Psychophysiology, 47 (2010), 289–298. Wiley Periodicals, Inc. Printed in the USA.

DOI: 10.1111/j.1469-8986.2009.00936.x

Effects of discrepancy between imagined and perceived sounds on the N2 component of the event-related potential

JIANHUI WU,[a] XIAOQIN MAI,[b] ZULIN YU,[c] SHAOZHENG QIN,[d] AND YUE-JIA LUO[e]

[a]Institute of Psychology, Chinese Academy of Sciences, Beijing, China
[b]Center for Human Growth and Development, University of Michigan, Ann Arbor, Michigan, USA
[c]State Key Laboratory of Brain and Cognitive Science, Institute of Biophysics, Chinese Academy of Sciences, Beijing, China
[d]Centre for Cognitive Neuroimaging at Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, Nijmegen, The Netherlands
[e]State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, Beijing, China

Abstract

Two experiments were conducted to examine whether the N2 component of the event-related potential (ERP), typically elicited in a S1-S2 matching task and considered to reflect mismatch process, can still be elicited when the S1 was imagined instead of perceived and to investigate how N2 amplitude varied with the degree of S1-S2 discrepancy. Three levels of discrepancy were defined by the degree of separation between the heard (S2) and imagined (S1) sounds. It was found that the N2 was reliably elicited when the perceived S2 differed from the imagined S1, but whether N2 amplitude increased with the degree of discrepancy depended in part on the S1-S2 discriminability (as evidenced by reaction time). Specifically, the effect of increasing discrepancy was attenuated as discriminability increased from hard to easy. These results, together with the dynamic ERP topography observed within the N2 window, suggest that the N2 effect reflects two sequential but overlapping processes: automatic mismatch and controlled detection.

Descriptors: Discrepancy, Automatic mismatch, Controlled detection, Auditory imagery, N2, Event-related potentials

In a S1-S2 matching task, when the second stimulus (S2) is different from the first stimulus (S1), a negative event-related potential (ERP) component, N2, can be elicited approximately 250 ms after the onset of S2. This N2 component has generally been considered as an index of the mismatch or conflict processing (Cui, Wang, Wang, Tian, & Kong, 2000; Mao & Wang, 2007; Wang, Cui, Wang, Tian, & Zhang, 2004; Wang, Wang, Kong, Cui, & Tian, 2001; Wang et al., 2003; Yang & Wang, 2002; Zhang, Wang, Li, Wang, & Tian 2005; Zhang et al., 2001; for a review, see also Folstein & Van Petten, 2008).

In most of the previous studies using the S1-S2 paradigm, both S1 and S2 were real presented stimuli; the N2 was elicited by a discrepancy between the physical attributes of these two sequentially presented stimuli, such as color (Wang et al., 2003, 2004), shape (Cui et al., 2000; Zhang et al., 2005), and spatial location (Mao & Wang, 2007; Yang & Wang, 2002). In addition, in Wang, Kong, Tang, Zhuang, and Li's (2000) study, the N2 was elicited by a false presented answer to the preceding mentally calculated arithmetic problem, suggesting that internally generated number information can also lead to the mismatch effect. It is of interest to determine whether the N2 can still be elicited when a physical attribute of S2 is different from that of a purely imagined S1.

According to the hypothesis of mismatch or conflict processing, it is logical to predict that the N2 amplitude would be directly proportional to the degree of discrepancy between S1 and S2, but empirical evidence is needed to support this prediction. To demonstrate such an effect would require at least two levels of discrepancy between the S1 and S2 along a single perceptual feature dimension (e.g., pitch or loudness of a sound).

