Bottom-Up Driven Involuntary Auditory Evoked Field Change: Constant Sound Sequencing Amplifies But Does Not Sharpen Neural Activity

Hidehiko Okamoto, Henning Stracke, Lothar Lagemann, and Christo Pantev
Institute for Biomagnetism and Biosignalanalysis, University of Muenster, Muenster, Germany

Submitted 27 May 2009; accepted in final form 31 October 2009

Okamoto H, Stracke H, Lagemann L, Pantev C. Bottom-up driven involuntary auditory evoked field change: constant sound sequencing amplifies but does not sharpen neural activity. J Neurophysiol 103: 244–249, 2010. First published November 4, 2009; doi:10.1152/jn.00530.2009. The capability of involuntarily tracking certain sound signals during the simultaneous presence of noise is essential in human daily life. Previous studies have demonstrated that top-down auditory focused attention can enhance excitatory and inhibitory neural activity, resulting in sharpening of frequency tuning of auditory neurons. In the present study, we investigated bottom-up driven involuntary neural processing of sound signals in noisy environments by means of magnetoencephalography. We contrasted two sound signal sequencing conditions: “constant sequencing” versus “random sequencing.” Based on a pool of 16 different frequencies, either identical (constant sequencing) or pseudorandomly chosen (random sequencing) test frequencies were presented blockwise together with band-eliminated noises to nonattending subjects. The results demonstrated that the auditory evoked fields elicited in the constant sequencing condition were significantly enhanced compared with the random sequencing condition. However, the enhancement was not significantly different between different band-eliminated noise conditions. Thus the present study confirms that by constant sound signal sequencing under nonattentive listening the neural activity in human auditory cortex can be enhanced, but not sharpened. Our results indicate that bottom-up driven involuntary neural processing may mainly amplify excitatory neural networks, but may not effectively enhance inhibitory neural circuits.

INTRODUCTION

In daily life, we are usually simultaneously exposed to several sounds. However, we can effortlessly separate these concurrent sound signals during perception. Top-down driven voluntary attention and bottom-up driven involuntary neural processing seem to contribute to this process.

Recent magnetoencephalography (MEG) (Okamoto et al. 2007c; Stracke et al. 2009) and intracranial electroencephalography (Bidet-Caulet et al. 2007) studies have demonstrated that focused auditory attention not only amplified, but also sharpened population-level neural responses within human auditory cortex, improving auditory performance in noisy environments. Moreover, another study (Okamoto et al. 2009b) revealed that these gain and sharpening effects of focused attention depend on the sequencing of the test sounds used. As compared with a random sequencing (RS) condition, in which test sounds with different frequencies were presented randomly, a constant sequencing (CS) condition, in which identical test sounds were presented in succession, resulted in larger neural responses and sharper population-level frequency tuning of auditory neurons. In this study, however, participants had always focused their attention on the sound signals, during both CS and RS conditions. Therefore the question remains unanswered whether the auditory evoked response changes were caused by top-down driven voluntary attentional effects or bottom-up driven involuntary sound signal sequencing effects.

Not only top-down but also bottom-up driven neural activities play an important role for auditory neural processing. Even in noisy environments, we can automatically segregate a specific auditory stream from other streams based on some auditory features (auditory scene analysis; Alain 2007; Alain and Bernstein 2008; Bregman 1990; Micheyl et al. 2007; Snyder and Alain 2007) without voluntary focused auditory attention (Nager et al. 2003; Sussman et al. 2007). Bottom-up driven involuntary neural processing seems to play an important role for auditory scene analysis. Among the well-known event-related potentials/fields evoked by bottom-up driven neural processing is the mismatch negativity, which involuntarily appears about 100 to 200 ms after the occurrence of a change in a series of repeatedly presented identical auditory stimuli (Näätänen and Winkler 1999; Picton et al. 2000). The mismatch negativity seems to reflect the violation of an automatically (i.e., without focused attention) generated trace left by bottom-up auditory inputs. However, it still remains elusive whether bottom-up driven involuntary neural processing could amplify and sharpen population-level neural activity.

The goal of the present study was to investigate bottom-up driven sound signal sequencing effects on the auditory evoked N1m response. To measure the population-level frequency tuning in human auditory cortex, we used test stimuli (TS) that were overlaid with band-eliminated noises (BENs), which contained a spectral stop-band centered at the TS frequency (Okamoto et al. 2007c, 2009b; Sams and Salmelin 1994; Stracke et al. 2009).

