

Neural Mechanisms of the Time-order Error: An MEG Study

Ilana S. Hairston¹ and Srikantan S. Nagarajan²

Abstract

■ The time-order error (TOE) refers to the influence of presentation order on performance accuracy in a discrimination task. Despite it being a well-documented perceptual bias, the underlying mechanisms have not been studied. In this study, observers were trained on a two-interval forced-choice procedure. The stimuli presented for discrimination were a standard, consisting of four tones presented at a 5-Hz rate, and targets, consisting of various rates higher than 5 Hz. Psychometric functions were measured for discrimination of the trained standard and targets, a novel standard of 13 Hz with higher target rates; and the trained 5 Hz standard with novel targets with rates below 5 Hz. Discrimination did not improve with training; in fact, accuracy declined when standard was presented in the first

interval during the session, resulting in a TOE. The TOE was specific to the 5-Hz standard generalizing to the novel targets slower than 5 Hz, but not to the 13-Hz STANDARD. Analysis of the event-related magnetic field responses (ERFs) revealed a waveform to the whole stimulus, rather than to each tone in the train. Although ERFs in the second interval were attenuated independent of stimulus type, the M300 component in the second interval was attenuated only when the standard was first, but remained of equivalent magnitude when the standard was second. This was observed only in the two 5-Hz conditions. Combined, these results suggest that the TOE reflects the emergence of an internal representation of the standard, and that the M300 is potentially a neural correlate of plasticity. ■

INTRODUCTION

The time-order error (TOE) is a population bias in discriminating between successively presented stimuli arising from the order of presentation (Allan & Gibbon, 1994; Hellström, 1985; Needham, 1935). The phenomenon was first described in the mid 19th century by Gustav Fechner, who demonstrated judgment asymmetry when determining just noticeable differences (JND) in weight estimation (Gescheider, 1997); that is, observers tended to overestimate the weight of the second stimulus, resulting in an order-dependent bias in their JND.

TOE is most often reported in two-alternative forced-choice (2AFC) discrimination tasks, where observers are presented with *standard* and *comparison* stimuli, and required to determine either if they are different, or the relationship between the two (e.g., “which is brighter?”). However, the phenomenon is not restricted to discrimination paradigms. For example, Nakajima et al. (2004) and Sasaki, Suetomi, Nakajima, and ten Hoopen (2002) asked observers to reproduce the duration of a time-interval standard, when the standard either preceded or followed another longer or shorter interval. Observers demonstrated biased estimates of the duration of the standard depending on order of presentation.

Similarly, TOE is reported for several perceptual modalities (for review, see Hellström, 1985), as well as in higher cognitive judgments such as valence. Koh (1967), for example, reported that observers preferred the second of two musical excerpts, previously rated as having similar pleasantness. Likewise, Hellström (2001) reported that a second of two visual stimuli was judged more attractive than the first. However, most investigations of the TOE focused on the effects of order on interval discrimination. Pioneered by Wolfgang Köhler in the early 1920s (see Woodrow, 1935), and revived in studies by contemporary researchers (e.g., Grondin, 2001; Allan, 1977; Jamieson & Petrusic, 1975), TOEs have been reliably reported (Allan, 1979).

Despite the robustness of the TOE phenomenon, its underlying neural mechanisms have not been explored. Here we used rate-modulated tone trains in a 2AFC discrimination task, where evoked responses to the stimuli were used to assess the neural correlates of TOE. The tone trains containing four 1-kHz tones, where observers were required to discriminate between modulation rates. We tested whether the TOE was a function of familiarity with the *standard* and/or *comparison* stimuli. We found that observers were more accurate when the standard was presented in the second interval, and only after training, suggesting that the TOE is an acquired response. Analysis of the event-related magnetic fields (ERFs) showed that the behavioral TOE was associated

¹University of California, Berkeley, ²University of California, San Francisco

with modulation of the M300 component, suggesting a neural correlate of this behavioral phenomenon.

METHODS

The experiment was performed in accordance with the Declaration of Helsinki, and was approved by the UCSF Office for the Protection of Human Subjects. Eleven healthy participants (7 men, 4 women), mean age 29.6 years (range: 22–40 years), were trained on an auditory 2AFC modulation rate discrimination task. A trial consisted of two trains of four 1-kHz tones (25 msec long, 5 msec ramp, 75 dB SPL) each, presented about 1 sec apart. One train, standard, was presented at a 5-Hz (200 msec interstimulus interval [ISI]) rate. The standard was presented on all trials, and was meant as a comparison stimulus. The second train, target, was presented at 5.125, 5.25, 5.5, 5.75, or 6 Hz. These rates were chosen to obtain a psychometric function. On each trial, observers were instructed to identify the slower rate, which always was the standard. The order of standard and target tone trains randomly varied, and observers indicated their choice by pressing one of two buttons (left, train in first interval was slower; right, train in second interval was slower). Observers received visual feedback on each trial indicating their accuracy, and the next trial began within 20 msec thereafter (Figure 1). Thus, the duration of each trial was approximately 3.5 sec. There were seven blocks of 100 trials, and an additional test block of 150 trials. To assess whether the TOE was associated with familiarity with the standard or with the

task, two additional stimulus conditions were presented for a single block of 150 trials each. In one, a 13-Hz standard was used with targets of 13.65, 14.3, 14.95, 15.6, 16.25 Hz. In the other, the standard was the same 5 Hz, but the targets were slower (4, 4.25, 4.5, 4.75, 4.875 Hz). In this block, observers were still required to detect the slower of the two comparison stimuli, however, the correct response was, in this case, always the target. This condition is referred to as 5 Hz-inv. Percent correct responses for each block were quantified separately for presentation order (standard 1st, standard 2nd).

The experiment was conducted in a magnetic shielded room. Visual instructions were projected on a screen and tones were delivered through pneumatic earphones connected to a custom-built auditory stimulus presentation system (Tucker Davis Technologies). Magnetoencephalographic (MEG) data were collected on 275-channel SQUID array (CTF Systems, Port Coquitlam, Canada). Acquisition was continuous for each block, which lasted 4 to 6 min. Fiducial skin markers were placed on the participants' nasion and at bilateral preauricular points for coregistration with the participants' structural magnetic resonance images, acquired on a different day.

ERFs were averaged off-line, baseline was set to zero, and filtered with a band-pass filter of 1–40 Hz. Averaged ERFs were divided by type and order (standard 1st; standard 2nd; target 1st; target 2nd) for each of the stimulus conditions (i.e., 5 Hz, 13 Hz, 5 Hz-inv). For analysis of latency and amplitude of ERFs, only temporal sensors were used. To correct for intersubject variability, peak values were normalized to percent change of mean root