There were two main objectives in the present study. First, we explored whether the N2 can be observed when S1 is mentally imagined instead of actually perceived in the S1-S2 matching task. Second, and more importantly, we investigated the relationship between the N2 amplitude and the degree of discrepancy between S1 and S2. To address these issues, we conducted this study using a modified S1-S2 paradigm that we entitled "imagined-S1 perceived-S2 paradigm." Before initiating the ERP recordings, participants were trained to associate each of three geometrical shapes with one of three pure tones that varied in a single parameter (in Experiment 1 the pitch of the tones varied

---

This work was supported by the Ministry of Education, China (PCSIRT, IRT0710), the NSF China (30930031, 30900442), and the IP, CAS (O9CX042004). We acknowledge Professor Jing-Han Wei for his valuable discussions throughout the experiments and Victoria S. Arch for her editorial assistance. We sincerely thank Professor Steven Hillyard and Professor Erich Schröger and two anonymous reviewers for their helpful suggestions.

Address reprint requests to: Yue-jia Luo, State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, 19 Xin Jie KouWai Street, Beijing, 100875, China. E-mail: luoyj@bnu.edu.cn

but the loudness was constant; in Experiment 2 the loudness varied but pitch was constant). During the recording session, participants were presented with one of these three geometrical shapes and asked to imagine hearing the corresponding sound (imagined-S1); after a short delay one of these three sounds was presented (perceived-S2), and participants were required to make a same–different judgment between the perceived sound and the previously imagined one. This comparison led to three different levels of discrepancy defined by degree of separation between the heard and imagined sound: no, small, and large discrepancies. The effective use of imagery was ensured by the imagery training before the recordings and evaluated by means of questionnaire afterward. It was predicted that both the small and large discrepancy conditions would elicit an N2 component, and the large discrepancy would elicit a higher amplitude of N2 than the small discrepancy condition.

Methods

Participants

Data from 22 participants (mean age 21.6 ± 0.9 years, 10 men, all right-handed) in Experiment 1 and 23 participants (mean age 21.3 ± 1.3 years, 11 men, all right-handed) in Experiment 2 were used, after excluding 2 participants in Experiment 1 and 1 participant in Experiment 2 due to excessive movement artifacts. All were undergraduates from China Agricultural University and Beijing Forestry University who gave informed consent and were paid for their participation. None of them had a history of neurological or psychiatric disorders. All reported normal hearing and normal or corrected-to-normal vision.

Stimuli

Three geometrical shapes (square, circle, and triangle) were chosen as visual cues to induce auditory imagery. Pictures (cues) were presented on a computer monitor placed a distance of 75 cm from the participants' eyes and subtended a visual angle of approximately 2° horizontally and vertically. In Experiment 1, three pure-tone bursts of different pitch (400 Hz for low, 1000 Hz for medium, and 2500 Hz for high-pitch sound) at 75 dB SPL were chosen as imagined auditory stimuli. In Experiment 2, tones were at a constant pitch (400 Hz) and their loudness varied as 50 dB SPL (soft), 75 dB SPL (medium), and 85 dB SPL (loud). Tone bursts (250 ms duration, 25 ms rise and fall times) were broadcasted from a loudspeaker (Fostex FE107E, Japan) positioned beside the computer monitor. The SPLs of the stimulation system were measured with a condenser microphone (Brüel and Kjaer 4135) and a sound level meter (Brüel and Kjaer 2610), accurate to ± 1 dB over 0.1–10 kHz. The relationship between visual cues and pure tones was counterbalanced across participants.

Procedure

The procedure in Experiment 1 was as follows: Participants were seated in a relaxed position on a comfortable chair in a dimly lit and electrically isolated room. There were three practice sessions before the ERP recordings began. The first was a familiarization session during which the visual cues and corresponding sounds were presented simultaneously and repeatedly (at least 50 times for each pair) until participants reported that they felt capable of associating each of the three geometrical shapes with its corresponding pure tone. The second was an imagery training session during which only the visual stimulus was presented and the participants were encouraged to imagine hearing the corresponding sound as vividly as possible; the real sound was then presented, and the participants were asked to compare and adjust their previously imagined sound to this presented sound. Finally, an imagery-comparison training session was conducted during which the participants were presented with one of the geometrical shapes and asked to imagine hearing the corresponding sound; after a short delay one of these three sounds was presented and participants were required to make a same–different comparison judgment between the real sound and the previously imagined one and to indicate their answer by pressing a button as quickly and accurately as possible (see Figure 1). A correct/incorrect feedback signal was presented following the button press to encourage both response speed and accuracy.