METHODS

Subjects

Fifteen healthy subjects (nine females; age range: 21–29 yr) participated in the present study. All subjects were right-handed, as assessed with the Edinburgh Handedness Inventory (Oldfield 1971) and had normal hearing. All subjects were fully informed about the study and gave written informed consent for their participation in accordance with procedures approved by the Ethics Commission of the Medical Faculty, University of Muenster.

Address for reprint requests and other correspondence: H. Okamoto, Institute for Biomagnetism and Biosignalanalysis, University of Muenster, Medizinweg 15, 48149 Muenster, Germany (E-mail: okamotoh@uni-muenster.de).
Stimuli and experimental design

The experimental stimuli and design are similar to those of a previous study (Okamoto et al. 2009b), except for subjects’ attentional states. Test stimuli (TS) were pure tones with a frequency of either 250, 350, 450, 570, 700, 840, 1000, 1170, 1370, 1600, 1850, 2150, 2500, 2900, 3400, or 4000 Hz. These frequencies were chosen based on the critical band (CB) scale (Zwicker and Fastl 1999). The TS had a duration of 600 ms with 10-ms onset and offset ramps. All of the TS were presented simultaneously with band-eliminated noises (BENs) (cf. Fig. 1 and Supplemental Materials S1 and S2). Initially, the simultaneously presented BENs were 8,000-Hz low-pass filtered from a white noise (sampling rate: 48,000 Hz), then spectrally notched with widths of either 1/4CB, 1/2CB, or 1CB, centered at the frequency corresponding to the simultaneously presented TS. All BENs had a duration of 3,000 ms (10-ms rise and fall ramps). TS started 2.200 ms after the BEN onset and ceased 200 ms prior to the BEN offset (cf. Fig. 1). We used Presentation (Neurobehavioral Systems, Albany, CA) to control the timing of sound presentation and SRM-212 electrostatic earphones (Stax, Saitama, Japan) to transduce air-conducted sounds. All sounds were delivered through silicon tubes (length: 60 cm; inner diameter: 5 mm) and silicon earpieces, which were adjusted to fit each individual subject’s ears.

Before starting the MEG data acquisition, each subject’s hearing threshold for the 1,000-Hz frequency TS was determined for each ear. During the MEG session, the 1,000-Hz TS was presented at an intensity of 35 dB above individual sensation level, and other TS were presented at the same power level as that of the 1,000-Hz TS. The total power of the simultaneously presented BENs was adjusted to be 15 dB larger than TS power. All of the TS and BENs were presented diotically.

To investigate the effects of bottom-up driven involuntary neural processing on the auditory evoked response, we contrasted two different sequencing conditions within subjects under passive listening (Fig. 1): “constant sequencing” (CS; cf. Supplemental Material S1) and “random sequencing” (RS; cf. Supplemental Material S2). In the constant sequencing block, 30 identical frequency TS were successively presented simultaneously with the pseudorandomly selected 1/4, 1/2, or 1 CB BEN. The TS frequency was different between blocks (250, 350, 450, 570, 700, 840, 1,000, 1,170, 1,370, 1,600, 1,850, 2,150, 2,500, 2,900, 3,400, or 4,000 Hz, respectively). In the RS session, 30 TS were pseudorandomly chosen from the above-listed frequencies and presented together with a BEN. During MEG measurements, subjects were instructed to watch a self-chosen silent movie and fix their eyes on the screen. Crucially, the inputs were balanced between CS and RS conditions. Only the patterning of auditory stimuli was different: constant or random. The first block of the MEG measurement was pseudorandomized between CS and RS and CS and RS blocks alternated after that. In total, 160 trials (10 trials for 16 frequencies) for each BEN condition (1/4CB, 1/2CB, and 1CB) in each sequencing condition (CS and RS) were presented and subjected to data analysis.

Data acquisition and analysis

We obtained auditory evoked fields with a whole-head 275 channels MEG system (Omega; CTF Systems, Coquitlam, British Columbia, Canada) in a magnetically shielded and acoustically silent room. Subjects were instructed not to move their heads and their compliance was monitored through video camera by the experimenter. The magnetic fields were digitally recorded (sampling rate = 600 Hz). The magnetic fields evoked by TS were averaged selectively for each BEN and for each sequencing condition (irrespective of frequency), starting 300 ms prior to TS onset and ending 300 ms after TS onset. Epochs containing field changes >3 pT were rejected as artifact-contaminated epochs. Thereafter, 152.8 ± 7.3 (mean ± SD) of 160 epochs were averaged for each condition. The source locations and orientations of the auditory evoked fields were determined in a head-based Cartesian coordinate system. The origin was located at the middle of the mediolateral–lateral, connecting the center of the entrances of the left and right ear canals. The posterior–anterior axis connected nasion and origin and the inferior–superior axis ran through the origin perpendicularly to the mediolateral and posterior–anterior plane.