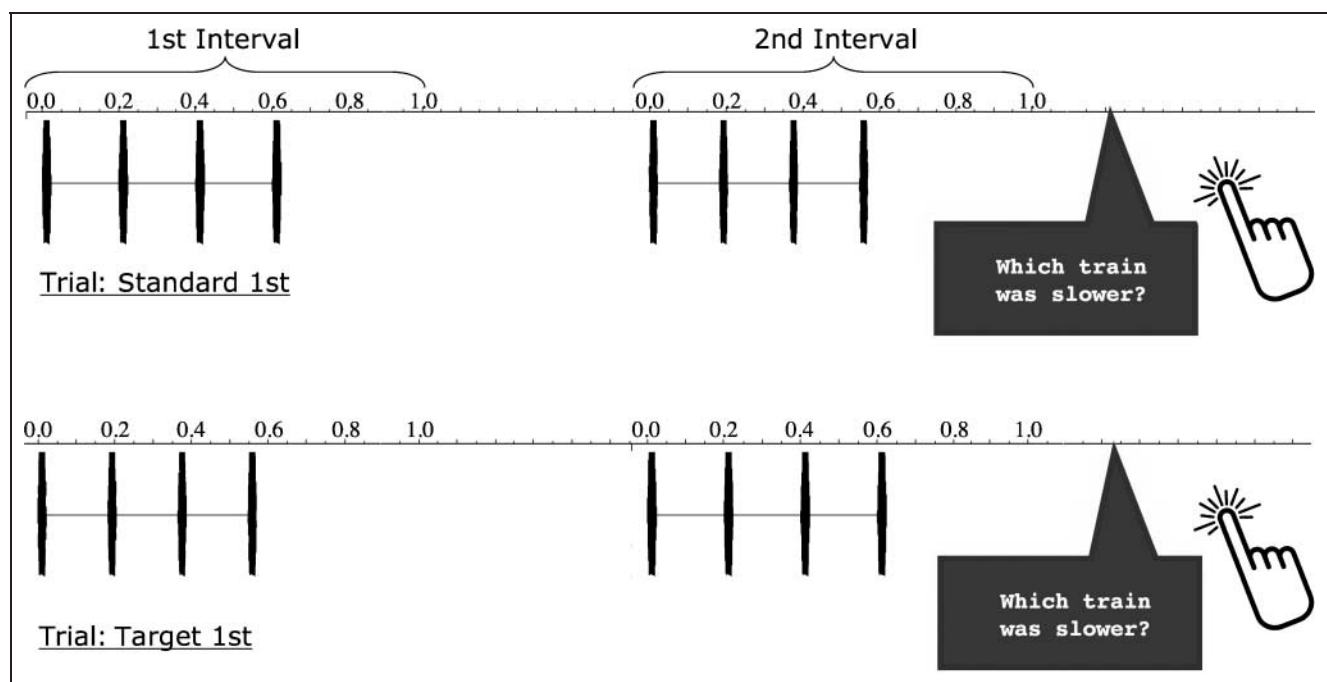


Figure 1. Structure of a trial: Cartoon of 5-Hz trials. Top trace: Standard 1st; Bottom: Standard 2nd. In all conditions, observers were presented on each trial with the standard stimulus and one of the targets, and were required to determine which of the stimuli was at a slower rate. About 1 sec after the second stimulus, observers were asked to make a judgment and respond by pressing the appropriate button.

mean squared (RMS) within each block, and the threshold amplitude for a component was 15% greater than the mean RMS. For each tone train, 900 msec from onset of the first tone was used for analysis of MEG data where the amplitude and latency of the peaks were characterized.

To localize the neuronal sources of the magnetic fields, an equivalent current dipole (ECD) model, in a spherical volume conductor, using all the sensors, was applied to M50, M100, M200, and M300 components of the ERF.

RESULTS

Discrimination Accuracy

To assess whether discrimination accuracy improved during the session, a repeated-measures analysis of variance (ANOVA) with blocks as the repeated-measures

variable, and rate and order of presentation as the independent variables was done on the 5-Hz condition. A main effect of order was found [$F(1, 120) = 4.171, p = .054$], but there was no main effect of blocks [$F(6, 120) = 1.463, p = ns$]. There were no significant interactions of block with order [$F(6, 120) = 1.416, p = ns$]; nor block with target rates [$F(24, 480) = 0.882, p = ns$]. A significant interaction of rate with order was found [$F(1, 480) = 30.226, p < .001$] due to a decline in accuracy in the slowest targets when standard was presented first [5.125 Hz: $F(1, 120) = 29.626, p < .001$; 5.25 Hz: $F(1, 120) = 14.761, p = .001$]. Thus, observers demonstrated diminished accuracy in an order-dependent manner, wherein observers performed better when the target was the first stimulus (Figure 2).

All three stimulus conditions yielded psychometric functions with performance at ~90% accuracy at the

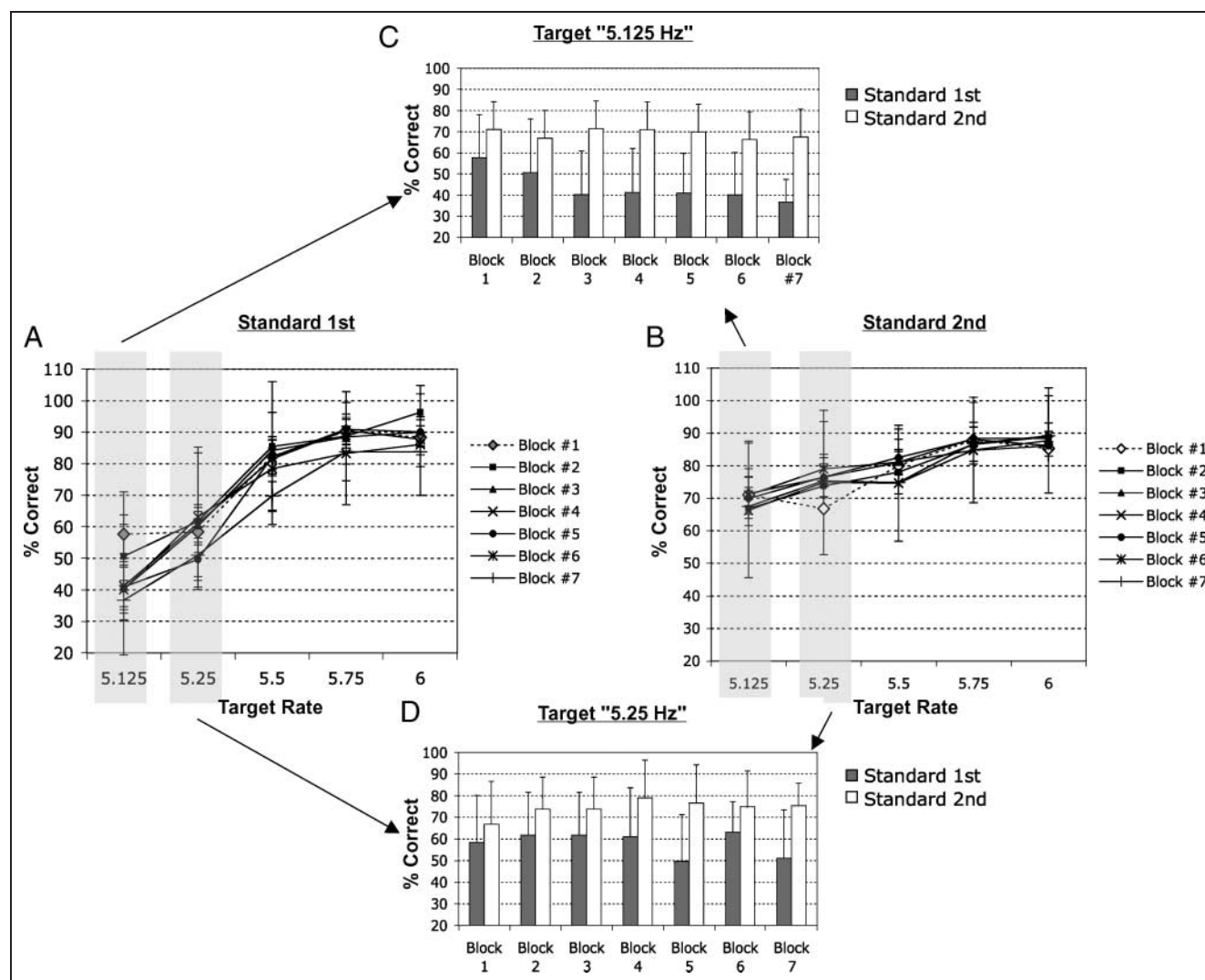


Figure 2. Effects of training on discrimination accuracy. Observers were trained on 7 blocks of the 5-Hz condition. Performance did not improve across the session. (A) Performance on trials where standard was first; (B) Performance on trials where standard was second. Performance declined when standard was first due to reduced accuracy in the hardest discrimination trials (shaded areas in A, B). (C) Performance accuracy in the hardest discrimination trials (Target = 5.125 Hz) over the session; (D) Performance accuracy in the second hardest discrimination trials (Target = 5.25 Hz) over the session.

easiest discrimination trials, and around chance on the hardest trials (Figure 3). In both the last block of 5-Hz and in the 5-Hz-inv conditions, a significant order effect was observed, with better performance in trials where the standard was second [5 Hz, paired $t(449) = -5.68$, $p < .0001$, Figure 3A; 5 Hz-inv, paired $t(34) = -3.22$, $p = .003$, Figure 3C]. This effect was not apparent in the 13-Hz condition [paired $t(34) = -0.75$, $p = ns$, Figure 3B]. To determine whether TOE was associated with familiarity with the standard, or an inherent property of the lower-frequency stimuli, a similar analysis was performed for the first block in the 5-Hz condition. No order effect was apparent in this block [paired $t(34) = 0.05$, $p = ns$, insert in Figure 3A]. Thus, the TOE developed during the session, and was generalized to the 5-Hz-inv condition, but not to the 13-Hz condition.