After these three practice sessions, the participants performed 10 blocks of the ERP recording experiment with short breaks between blocks. Each block began with a short familiarization and imagery training session again. Subsequently, the imagery-comparison task was performed. Unlike the training session, however, no response feedback was given during the experiment

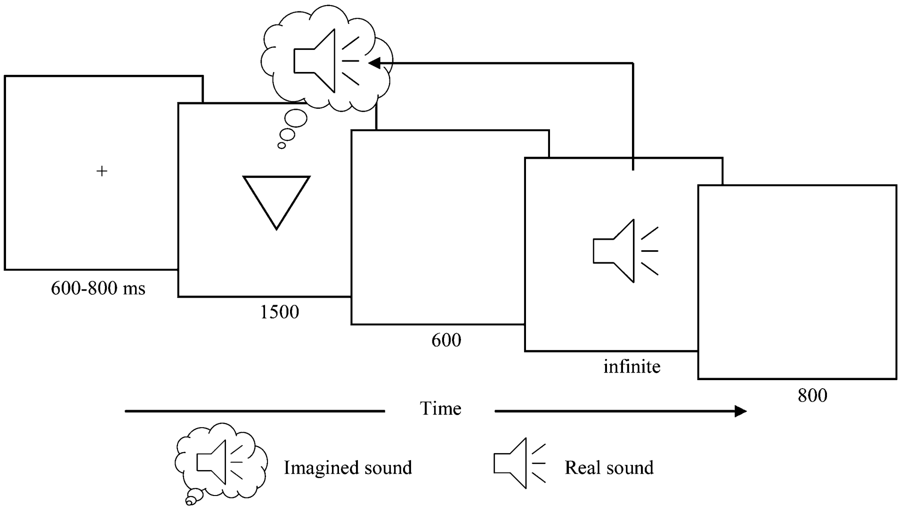

Figure 1. Schematic description of the experimental paradigm. The subjects' task was to compare the real sound with the preceding imagined sound.

(see Figure 1). The different comparison pairings between presented sounds and imagined sounds led to nine different conditions, of which only six conditions were analyzed. When the participant heard the high-pitch sound and the imagined sound was also high, then it was the no discrepancy condition; if the imagined sound was medium, then it was the small discrepancy condition; if the imagined sound was low, then it was the large discrepancy condition. The same logic applied when the low-pitch sound was presented. When the subject heard the medium-pitch sound, however, the pitch discrepancy conditions were indistinguishable and thus were excluded from analysis. Therefore, the analyzed experimental factors and levels were discrepancy (no, small, or large) and tone pitch (low- or high-pitch sound). The sequential effects for trial-to-trial transitions were counterbalanced within each block. The trials for same and different responses were presented with equal probability. For the whole experiment, each sound was both imagined and heard 240 times, leading to a total of 720 trials performed by each participant.

After the 10 blocks of recording, subjects were given a questionnaire to rate the vividness of their imagery on a 7-point scale (1 = *no imagery at all*, 7 = *very vivid imagery*) and to report whether they had experienced subvocalization (i.e., silent movements of their lips, tongue, or larynx) while imagining sounds.

The procedure for Experiment 2 was identical to that of Experiment 1 with the exception of the auditory stimuli, and thus the analyzed experimental factors and levels were discrepancy (no, small, or large) and tone loudness (hearing soft or loud sound).

EEG Recording and Analysis

During the 10 blocks of the experiments, the electroencephalogram (EEG) was recorded from 64 scalp sites using Ag/AgCl electrodes mounted in an elastic cap (Neuroscan Inc.), with an online reference to the left mastoid and off-line algebraic rereference to the average of the left and right mastoids. The vertical (VEOG) and horizontal electrooculogram (HEOG) were recorded from two pairs of electrodes, one placed above and below the left eye and another 10 mm from the outer canthi of each eye. All interelectrode impedance was maintained at <5 kΩ. Signals were amplified with a 0.05–100 Hz bandpass filter and digitized at 500 Hz.

