story processing tasks in children," *J. Magn. Reson. Imaging* **29**, 971–976 (2009).
20. D. A. Hall, M. P. Haggard, M. A. Akeroyd, A. Q. Summerfield, A. R. Palmer, M. R. Elliott, and R. W. Bowtell, "Modulation and task effects in auditory processing measured using fMRI," *Hum. Brain Mapp.* **10**, 107–119 (2000).
21. L. Gootjes, A. Bouman, J. W. Van Strien, P. Scheltens, and C. J. Stam, "Attention modulates hemispheric differences in functional connectivity: Evidence from MEG recordings," *Neuroimage* **30**, 245–253 (2006).
22. M. J. Herrmann, E. Woldich, T. Schreppel, P. Pauli, and A. J. Fallgatter, "Brain activation for alertness measured with functional near infrared spectroscopy (fNIRS)," *Psychophysiology* **45**, 480–486 (2008).
23. H. Kojima and T. Suzuki, "Hemodynamic change in occipital lobe during visual search: visual attention allocation measured with NIRS," *Neuropsychologia* **48**, 349–352 (2010).
24. G. B. Remijn and Y. Nakajima, "The perceptual integration of auditory stimulus edges: an illusory short tone in stimulus patterns consisting of two partly overlapping glides," *J. Exp. Psychol. Human* **31**, 183–192 (2005).
25. B. R. Glasberg and B. C. J. Moore, "Derivation of auditory filter shapes from notched-noise data," *Hear. Res.* **47**, 103–138 (1990).
26. A. P. Weiss, M. Duff, J. L. Roffman, S. L. Rauch, and G. E. Strangman, "Auditory stimulus repetition effects on cortical hemoglobin oxygenation: a near-infrared spectroscopy investigation," *NeuroReport* **19**, 161–165 (2008).
27. B. C. J. Moore, "Perception of loudness," *An Introduction to the Psychology of Hearing*, 5th ed, Chap. 4, pp. 128–130, Academic Press, London (2003).
28. A. Villringer, J. Planck, C. Hock, L. Schleinkofer, and U. Dirnagl, "Near infrared spectroscopy (NIRS): A new tool to study hemodynamic changes during activation of brain function in human adults," *Neurosci. Lett.* **154**, 101–104 (1993).
29. Y. Hoshi, M. Shimada, C. Sato, and Y. Iguchi, "Reevaluation of near-infrared light propagation in the adult human head: implications for functional near-infrared spectroscopy," *J. Biomed. Opt.* **10**, 1–10 (2005).
30. H. H. Jasper, "The ten-twenty electrode system of the International Federation," *Electroencephalogr. Clin. Neurophysiol.* **10**, 367–380 (1958).
31. M. Okamoto and I. Dan, "Automated cortical projection of head-surface locations for transcranial functional brain mapping," *Neuroimage* **26**, 18–28 (2005).
32. S. L. Bengtsson, F. Ullén, H. H. Ehrsson, T. Hashimoto, T. Kito, E. Naito, H. Forssberg, and N. Sadato, "Listening to rhythms activates motor and premotor areas," *Cortex* **45**, 62–71 (2008).
33. J. A. Grahn and M. Brett, "Rhythm and beat perception in motor areas of the brain," *J. Cogn. Neurosci.* **19**, 893–906 (2007).

34. J. L. Chen, V. B. Penhume, and R. J. Zatorre, "Listening to musical rhythms recruits motor regions of the brain," *Cereb. Cortex* **18**, 2844–2854 (2008).
35. E. Geiser, E. Ziegler, L. Jancke, and M. Meyer, "Early electrophysiological correlates of meter and rhythm processing in music perception," *Cortex* **45**, 93–102 (2009).
36. C. J. Limb, S. Kemeny, E. B. Origoza, S. Rouhani, and A. R. Braun, "Left hemispheric lateralization of brain activity during passive rhythm perception in musicians," *Anat. Rec. Part A* **288**, 382–389 (2006).
37. K. Sakai, O. Hikosaka, S. Miyauchi, R. Takino, T. Tamada, and N. K. Iwata, "Neural representation of a rhythm depends on its interval ratio," *J. Neurosci.* **19**, 10074–10081 (1999).
38. A. Solodkin, P. Hlustik, D. C. Noll, and S. L. Small, "Lateralization of motor circuits and handedness during finger movements," *Eur. J. Neurosci.* **8**, 425–434 (2001).
39. C. Babiloni, F. Carducci, C. Del Gratta, M. Demartin, G. L. Romain, F. Babiloni, and P. M. Rossini, "Hemispherical asymmetry in human SMA during voluntary simple unilateral movements: an fMRI study," *Cortex* **39**, 293–305 (2003).
40. Y. I. Fishman, D. H. Reser, J. C. Arezzo, and M. Steinschneider, "Neural correlates of auditory stream segregation in primary auditory cortex of the awake monkey," *Hear. Res.* **151**, 167–187 (2004).
41. C. Micheyl, B. Tian, R. P. Carlyon, and R. P. Rauschecker, "Perceptual organization of tone sequences in the auditory cortex of awake macaques," *Neuron* **488**, 139–148 (2005).
42. A. Gutschalk, C. Micheyl, J. R. Melcher, A. Rupp, M. Scherg, and A. J. Oxenham, "Neuromagnetic correlates of streaming in human auditory cortex," *J. Neurosci.* **25**, 5382–5388 (2005).
43. E. C. Wilson, J. R. Melcher, C. Micheyl, A. Gutschalk, and A. J. Oxenham, "Cortical fMRI activation to sequences of tones alternating in frequency: relationship to perceived rate and streaming," *Neurophysiology* **97**, 2230–2238 (2007).
44. M. G. Woldorff, C. C. Gallen, S. A. Hampson, S. A. Hillyard, C. Pantev, D. Sobel, and F. E. Bloom, "Modulation of early sensory processing in human auditory cortex during auditory selective attention," *Proc. Natl. Acad. Sci. U.S.A.* **90**, 8722–8726 (1993).
45. N. Fujiwara, T. Nagamine, M. Imai, T. Tanaka, and H. Shibasaki, "Role of the primary auditory cortex in auditory selective attention studied by whole-head neuromagnetometer," *Cogn. Brain Res.* **7**, 99–109 (1998).
46. J. S. Lewin, L. Friedman, and D. Wu, "Cortical localization of human sustained attention: detection with functional MR using a vigilance paradigm," *J. Comput. Assist. Tomogr.* **20**, 695–701 (1996).
47. K. R. Pugh, B. A. Shaywitz, S. E. Shaywitz, R. K. Fulbright, D. Byrd, P. Skudlarski, D. P. Shankweiler, L. Katz, R. T. Constable, J. Fletcher, C. Lacadie, K. Marchione, and J. C. Gore, "Auditory selective attention: an fMRI investigation," *Neuroimage* **4**, 159–173 (1996).
48. R. P. Carlyon, R. Cusack, J. M. Foxton, and I. H. Robertson, "Effects of attention and unilateral neglect on auditory stream segregation," *J. Exp. Psychol.* **27**, 115–127 (2001).
49. J. S. Snyder, C. Alain, and T. W. Picton, "Effects of attention on neuroelectric correlates of auditory stream segregation," *J. Cogn. Neurosci.* **18**, 1–13 (2006).