For the analysis of the N1m component, which is the major reflection of the auditory evoked response (Na¨a ¨ta¨nen and Picton 1987), the averaged auditory evoked fields were 30-Hz low-pass filtered and the baseline was corrected relative to the 300-ms prestimulus interval. Initially, the peak N1m response was identified as the maximal root-mean-square value of the global field power around 150 ms after TS onset. The 10-ms time window around the N1m peak was used for source estimation. The source locations and orientations were estimated by means of single equivalent current dipole modeling (one dipole for each hemisphere) for each subject individually. The estimated source was fixed in its location and orientation for each
hemisphere of each subject and served as spatial filter (Tesche et al. 1995) during the calculation of the source strength for each BEN (BEN_1/4CB, BEN_1/2CB, and BEN_1CB) and each stimulus sequencing (CS and RS) condition, irrespective of the TS frequency. The maximal source strength in each hemisphere and BEN condition in the time range between 100 and 300 ms was used for statistical analysis. When the source strength waveform contained several peaks, the peak with the latency closest to the average peak latency across the single peaks for this condition and hemisphere was selected as N1m response.

To evaluate bottom-up driven involuntary effects on the N1m response, the maximum source strengths and latencies of the N1m responses elicited by the TS for each condition were analyzed separately via repeated-measures ANOVA, using the three factors BEN-TYPE (BEN_1/4CB, BEN_1/2CB, and BEN_1CB); HEMISPHERE (Left and Right), and SEQUENCING (Constant and Random). However, bottom-up driven effects on the N1m might not necessarily be additive, but multiplicative. Therefore we additionally normalized the N1m source strength in each condition with respect to the mean of the N1m source strengths in all BEN-TYPE and SEQUENCING conditions for each hemisphere and for each subject. The normalized N1m source strengths were also analyzed via repeated-measures ANOVA using three factors (BEN-TYPE, HEMISPHERE, and SEQUENCING).

RESULTS

Clear dipolar patterns over left and right hemispheres were observed in all subjects (cf. Fig. 2); we therefore used the single-dipole source estimation method. Furthermore, the source location estimation results obtained with standardized low-resolution brain electromagnetic tomography (sLORETA; Pascual-Marqui 2002; Pascual-Marqui et al. 2002) were focused at the estimated equivalent current dipole locations within the auditory cortex (Fig. 2). The goodness-of-fit of the underlying dipolar source model for the grand-averaged MEG waveforms of all sensors was 94.6 ± 2.6% (mean ± SD). The estimated dipolar sources were located to the superior temporal plane, a location that corresponds to the N1m generator (Eggersmont and Ponton 2002; Pantev et al. 1995).

N1m cortical source strength and latency

The time courses (time range from -100 to +300 ms) of the source strengths averaged across all subjects and hemispheres are displayed in Fig. 3. The N1m responses are larger and have shorter latencies in the CS and the wide BEN conditions compared with those in the RS and the narrow BEN conditions.

Figure 4 represents the mean N1m source strengths and latencies for the left and right hemispheres in each condition with 95% confidence limits. The repeated-measures ANOVAs evaluating N1m source strength and N1m latency resulted in significant main effects for HEMISPHERE [Source strength: $F(1,14) = 5.54, P < 0.04$; Latency: $F(1,14) = 0.11, P = 0.75$], SEQUENCING [Source strength: $F(1,14) = 24.98, P < 0.001$; Latency: $F(1,14) = 50.99, P < 0.001$], and BEN-TYPE [Source strength: $F(2,28) = 20.96, P < 0.001$; Latency: $F(2,28) = 54.36, P < 0.001$]. The N1m source strengths were significantly larger in the left hemisphere, during constant sequencing, and in the wide BEN conditions. However, there was no significant interaction between SEQUENCING and BEN-TYPE [Source strength: $F(2,28) = 0.226, P = 0.80$; Latency: $F(2,28) = 2.99, P = 0.07$].