The EEG data were digitally filtered with a 30-Hz low-pass filter and were epoched into periods of 1000 ms (including a 200-ms prestimulus baseline) time-locked to the onset of the presented sound. Ocular artifacts were removed from the EEG signal using a regression procedure implemented in the Neuroscan software (Semlitsch, Anderer, Schuster, & Presslich, 1986). Trials with various artifacts were rejected, with a criterion of ± 70 μV. The ERPs were then averaged separately for each experimental condition.

The mean amplitude of the N2 was measured in each condition in the time window of 170–270 ms after sound onset at the following 21 sites: Fz, FCz, Cz, CPz, Pz, POz, Oz, F3, FC3, C3, CP3, P3, PO3, O1, F4, FC4, C4, CP4, P4, PO4, and O2. The repeated measures analysis of variance (ANOVA) factors included Discrepancy (no, small, and large) × Pitch (hearing low- and high-pitch sound) or Loudness (hearing soft and loud sound) × Laterality (left, midline, and right) × Anterior-posterior scalp location (F, FC, C, CP, P, PO, and O). This time window was chosen because it best captured the N2 difference between the large/small discrepancy and no discrepancy conditions and was relatively free from overlap with adjacent ERPs. To reveal the dynamic ERP topography within the N2 time window, we further illustrated and measured the mean amplitudes of large/small–no discrepancy difference waves for every 20 ms from 170 to 270 ms after the onset of the sound. Due to the consideration that the N2 effects could be attributable to some extent to effects on the overlapping P3 component, we also measured the mean amplitudes in the time window of the P3 (280–400 ms) at these same 21 sites. Behaviorally incorrect trials were not analyzed. The Greenhouse–Geisser correction was used to compensate for sphericity violations. Post hoc analyses were conducted to explore interaction effects.

Results

Experiment 1

Behavior and postexperimental questionnaire. The accuracy of comparison was significantly different among the three discrepancy conditions, $F(2,42) = 31.09$, $p < .001$. Pairwise comparisons indicated that subjects responded more accurately to large discrepancy than to both no discrepancy (98.7 ± 1.0% vs. 98.1 ± 1.4%, $p < .05$) and small discrepancy (98.7 ± 1.0% vs. 94.8 ± 3.1%, $p < .001$) and also more accurately to no discrepancy than to small discrepancy ($p < .001$).

The reaction time (RT) was also significantly different among the three discrepancy conditions, $F(2,42) = 69.44$, $p < .001$. Pairwise comparisons indicated that subjects reliably responded faster on no discrepancy trials than both small discrepancy (435 vs. 557 ms, $p < .001$) and large discrepancy trials (435 vs. 476 ms, $p < .001$) and also faster on large discrepancy than on small discrepancy trials ($p < .001$). The Discrepancy (small and large) × Pitch (low- and high-pitch sound) interaction reached marginal significance, $F(1,21) = 3.40$, $p < .1$. Post hoc tests revealed that the RT difference between large and small discrepancy was greater when hearing low-pitch sound than high-pitch sound (see Figure 3, top left corner).

The postexperimental questionnaire revealed that all subjects had experienced vivid auditory imagery (5.98 ± 0.54) when the geometrical shapes were presented. Fifteen out of the 22 subjects reported that they had experienced subvocalization when imagining sounds.

ERP results: N2 amplitude. An ANOVA revealed a significant main effect of discrepancy on N2 amplitude, $F(2,42) = 61.31$, $p < .001$, and pairwise comparisons showed that both large and small discrepancy conditions were more negative than no discrepancy ($ps < .001$), and large discrepancy showed a trend toward a more negative N2 than small discrepancy ($p < .1$). The interaction between discrepancy (no, small, and large) and pitch (low- and high-pitch sound) was significant, $F(2,42) = 3.90$, $p < .05$. Post hoc tests revealed that N2 amplitude increased with the degree of discrepancy when hearing high-pitch sound (large discrepancy $>$ small discrepancy, $p = .001$; small discrepancy $>$ no discrepancy, $p < .001$; see Figure 2, left panel). When participants heard low-pitch sound, both small and large discrepancies showed a more negative N2 than in the no discrepancy condition ($ps < .001$), but the difference between small and large discrepancies was not significant (see Figure 2, left panel; also see Figure 3, middle left).