Event-related Magnetic Field Response Analysis

ERFs were easily characterized as M50, M100, M200, and M300 (Figure 4A). Averaging was performed separately for left and right hemispheres to standard first, standard second, target first, and target second from the last block of 5 Hz, and the single 13-Hz and 5-Hz-inv blocks. To assess differences in the ERFs to the different stimuli, repeated-measures ANOVA, with ERF components and hemisphere as the within variables, was performed on the latencies and peak amplitude values. Across stimuli, the M100 had the highest amplitude compared to other components [main effect of peak: $F(3, 321) = 30.723$, $p < .001$]. In addition, the responses from the right temporal sensors were stronger [$F(1, 321) = 27.384$, $p < .0001$], and tended to occur earlier than from the left hemisphere [$F(1, 321) = 9.548$, $p = .018$, Figure 4B]. This laterality was most pronounced in the amplitude of the M100 [interaction of amplitude by side: $F(3, 321) = 9.377$, $p < .001$]. The ERFs to the stimulus in the first interval tended to be stronger [$F(1, 321) = 16.030$, $p = .0001$], independent of stimulus type (standard/target) and condition [5 Hz/5 Hz-inv/13 Hz: $F(1, 321) = 0.824$, $p = ns$]; and an interaction of stimulus condition by component was found [$F(6, 321) = 4.239$, $p < .001$] due to higher amplitude responses to the 13-Hz stimuli [post hoc $F(1, 255) = 4.24$, $p = .043$], with no difference between the 5-Hz and 5-Hz-inv conditions [$F(1, 255) = 0.189$, $p = ns$; see also below]. No main effects of stimulus condition or type were found for latency [condition: $F(2, 321) = 2.310$, $p = ns$; type: $F(1, 321) = 1.00$, $p = ns$]. Significant interactions of latency with stimulus condition [$F(2, 321) = 3.229$, $p = .021$], and latency with stimulus condition and side [$F(96, 321) = 3.038$, $p = .027$] were found due to slower M200 to the targets 5 Hz-inv, especially from the right hemisphere (Figure 5F).

To determine whether stimulus condition was reflected in the relative latencies of the different components, the latency to M100 to the 5-Hz and 13-Hz standard, and 5 Hz was subtracted from the latency to M200 and M300.

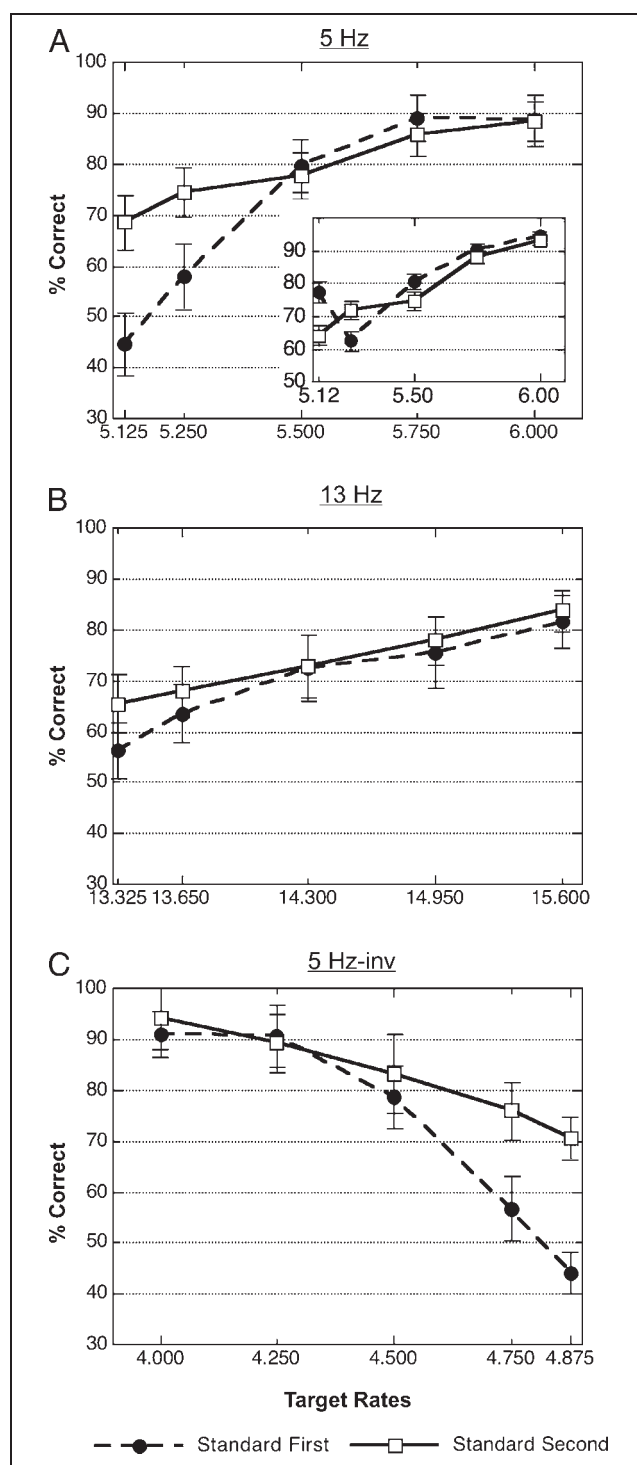


Figure 3. Psychometric functions of performance. (A) The average responses in the last block with base rate of 5 Hz. Observers performed significantly better when standard was second (filled squares). Insert: Average responses in the first block of 5 Hz, no TOE is observed. (B) The average response to the single block of 13-Hz standard; no TOE is observed. (C) The average response to the single block of 5 Hz-inv. Observers performed significantly better when the standard was presented second. *Depicts $p < .005$.