The repeated-measures ANOVA evaluating the N1m source strength ratio (the N1m source strength was normalized with respect to the mean N1m source strength) resulted in patterns similar to those of the N1m source strength. There were significant main effects for SEQUENCING [$F(1,14) = 18.46, P < 0.001$] and BEN-TYPE [$F(2,28) = 49.80, P < 0.001$]. There was no significant interaction between SEQUENCING and BEN-TYPE [$F(2,28) = 0.86, P = 0.43$]. The significant hemispheric difference observed in the N1m source strength was not found for the N1m source strength ratio due to the normalization procedure.

The nonsignificant interactions between SEQUENCING and BEN-TYPE for the N1m source strength and N1m source strength ratio demonstrate that the N1m response was similarly amplified by the constant presentation of TS compared with the random TS presentation, irrespective of the type of the simultaneously presented BEN. The present results reflect a significant additive gain effect driven by constant sound signal sequencing. However, dissimilar to the previous studies (Okama...
We hypothesized that constant stimulus sequencing would enhance bottom-up driven neural processing, which let the listeners involuntarily focus their resources on neurons processing the constant test sounds. Our present results support this hypothesis by providing experimental evidence that bottom-up driven sensory inputs can amplify neural activity: the N1m responses were significantly larger when the test stimulus (TS) was presented in constant sequencing (CS) compared with random sequencing (RS), irrespective of the notch width of the simultaneously presented band-eliminated noises (BENs).

In the present study, we used the combination of TS and simultaneously presented BEN (cf. Fig. 1 and Supplemental Materials S1 and S2) to investigate population-level gain and sharpening effects on the auditory evoked N1m response. The neural activity evoked by the TS partially overlapped with the neural activity evoked by the BENs. However, the N1m response that was analyzed in the present study does not stem from this overlapping neural activity, since the neurons that are evocable by both TS and BEN had already been activated by the continuously presented BEN when TS appeared. Thus they did not contribute to the auditory evoked fields elicited by the TS onset. The neural activity overlap would be rather large and thus a comparably small N1m response would be elicited by the TS onset. As described in our previous studies (Okamoto et al. 2007c, 2009b; Stracke et al. 2009), the neural gain effect driven by constant sequencing is reflected in an additive increment of N1m source strength across BEN conditions, whereas sharpening of population-level frequency tuning results in larger N1m source strength gain in the narrow BEN compared with the wider BEN conditions. The significant main effect of SEQUENCING found in the present study confirmed that constant sequencing could enhance the corresponding neural activity. However, since a significant interaction between SEQUENCING and BEN-TYPE was missing, a sharpc

![Source Waveform](image)

**FIG. 3.** Mean of source strength waveforms across all subjects ($n = 15$) and both hemispheres. Solid lines represent the CS condition and dotted lines represent the RS condition. Line thickness represents the BEN condition (thick: 1 critical band; medium: 1/2 critical band; thin: 1/4 critical band).

moto et al. 2007c, 2009b; Stracke et al. 2009), a significant sharpening effect of constant sequencing cannot be seen in the present results.

**DISCUSSION**

We hypothesized that constant stimulus sequencing would enhance bottom-up driven neural processing, which let the listeners involuntarily focus their resources on neurons pro-

![Source Waveform](image)

**FIG. 4.** Group means ($n = 15$) of the N1m source strengths (top graph) and latencies (bottom graph) in the left and the right hemispheres for each experimental condition. The error bars denote the 95% confidence limits of the difference to the mean across all conditions. Filled circles denote the N1m responses elicited by the test stimulus during the CS condition and open circles denote the N1m responses during the RS condition.
ening effect of constant sequencing on neural activity could not be observed here (Fig. 4).

In the present study, during CS blocks the subjects could foresee the frequency of the upcoming TS, whereas this was impossible in the RS condition. The repeated presentation of identical auditory signals in a noisy environment could have driven the constitution of a specific auditory object, which could then have been effortlessly segregated from the noise by the listeners (Alain and Arnott 2000; Bregman 1990); however, this was impossible in the case of the RS condition. The auditory object formation might have amplified neural activity corresponding to the constant test sound and might have facilitated the involuntary tracking of the object in the noisy environment. In a silent environment, the repeated presentation of an identical auditory stimulus would lead to a gradual decline of evoked N1m response amplitude; the presentation of a deviant auditory stimulus would elicit a reenlarged N1m response as well as a mismatch negativity (Garrido et al. 2009; Jääskeläinen et al. 2004). The present results, showing larger N1m responses during the CS condition compared with the RS condition, seem contradictory to these previous findings at first sight. However, it could be that in a silent situation, bottom-up driven involuntary neural processes would suppress neural activity corresponding to repeatedly presented auditory signals, to allow for an effective picking up of deviant sound signals, which might be salient for the listener. In contrast, in a noisy environment bottom-up driven involuntary neural processes might facilitate the tracking of a repeatedly presented sound signal by enhancing corresponding neural activity. This mechanism would enable effective tracking and change detection of regular auditory signals in a noisy environment.