The interaction between three factors (pitch: low- and high-pitch sound; discrepancy: no, small, and large; and anterior-

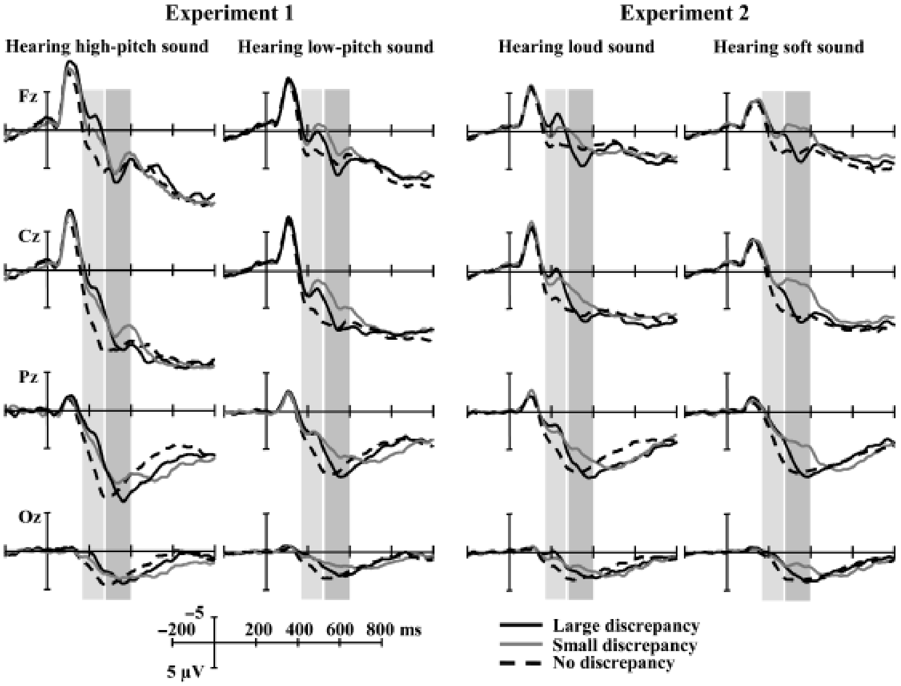

Figure 2. Grand-averaged ERPs to three levels of discrepancy when participants heard high-/low-pitch sounds in Experiment 1 (left panel) and loud/soft sounds in Experiment 2 (right panel). The data from four electrodes (Fz, Cz, Pz, and Oz) are presented, and the gray areas highlight the time windows of N2 (170–270 ms) and P3 (280–400 ms) used for statistical analysis.

posterior electrodes) was significant, $F(12,252) = 4.34$, $p < .01$, $\varepsilon = .29$. Further analysis revealed that the effects of both small (small vs. no) and large (large vs. no) discrepancies were broadly distributed along the anterior–posterior dimension, but were minimal at occipital areas. Figure 4 (upper panel) revealed the subtle change of scalp distributions over the time window of N2 for the large discrepancy effect when participants heard high-pitch sound, from fronto-central areas to centro-parietal areas; this trend, however, disappeared for the discrepancy effects under other conditions. Accordingly, when participants heard the high-pitch sound, the N2 effect for the large discrepancy at the earlier time intervals (170–190 ms) was maximally located at fronto-central areas, whereas at the later time intervals (250–270 ms), the maximal discrepancy effect shifted to the centro-parietal areas (see Figure 5, left).

ERP results: P3 amplitude. The main effect of discrepancy on P3 amplitude was significant, $F(2,42) = 17.20$, $p < .001$, and pairwise comparisons showed that both no and large discrepancies evoked more positive P3 amplitudes than small discrepancy (ps $< .001$), but the difference between no and large discrepancies was not significant (see Figure 2, left panel). The interaction between discrepancy (small and large) and pitch (low- and high-pitch sound) was also significant, $F(1,21) = 16.55$, $p = .001$. Inspection of Figure 3 (bottom left corner) suggests that the interaction occurred because the difference between small and large discrepancies was greater when hearing low-pitch sound than high-pitch sound.