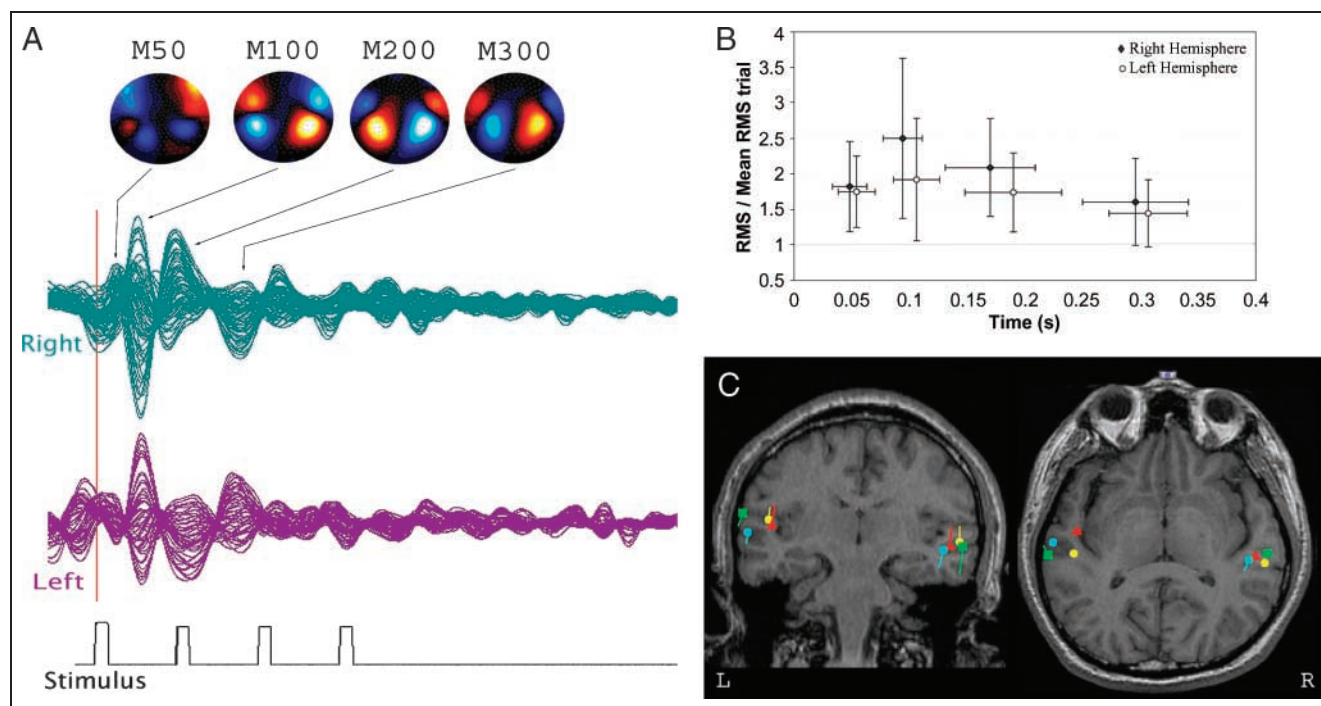


Figure 4. Event-related magnetic fields: (A) The averaged ERF to 5 Hz (bottom line). In this observer, the evoked responses were bilateral for all four components (M50, M100, M200, M300). (B) Right (gray diamonds) and left (white circles) hemisphere corrected RMS values. The abscissa depicts mean latency for each component, and the ordinate depicts the mean peak value. Error bars depict standard deviations for latency (abscissa) and amplitude (ordinate). (C) Results of ECD, averaged across stimulus conditions, from a single participant with bilateral responses. Each line depicts the moment (length) and direction (angle) for the dipoles: Yellow = M50; green = M100; red = M200; cyan = M300.

These temporal distances were analyzed with an ANOVA, with stimulus condition as independent factor. No differences were found in the temporal relationships between the components as a function of modulation rate [right hemisphere—M200 – M100: $F(1, 9) = 1.12, p = ns$; right hemisphere—M300 – M100: $F(1, 9) = 0.12, p = ns$; left hemisphere—M200 – M100: $F(1, 9) = 0.37, p = ns$; right hemisphere—M300 – M100: $F(1, 9) = 0.06, p = ns$, Figure 5].

ECD analysis yielded reliable localizations for 30% to 50% of detected peaks within the right auditory cortex (Brodmann's area 41 and 42; Figure 4C). The direction of the dipole was reversed for neighboring peaks, and statistical analysis of the x, y, z coordinates for the different components revealed that, in the right hemisphere M300 tended to localize more anterior to the other components ($p < .02$), whereas M50 tended to be the most medial ($p < .04$). In the left hemisphere, M100 was the most anterior ($p < .01$), and M100 and M200 were more dorsal than M100 and M50 ($p < .03$). These differences in average location were less than 1 cm. Our localizations of the M100 and M200 were equivalent to previous reports which demonstrated that with frequent stimulation, the main contributing sources of the signal arise from auditory projection regions (Hari, Kaila, Katila, Tuomisto, & Varpula, 1982).

The different components were not always apparent. For each observer, the probability for the occurrence of

each component was calculated across blocks. The M100 component occurred at highest frequency, and M300 at the lowest. The probability for M50 and M100 was sensitive to order, with higher percentages in the first interval [M50: $F(1, 80) = 6.234, p = .015$; M100: $F(1, 80) = 4.142, p = .045$]. M200 was marginally affected by order [$F(1, 80) = 3.644, p = .059$], and M300 not all [$F(1, 90) = 0.907, p = ns$]. All but M50 occurred more frequently in the right hemisphere [M50: $F(1, 80) = 0.393, p = ns$; M100: $F(1, 80) = 11.339, p = .001$; M200: $F(1, 80) = 6.400, p = .013$; M300: $F(1, 80) = 13.051, p = .001$]. Frequency of occurrence was not associated with stimulus type (i.e., standard, target; see Table 1).

The targets in the 5-Hz-inv condition, and both targets and standard in the 13-Hz condition, were relatively novel compared with the standard and targets of the 5-Hz condition, which were presented in the seven training blocks. As noted above, the responses to the 13-Hz stimuli were of higher amplitude compared with the 5-Hz and 5-Hz-inv conditions. To further determine whether amplitudes of the different components reflected the novelty of stimuli, we compared the peak values of the components of the ERF for 5-Hz-inv targets and 5-Hz targets, pooling the responses to the two 5-Hz standard; and performed a similar comparison for 13-Hz standard to 5-Hz standard. Table 2 summarizes the results of a repeated-measures ANOVA with peak values and hemisphere as within-subject variables. ERFs to the two

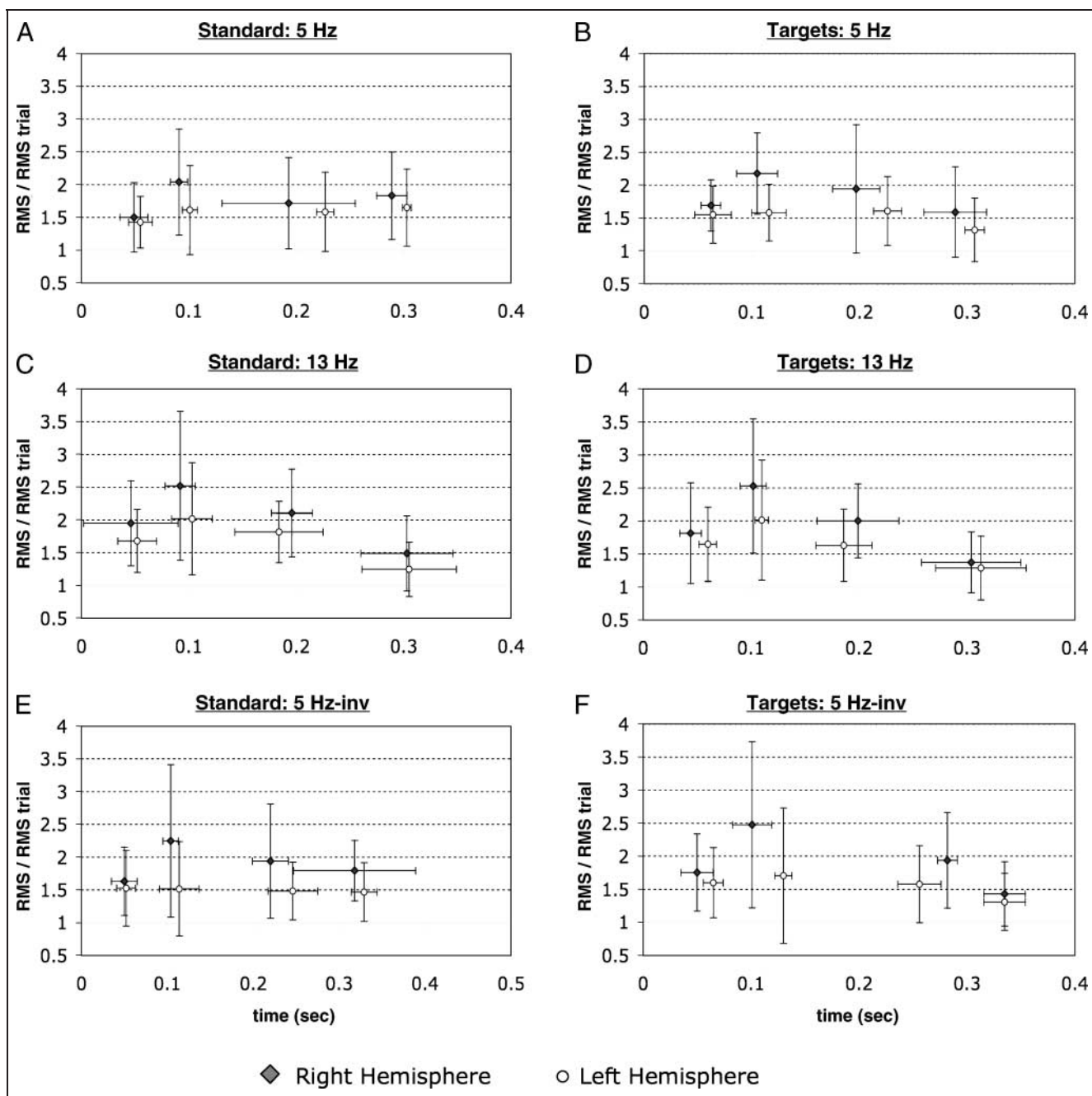


Figure 5. Evoked magnetic fields in different conditions. (A) Standard 5 Hz; (B) Targets 5 Hz; (C) Standard 13 Hz; (D) Targets 13 Hz; (E) Standard 5 Hz-inv; (F) Targets 5 Hz-inv. Right (gray diamonds) and left (white circles) hemisphere corrected RMS values. The abscissa depicts mean latency for each component, and the ordinate depicts the mean peak value. Error bars depict standard deviations for latency (abscissa) and amplitude (ordinate).