Another possibility is that stimulus-specific adaptation might have differentially reduced the neural activity between the CS and RS conditions, since successive presentation of identical sound signals reduces the activity of the corresponding neurons (Brosch and Schreiner 1997; Butler 1968; Perez-Gonzalez et al. 2005; Sörös et al. 2009; Ulanovsky et al. 2003, 2004). However, in the present study the neural responses elicited by the repeatedly presented identical test sounds (constant sequencing) were larger than those evoked by the randomly presented sounds (random sequencing). Therefore stimulus-specific adaptation corresponding to the test stimuli (TS) cannot explain the obtained results. The simultaneously presented BEN, containing a spectral notch around the TS frequency, was identical between the CS and RS conditions. However, in the RS condition, 94% of the preceding (not the simultaneously presented) BENs overlapped in frequency spectrum with the subsequent TS. On the contrary, none of the preceding BENs had any spectral overlap with the subsequent TS in the CS condition. Thus the receptive fields overlapping between the preceding BEN and the consequent TS in the RS condition might have caused stronger adaptation of auditory neurons corresponding to the TS and smaller N1m responses compared with the CS condition. However, our previous study (Okamoto et al. 2005) demonstrated that a preceding noise, which is band-eliminated around 1,000 Hz, could effectively decrease the N1m response amplitude elicited by a following 1,000-Hz pure tone, possibly via the lateral inhibitory system. Additionally, the long sound onset asynchrony between the termination of a preceding BEN and the subsequent TS onset (2,200 ms) and the simultaneously presented BEN would have decreased the effect of the preceding BEN on the consequent TS. Previous studies have demonstrated that the effects of adaptation and lateral inhibition decay when the time interval between sounds is increased (Brosch and Schreiner 1997; Okamoto et al. 2004). Moreover, the auditory evoked fields elicited by the combination of BEN offset and onset did not differ between the CS and RS conditions. Therefore it seems unlikely that the BEN offset and onset caused the differential TS onset-related N1m responses between the CS and RS conditions.

In the present study, we observed significantly larger N1m source strengths in the left compared with the right hemisphere, as previously shown (Okamoto et al. 2007b,c, 2009b; Stracke et al. 2009). In a silent environment, the N1m source strength elicited by a pure tone is usually similar between hemispheres or even right hemispheric dominant (Kanno et al. 1996; Roberts et al. 2000). Thus the left hemisphere seems to play a particularly important role in processing auditory signals in noisy environments (Okamoto et al. 2007b). This left hemispheric dominance in noisy environments might originate from shorter temporal integration windows in the left compared with the right hemisphere (Belin et al. 1998; Poeppel 2003), causing left hemispheric dominance for temporal processing and right hemispheric dominance for spectral processing (Jamison et al. 2006; Okamoto et al. 2009a; Zatorre and Belin 2001).

In conclusion, our present findings complement the results of our previous study (Okamoto et al. 2009b), which used almost identical auditory stimuli and identical experimental settings. However, whereas participants had listened attentively in the previous study, they listened nonattentively here. It may be that top-down voluntary auditory attention not only enhances excitatory neural activity, but also activates inhibitory neural circuits within the auditory pathway, which evidently play an important role in sharpening frequency tuning on both the single-cell (Higley and Contreras 2006; Sutter and Loftus 2003; Sutter et al. 1999; Wehr and Zador 2003; Wu et al. 2008) and the population levels (Okamoto et al. 2004, 2005, 2007a; Pantev et al. 2004; von Békésy 1967). In contrast, it seems that bottom-up driven involuntary neural processes can indeed significantly enhance neural activity, but may have merely little effect on population-level frequency tuning. Therefore bottom-up driven involuntary neural processes might ease the automatic tracking of auditory signals without strengthening of inhibitory mechanisms, which may suppress neural activity elicited by other, potentially salient sound signals.

ACKNOWLEDGMENTS

We thank A. Wollbrink and K. Berning for technical assistance.

GRANTS

This work was supported by Deutsche Forschungsgemeinschaft Grant Pa 392/10-3 and Tinnitus Research Initiative Grant BD169968.

REFERENCES