The interaction between discrepancy (large and no discrepancies) and anterior-posterior electrodes was not significant when participants heard both low- and high-pitch sounds, whereas the interaction between discrepancy (small and no discrepancies) and anterior–posterior electrodes was significant when participants heard low-pitch sound, $F(6,126) = 8.77$, $p = .001$, $\varepsilon = .33$ and high-pitch sound, $F(6,126) = 3.89$, $p < .05$, $\varepsilon = .31$. Post hoc tests revealed that when participants heard low-pitch sound, this small discrepancy effect was broadly distributed along the anterior–posterior dimension, but was relatively small at frontal, parieto-occipital, and occipital areas and when participants heard high-pitch sound, this small discrepancy effect was distributed from fronto-central to parietal scalp areas, but not at frontal, parieto-occipital, and occipital areas.

Experiment 2

Behavior and postexperimental questionnaire. The results from Experiment 2 were generally consistent with those from Experiment 1. The accuracy of comparison was significantly different among the three discrepancy conditions, $F(2,44) = 11.49$, $p = .001$. Pairwise comparisons showed that accuracy in the large discrepancy condition was significantly greater than in both small discrepancy (98.9 ± 1.2% vs. 96.1 ± 3.0%, $p = .001$) and no discrepancy conditions (98.9 ± 1.2% vs. 97.0 ± 1.6%, $p < .001$). The difference between no and small discrepancies, however, did not achieve significance in Experiment 2.

RT was also significantly different among the three discrepancy conditions, $F(2,44) = 125.97$, $p < .001$. Pairwise comparisons indicated that subjects responded faster on no discrepancy trials than on both small discrepancy (507 vs. 609 ms, $p < .001$) and large discrepancy trials (507 vs. 532 ms, $p < .001$) and also faster on large discrepancy than on small discrepancy trials ($p < .001$). The Discrepancy (small and large) × Loudness (soft and loud sound) interaction was significant, $F(1,22) = 28.15$, $p < .001$. Post hoc tests revealed that the RT difference between large and small discrepancies was greater when participants heard soft sound than loud sound (see Figure 3, top right corner).