targets and pooled standard in the 5-Hz and 5-Hz-inv conditions did not differ significantly [$F(2, 216) = 1.841, p = ns$]. A significant interaction of stimulus type with component was found [$F(6, 216) = 3.251, p = .004$] due to significantly stronger M50 in the 5-Hz-inv condition (Fisher PSLD: Standard – Target 5 Hz-inv = $-0.56, p = .003$; others = ns), and a significantly increased M300 to the standard (Fisher PSLD: Standard – Target 5 Hz-inv = $2.85, p = .048$; Standard – Target 5 Hz = $2.58, p = .047$; Target 5 Hz-inv – Target 5 Hz = $-0.07, p = ns$). As

the M50 occurs prior to the second tone of the trains in the 5-Hz conditions, it is unlikely that this difference reflects encoding of rate. The M300 for both 5-Hz and 5-Hz-inv targets were of similar magnitude, suggesting that the two 5-Hz-inv targets were not perceived as novel (Table 2A). Conversely, a significant interaction was found for the comparison of 5-Hz standard to 13-Hz standard [$F(3, 123) = 7.232, p = .010$]. The interaction was due to significantly higher RMS values to the 13-Hz condition in the right hemisphere at M50, M100,

Table 1. Percentage of Blocks in Which Each Component Occurred (Standard Deviations)

	<i>Frequency of Components of the ERF</i>							
	<i>M50</i>		<i>M100</i>		<i>M200</i>		<i>M300</i>	
	<i>Right</i>	<i>Left</i>	<i>Right</i>	<i>Left</i>	<i>Right</i>	<i>Left</i>	<i>Right</i>	<i>Left</i>
Standard 1st	83.6%	84.8%	94.8%	89.6%	87.1%	83.4%	78.8%	69.2%
	(23.7)	(15.3)	(11.3)	(13.1)	(19.5)	(15.8)	(16.9)	(20.6)
Standard 2nd	76.2%	73.7%	94.1%	79.9%	83.9%	71.0%	76.4%	58.5%
	(21.5)	(20.6)	(11.7)	(17.5)	(22.1)	(17.6)	(15.0)	(23.2)
Target 1st	87.5%	85.5%	95.5%	89.5%	93.9%	81.7%	77.0%	58.3%
	(17.5)	(14.0)	(12.6)	(13.1)	(17.5)	(19.8)	(21.9)	(18.2)
Target 2nd	77.8%	70.2%	94.2%	73.4%	86.2%	73.7%	74.0%	57.8%
	(19.6)	(27.5)	(13.5)	(28.2)	(14.9)	(24.3)	(20.0)	(24.3)

Percentages were calculated separately for order and stimulus type, for all components across stimulus conditions, and averaged across subjects. M100 was the most reliable response, whereas M300 was the least. Lateralization can be observed in the frequency of occurrence of the M100, M200, and M300, but not M50. M50 and M100 responses occurred less frequently in the second interval compared to the first. Frequency of the components was not affected by stimulus type.

and M200, and in the left hemisphere at M100. M300 of the left hemisphere was higher in the 5-Hz condition (Table 2B). Thus, ERFs were stronger only to the novel 13-Hz standard, but not to the novel 5-Hz-inv targets.

To assess whether stimulus type presented in the first interval impacted the event-related responses to the stimulus presented in the second interval, the RMS for each component was divided by the equivalent ERF to

the stimulus in the second interval. The products of the ratio were classified as standard leads (Standard 1st/Target 2nd) and target leads (Target 1st/Standard 2nd). A repeated-measures ANOVA with peaks and hemisphere as the within-subject variables was performed. A significant interaction of order with peak RMS values was found. This interaction was due to a significantly higher ratio of Standard 1st/Target 2nd compared with

Table 2. Comparison of the ERFs of Targets of 5 Hz to 5 Hz-inv, and the Standard of 5 Hz to 13 Hz

<i>Rate</i>	<i>M50</i>		<i>M100</i>		<i>M200</i>		<i>M300</i>	
	<i>Right</i>	<i>Left</i>	<i>Right</i>	<i>Left</i>	<i>Right</i>	<i>Left</i>	<i>Right</i>	<i>Left</i>
<i>A. Difference between ERFs of pooled standards, targets 5 Hz and targets 5 Hz-inv</i>								
Std	1.55 ± 0.6*	1.62 ± 0.5	2.18 ± 0.9	1.65 ± 0.7	1.95 ± 0.7	1.71 ± 0.5	1.69 ± 0.6	1.54 ± 0.5*
5	1.83 ± 0.6	1.75 ± 0.4	2.37 ± 1.0	1.84 ± 0.9	1.87 ± 0.6	1.58 ± 0.5	1.53 ± 0.5	1.28 ± 0.5
5 inv	2.11 ± 0.7	1.64 ± 0.6	2.86 ± 1.4	2.03 ± 1.0	2.28 ± 0.7	1.76 ± 0.7	1.46 ± 0.5	1.25 ± 0.3
<i>F</i>	4.972*	0.463	2.267	1.190	1.864	0.625	1.109	3.040*
<i>B. Difference between ERFs of standard 5 Hz and standard 13 Hz</i>								
5	1.50 ± 0.5	1.43 ± 0.4	2.04 ± 0.8	1.61 ± 0.7	1.72 ± 0.7	1.58 ± 0.6	1.83 ± 0.7	1.65 ± 0.6
13	1.95 ± 0.7	1.68 ± 0.5	2.52 ± 1.1	2.02 ± 0.9	2.11 ± 0.7	1.82 ± 0.5	1.49 ± 0.6	1.25 ± 0.4
<i>t</i>	-3.06*	-1.7	-3.27*	-2.79**	-2.65**	-1.63	2.31	2.72**

(A) The ERFs to the standards of the 5-Hz and 5-Hz-inv conditions were pooled and compared with the targets in the two conditions. For each subject, peak RMS values for each of the ERF components was divided by the mean RMS of the whole trial, and values were averaged across subjects. ERFs of the targets in the 5-Hz-inv condition did not differ significantly from the targets in the 5-Hz condition. The pooled standard ERFs differed from both target ERFs, wherein M50 of the STANDARD was significantly smaller than that of 5 Hz-inv, and M300 was significantly larger than both 5 Hz and 5 Hz-inv. (B) ERFs of the standard in the 13-Hz condition and 5 Hz yielded significantly higher RMS values in the 13-Hz condition in the right hemisphere at M50, M100, and M200, and in the left hemisphere in M100. M300 of the left hemisphere was higher in the 5-Hz condition.

* $p < .05$.

** $p < .01$.

the ratio Target 1st/Standard 2nd in M300 in a base rate 5 Hz [$F(1, 20) = 5.21, p = .034$] and 5 Hz-inv [$F(1, 18) = 18.95, p = .0007$], but not in 13 Hz [$F(1, 20) = 0.72, p = ns$; Figure 6]. Thus, only in the 5-Hz and 5-Hz-inv conditions, where TOE was observed, M300 in the second interval was attenuated when standard was first.

To further ascertain whether this effect followed the emergence of the TOE, a similar analysis was performed on the first block of the 5-Hz condition. A repeated-measures ANOVA with peaks and hemisphere as the within-subject variables was performed. Overall, the

components in the first interval were higher, but this was stimulus type-dependent. Thus, no TOE effect on any of the components were found in the first block [$F(3, 20) = 0.1, p = ns$, data not shown].

If the ratio of the M300 was associated with the order effect, then the difference between target leads and standard leads would be significant only in trials where the observers were accurate. In other words, when target leads in wrong trials, the M300 ratio would be larger than 1, and equivalent to the ratios of standard leads in correct trials. To assess this, data from the last two

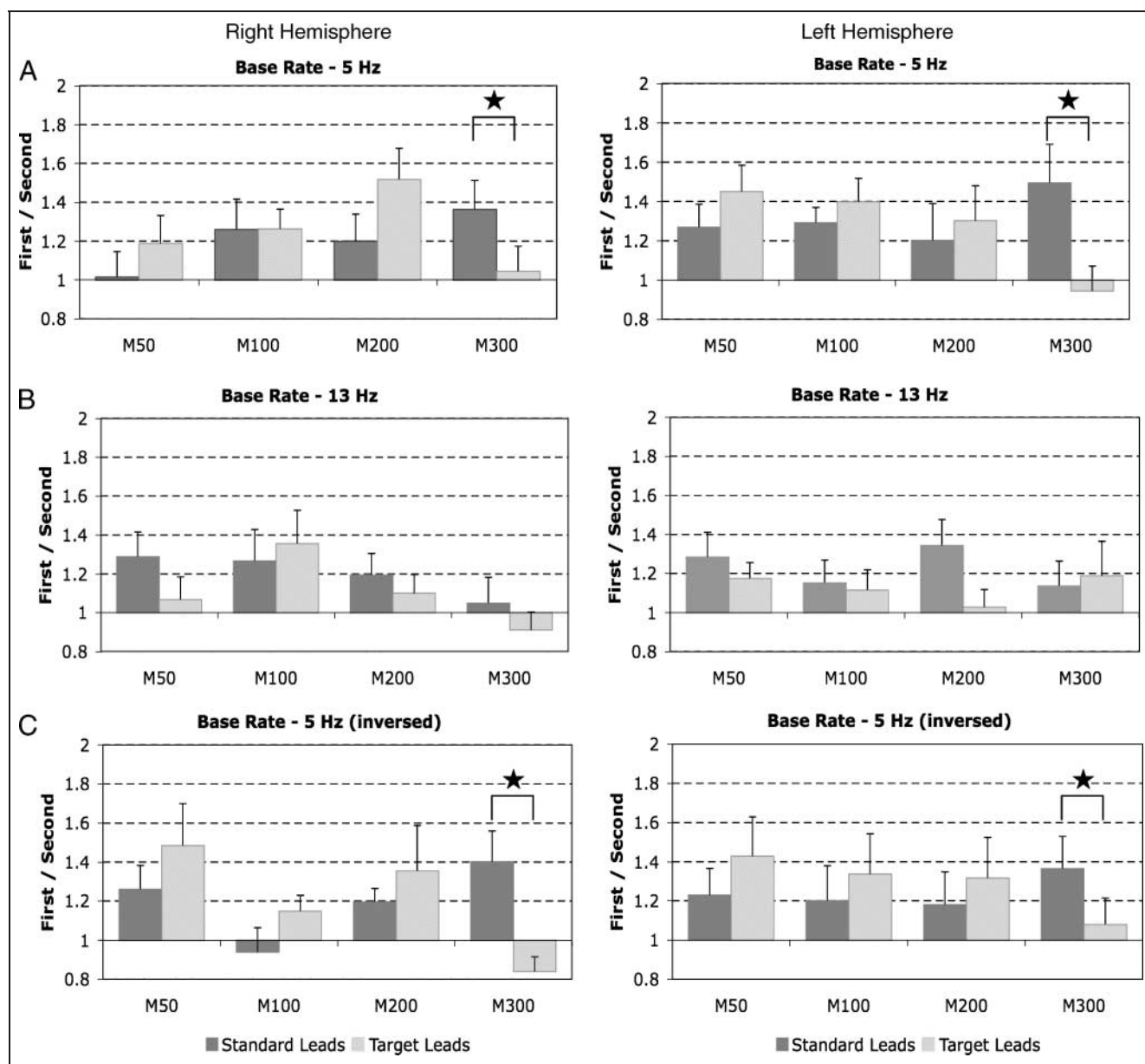


Figure 6. Ratios of the ERFs of the first interval to the ERFs to the second interval. Standard leads (filled bars)—each ERF to Standard 1st was divided by the equivalent ERF to Target 2nd. Target leads (hashed bars)—each ERF to Target 1st was divided by the equivalent ERF to Standard 2nd. Values <1 indicate that the leading stimulus had a smaller peak compared with the second; values >1 , the reverse. The ratios of ERFs for M50, M100, and M200 did not differ across condition. In the 5-Hz (A) and 5-Hz-inv (C) conditions, there was a significant difference in ratios of the M300 ERFs, wherein the M300 was greater than the M300 of the target when the standard was first, with no difference between the M300 ERFs when the target was first. “★” depict $p < .05$.

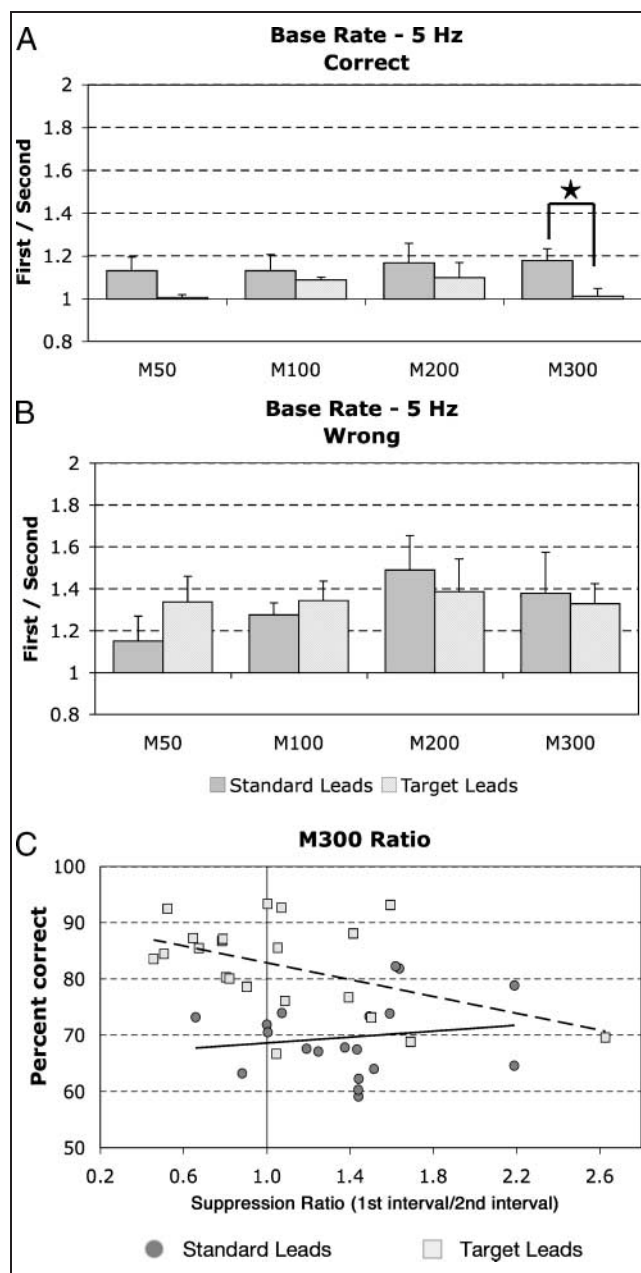


Figure 7. Ratios of the ERFs of the first interval to the ERFs to the second interval in the last two blocks of 5 Hz, separated by (A) successful (correct) and (B) unsuccessful (wrong) discrimination trials. Standard leads (filled bars)—each ERF to Standard 1st was divided by the equivalent ERF to Target 2nd. Target leads (hashed bars)—each ERF to Target 1st was divided by the equivalent ERF to Standard 2nd. Values < 1 indicate that the leading stimulus had a smaller peak compared with the second; values > 1 , the reverse. The ratios of ERFs for M50, M100, and M200 did not differ across condition. Only when observers performed successfully was there a significant difference in ratios of the M300 ERFs, wherein the M300 was greater than the M300 of the target when the standard was first, with no difference between the M300 ERFs when the target was first. “★” depict $p < .05$. (C) Percent correct on the last block was regressed against the level of suppression of the M300 component, where higher accuracy Standard 2nd trials was correlated with increased symmetry in the M300 of target (1st interval) and M300 of standard (light gray squares, hashed trend line). In Standard 1st trials, there was no correlation between performance accuracy and the ratio of the M300 component (dark gray circles, full line).