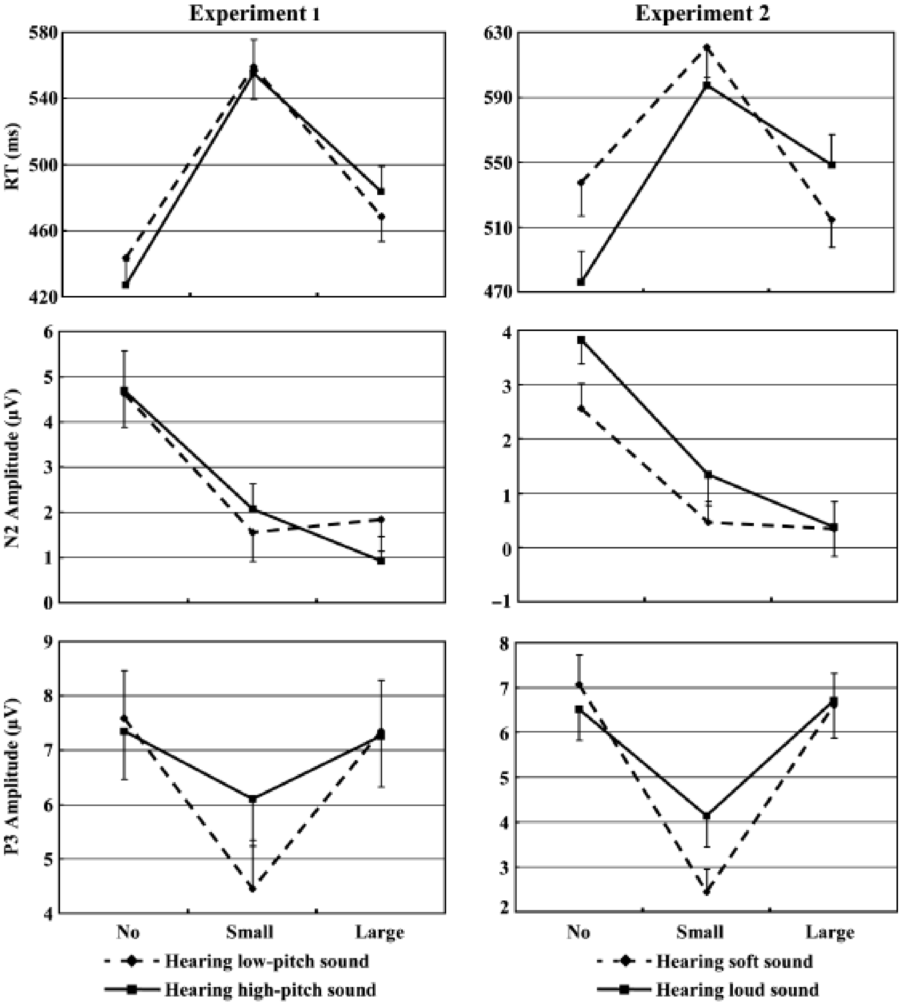

Figure 3. Mean reaction time and standard error in milliseconds (top), mean N2 amplitude and standard error across the 21 analyzed electrodes (middle), and mean P3 amplitude and standard error across the 21 analyzed electrodes (bottom) for three levels of discrepancy (No = no discrepancy, Small = small discrepancy, Large = large discrepancy). When participants heard low-pitch sounds in Experiment 1 and soft sounds in Experiment 2, the N2 amplitude from the small to the large discrepancy do not show significant change, whereas RT decreased and P3 amplitude increased to a greater extent than when participants heard high-pitch sound in Experiment 1 and loud sound in Experiment 2, respectively.

The postexperimental questionnaire revealed that all subjects had experienced vivid auditory imagery (5.94 ± 0.51) when the geometrical shapes were presented. Seventeen out of the 23 subjects reported that they had experienced subvocalization when imagining sounds.

ERP results: N2 amplitude. The results of Experiment 2 replicated those of Experiment 1. An ANOVA revealed a significant main effect of discrepancy on N2 amplitude, $F(2,44) = 71.45$, $p < .001$, and pairwise comparisons showed that both large and small discrepancy conditions were more negative than no discrepancy ($ps < .001$), and large discrepancy elicited a more negative N2 than small discrepancy ($p < .05$). The interaction between discrepancy (no, small, and large) and loudness (hearing soft and loud sounds) was significant, $F(2,44) = 4.32$, $p < .05$. Post hoc tests revealed that N2 amplitude increased with the degree of discrepancy when participants heard the loud sound (large discrepancy > small discrepancy, $p < .01$; small discrepancy > no discrepancy, $p < .001$; see Figure 2, right panel). When participants heard the soft sound, both small and large discrepancies showed a more negative N2 than in the no discrepancy condition ($ps < .001$), but the difference between small and large discrepancies was not significant (see Figure 2, right panel; also see Figure 3, middle right).

The interaction between three factors (loudness: hearing soft and loud sounds; discrepancy: no, small and large discrepancies; and anterior-posterior electrodes) was marginally significant, $F(12,264) = 2.41$, $p < .1$, $\varepsilon = .30$. Further analysis revealed that the effects of both small (small vs. no) and large (large vs. no) discrepancies were broadly distributed along the anterior–posterior dimension, but were minimal at occipital areas. Figure 4 (lower panel) further revealed the dynamic topography within the N2 time window for the large discrepancy effects when participants heard loud sound (i.e., from fronto-central to centro-parietal areas); this trend, however, disappeared for the discrepancy effects under other conditions. Accordingly, when participants heard the loud sound the N2 effect for the large discrepancy at the earlier time intervals (170–190 ms) was maximally located at fronto-central areas, whereas at the later time intervals (250–270 ms) the maximal discrepancy effect shifted to the centro-parietal areas (see Figure 5, right).