blocks of the 5-Hz condition were pooled and divided into trials in which observers performed accurately (correct) or not (wrong). Ratios were calculated as above, but hemispheres were pooled. A repeated-measures ANOVA with peak ratios and accuracy (correct/wrong) as the within-subject variables was performed. Overall, the ratios were higher for both ratio conditions, when observers were wrong [$F(1, 78) = 18.97, p < .001$], independent of order. This suggests that stronger suppression of the ERFs to the second stimulus reduced accuracy (Figure 7A and B). In addition, Scheffe’s post hoc comparison showed that the difference between target leads and standard leads maintained for the M300 peak was found only in correct trials [correct: $t(14) > t_{critical}, p = .039$; wrong: ns].

To assess this on a subject-by-subject basis, mean percent correct was regressed against the suppression measure of the M300 component. Regression analyses were run separately for Standard 1st against standard leads, and Standard 2nd against target leads. The data used were from the last block of 5 Hz and 5 Hz-inv combined. As can be seen in Figure 7C, in Standard 2nd trials, the level of suppression was negatively correlated with performance accuracy (F regression = 5.21, $p = .034$; $R^2 = .22$). No relationship was found for Standard 1st trials (F regression = 0.35, $p = ns$; $R^2 = .002$).

DISCUSSION

In this study, we demonstrated a TOE in an auditory 2AFC task with rate-modulated tone trains presented for discrimination. The TOE was manifested in a population asymmetry in the accuracy of discrimination, wherein observers performed better when the standard 5-Hz tone train was presented in the second interval. Our findings add to a long line of research in which TOE has been shown in various experimental paradigms (Allan & Gibbon, 1994), and in several sensory and perceptual modalities (Hellström, 1985). In contrast to previous findings (Allan, 1977; Jamieson & Petrusic, 1976), the TOE was present despite the fact that observers received feedback on their accuracy. It is possible, however, that observers ignored the feedback, as clearly it did not improve accuracy in the trained (5 Hz) condition.

The relative timing of the ERFs was not sensitive to the rate at which stimuli were presented, yielding a single waveform typical of auditory-evoked responses (Näätänen & Picton, 1987). This suggests that modulation rate may not be encoded temporally, but perhaps perceived as a gestalt. Accordingly, components of the ERF were easily identified, and the different components yielded dipoles localized in primary auditory regions. Notably, localization was done by applying an ECD model, and thus, may only reflect a rough estimation of the location of the more robust constituents of the network that underlie each component of the ERF. The

responses generated from the right hemisphere were stronger and occurred earlier than the responses observed from the left. The reason for this laterality is unclear, as evidence associating perception of rhythm, or beat, with right temporal regions is mixed (McAngus, Lee, & O'Boyle, 2002; Penhune, Zatorre, & Feindel, 1999; Platel et al., 1997), and assumed to occur mainly in other brain regions (reviewed in Janata & Grafton, 2003).

The TOE generalized to the 5-Hz-inv, but not to the 13-Hz condition. Thus, the asymmetry of the TOE was maintained despite the fact that, in the 5-Hz-inv condition, observers were required to respond to the targets as the correct response (i.e., they were always slower than the standard). Furthermore, the TOE was absent in the first block of the 5-Hz condition, when the stimuli were still relatively novel. Two interesting inferences can be drawn from these observations. First, as the TOE is affected by practice (see also Stott, 1935; Woodrow, 1935), its presence can be understood to reflect the emergence of an internal representation of the standard. Second, it was previously suggested that the TOE reflected a perceptual bias incurred by maintaining the first stimulus in memory during the ISI, causing the percept to “drift” toward the more common stimulus (reviewed in Allan, 1977; Woodrow, 1935), which—in the case of the 5-Hz condition—would be toward a higher rate. Thus, when the standard was presented first, it would be perceived as faster and harder to discriminate from the subsequent target. However, as the direction of the TOE was maintained in a single block in which the targets were slower, it is more likely that the TOE represents a form of interference; a conjecture more consistent with the development of an internal representation of the standard. It should be noted that the development of the TOE occurred with degradation in performance accuracy. Typically reduced accuracy is due to fatigue, however, as the decline was order-sensitive, it is unlikely that fatigue would be the sole explanation for this effect.

Furthermore, the ERFs to the 5-Hz-inv targets yielded responses that were not different from the responses to the targets in the 5-Hz condition, whereas the ERFs to the 13-Hz standard were significantly larger than the ERFs to the 5-Hz standard (Table 2B). This suggests that despite the fact that the targets in the 5 Hz-inv were novel, they were sufficiently similar to the 5-Hz standard as to elicit responses equivalent to the 5-Hz targets. This supports our interpretation that observers formed an internal representation of the 5-Hz standard, and suggests that the changes in the M300 may reflect plasticity associated with this representation.

Repetition (Lammertmann, Fijiki, Lütkenhöner, & Hari, 2000; Hari et al., 1982) and adaptation (Jääskeläinen et al., 2004; Näätänen, Jacobson, & Winkler, 2004; Näätänen & Picton, 1987) are known to affect the structure and amplitude of evoked responses to simple as well as complex stimuli, wherein higher amplitude is associated

with initial (i.e., first in a series) and/or novel stimuli. In our findings, the ERFs were reduced in the second interval up to 350 msec into the tone train, an effect that was accentuated in trials in which observers were incorrect in their discrimination. The M300 component reflected the behavioral TOE most reliably. In both 5-Hz and 5 Hz-inv conditions, the M300 to the target was significantly attenuated when presented in the second interval, but not when the target was first. In the 13-Hz condition, the amplitude of the M300 was higher in the first interval, independent of stimulus type. Moreover, although there was overall suppression of the ERFs in the second interval, the M300 was of equal magnitude, in Standard 2nd trials, when observers were accurate, but not when their discrimination was incorrect (Figure 7). It should be noted that in order to accurately discriminate the stimuli, an interval of at least 250 msec from stimulus onset is needed, in all stimulus conditions, as at least two tones of the train are necessary to determine whether it is the standard. Hence, the M300 potentially reflects the identification of the stimulus as belonging to one of the two categories, standard or target.

We propose that matching the presented stimulus to the internal representation of the standard acted as an attentional cue, resulting in an enhanced M300 response and higher accuracy. This is inline with literature on the role of the M300 (or P3) in attention allocation and performance accuracy (for review, see Soltani & Knight, 2000). For example, Shapiro et al. (2006) demonstrated that M300 amplitude was associated with the probability to detect a visual stimulus in the attentional blink paradigm, suggesting that when the target stimulus captured the observers' attention, performance was better and M300 was more robust. Thus, our findings add to the existing hypothesis as to the function of the 300-msec evoked response as reflecting attention allocation to a stimulus.

These findings may shed light on the nature of the TOE. In the past, several interpretations of TOE have been proposed. Köhler (see Neisser, 2002; Woodrow, 1935), and subsequently, Postman (1946) viewed the TOE as part of the physical properties of the sensory system, wherein prior stimulation affects perception of subsequent stimuli. Our MEG results support this hypothesis, as we show that the stimulus in the first interval generated stronger ERFs than in the second. Others have suggested that the stimuli presented for discrimination are compared with an internal representation of a *standard* which emerges during the experiment (e.g., Vos, Assen, & Franek, 1997; Woodrow, 1935; Needham, 1934). Our behavioral findings support this hypothesis as well, as the TOE occurred only with the familiar 5 Hz standard. In addition, we show that the amplitude of the M300 component was associated with correct detection of the standard, suggesting that the M300 is a neural correlate of the TOE, at least in the case of this type of stimulus.

Furthermore, although eight blocks were not sufficient for improved performance, the generalization of the TOE to novel targets in the 5-Hz-inv condition, but not to a novel standard (13 Hz), suggests that observers developed an internal representation of the standard. Thus, our findings indicate that the internal representation of the stimulus may interact with the natural tendency to have reduced processing capacity to stimuli presented serially. When a stimulus matching the internal representation is presented first, it allocates more resources, thus diminishing the ability to make an accurate decision when the comparison target appears.

Two caveats in this conclusion should be noted: First, the results obtained here are not counterbalanced to assess whether observers were repeatedly exposed to the 13-Hz condition, with the 5 Hz as a control, leaving open the possibility that our results are stimulus specific. Second, not all studies demonstrate the emergence of a TOE. For example, Jamieson and Petrusic (1975), using empty time intervals presented for discrimination, demonstrated that TOE was initially found, but subsequently eliminated, during a training session only if the observers received feedback. In subsequent work, however, these authors showed that the TOE was feedback independent (Jamieson & Petrusic, 1976, 1978). Further studies are necessary to clarify the differences between modalities, and other experimental factors, such as task and feedback conditions, that contribute to the TOE.

In sum, the TOE is a well-established phenomenon in psychophysics. Despite its robustness, it has received little attention in recent years, hence, little is known of its underlying neural processes. In this study, we have demonstrated TOE to a novel stimulus dimension (i.e., modulation rate). We found that evoked responses to the stimuli reflect the interaction of sensory and perceptual processes that may serve to encode the two stimuli presented for discrimination. Our findings further suggest that the TOE reflects plasticity in the absence of improved performance, and it remains to be determined whether this phenomenon precipitates improved performance in subsequent learning.

Acknowledgments

We thank Dr. Virginie van Wassenhove for her thoughtful comments, and Susanne Honma and Anne Findlay for their excellent technical assistance. The work was funded by R01 DC 004855-01.

Reprint requests should be sent to Ilana S. Hairston, Department of Psychology, UC Berkeley, 3210 Tolman Hall, MC 1650, Berkeley, CA 94720, or via e-mail: hairston@berkeley.edu, or to Srikantan S. Nagarajan via e-mail: sri@radiology.ucsf.edu.

REFERENCES

- Allan, L. G. (1977). The time-order error in judgments of duration. *Canadian Journal of Psychology*, *31*, 24–31.
- Allan, L. G. (1979). The perception of time. *Perception & Psychophysics*, *26*, 340–354.
- Allan, L. G., & Gibbon, J. (1994). A new temporal illusion or TOE once again? (comment). *Perception & Psychophysics*, *55*, 227–229.
- Gescheider, G. A. (1997). *Psychophysics: The fundamentals*. London: Erlbaum.
- Grondin, S. (2001). Discriminating time intervals presented in sequence marked by visual signals. *Perception & Psychophysics*, *63*, 1214–1228.
- Hari, R., Kaila, K., Katila, T., Tuomisto, T., & Varpula, T. (1982). Interstimulus interval dependence of the auditory vertex response and its magnetic counterpart: Implications for their neural generation. *Electroencephalography and Clinical Neurophysiology*, *54*, 561–569.
- Hellström, Å. (1985). The time-order error and its relatives: Mirrors of cognitive processes in comparing. *Psychological Bulletin*, *97*, 35–61.
- Hellström, Å. (2001). *Time-order effects for aesthetic preference*. Leipzig: Pabst Science Publishers.
- Jääskeläinen, I. P., Ahveninen, J., Bonmassar, G., Dale, A. M., Ilmoniemi, R. J., Levänen, S., et al. (2004). Human posterior auditory cortex gates novel sounds to consciousness. *Proceedings of the National Academy of Sciences, U.S.A.*, *101*, 6809–6814.
- Jamieson, D. G., & Petrusic, W. M. (1975). Presentation order effects in duration discrimination. *Perception & Psychophysics*, *17*, 197–202.
- Jamieson, D. G., & Petrusic, W. M. (1976). On a bias induced by the provision of feedback in psychophysical. *Acta Psychologica*, *40*, 199–206.
- Jamieson, D. G., & Petrusic, W. M. (1978). Feedback versus an illusion in time. *Perception*, *7*, 91–96.
- Janata, P., & Grafton, S. T. (2003). Swinging in the brain: Shared neural substrates for behaviors related to sequencing and music. *Nature Neuroscience*, *6*, 682–687.
- Koh, S. D. (1967). Time-error in comparisons of preferences for musical excerpts. *American Journal of Psychology*, *80*, 171–185.
- Lammertmann, C., Fijiki, N., Lütkenhöner, B., & Hari, R. (2000). *Short-term decrement of the auditory N1m response*. Paper presented at the BIOMAG, Helsinki, Finland.
- McAngus, N. T., Lee, C., & O'Boyle, D. (2002). A sensorimotor theory of temporal tracking and beat induction. *Psychological Research*, *66*, 26–29.
- Näätänen, R., Jacobson, T., & Winkler, I. (2004). Memory-based or afferent processes in mismatch negativity (MMN): A review of the evidence. *Psychophysiology*, *42*, 25–32.
- Näätänen, R., & Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound: A review and an analysis of the component structure. *Psychophysiology*, *24*, 375–425.
- Nakajima, Y., ten Hoopen, G., Sasaki, T., Yamamoto, K., Kadota, M., Simons, M., et al. (2004). Time shrinking: The process of unilateral temporal assimilation. *Perception*, *33*, 1061–1079.
- Needham, J. G. (1934). The time-error as a function of continued experimentation. *American Journal of Psychology*, *46*, 558–567.
- Needham, J. G. (1935). The effect of the interval upon the time-error at different intensive levels. *Journal of Experimental Psychology*, *18*, 539–543.
- Neisser, G. (2002). Wolfgang Köhler, 1887–1967. *Biographical Memoirs*, *81*, 1–13.
- Penhune, V. B., Zatorre, R. J., & Feindel, W. H. (1999). The role of auditory cortex in retention of rhythmic patterns as studied in patients with temporal lobe removals including Heschl's gyrus. *Neuropsychologia*, *37*, 315–331.

- Platel, H., Price, C., Baron, J. C., Wise, R., Lambert, R. S., Frackowiak, R. S., et al. (1997). The structural components of music perception. A functional anatomical study. *Brain*, *120*, 229–243.
- Postman, L. (1946). The time-error in auditory perception. *American Journal of Psychology*, *59*, 193–219.
- Sasaki, T., Suetomi, D., Nakajima, Y., & ten Hoopen, G. (2002). Time-shrinking, its propagation, and Gestalt principles. *Perception & Psychophysics*, *64*, 919–931.
- Shapiro, K., Schmitz, F., Mertens, S., Hommel, B., & Schnitzler, A. (2006). Resource sharing in the attentional blink. *NeuroReport*, *17*, 163–166.
- Soltani, M., & Knight, R. T. (2000). Neural origins of the P300. *Critical Reviews in Neurobiology*, *14*, 199–224.
- Stott, L. H. (1935). Time order errors in the discrimination of short tonal durations. *Journal of Experimental Psychology*, *18*, 741–766.
- Vos, P. G., Assen, M., & Franek, M. (1997). Perceived tempo change is dependent on base tempo and direction of change: Evidence for a generalized version of Schulze's (1978) internal beat model. *Psychological Research*, *59*, 240–247.
- Woodrow, H. (1935). The effect of practice upon time-order errors in the comparison of temporal intervals. *Psychological Review*, *42*, 127–152.