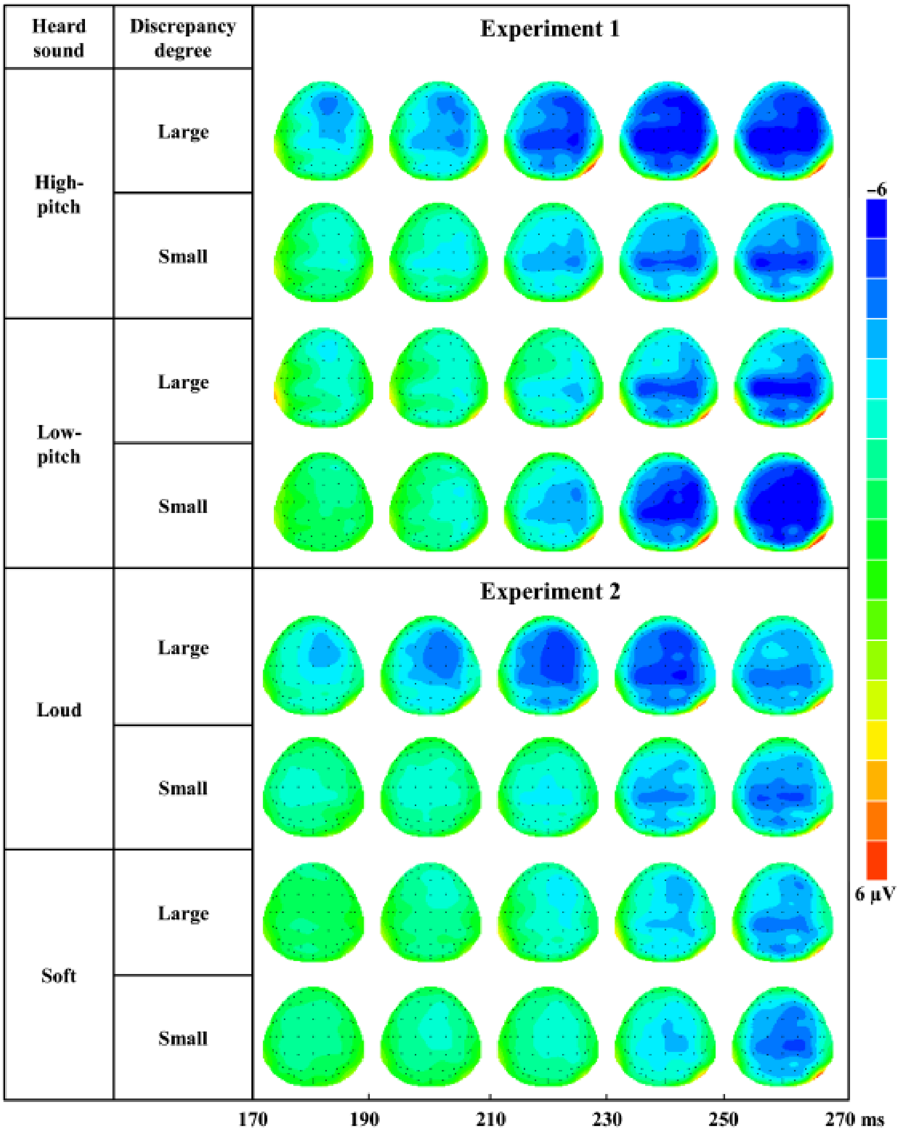

Figure 4. Sequential topographic voltage maps (generated every 20 ms from 170 ms to 270 ms) for large/small–no discrepancy difference waves when participants heard sounds with different pitch (Experiment 1, upper panel) and loudness (Experiment 2, lower panel).

ERP results: P3 amplitude. The general trend of discrepancy effect on the amplitude of P3 and its interaction with sound feature in Experiment 2 followed the same pattern as for those in Experiment 1. The main effect of discrepancy on P3 amplitude was significant, $F(2,44) = 53.87$, $p < .001$, and pairwise comparisons showed that both no and large discrepancies evoked more positive P3 amplitude than small discrepancy ($ps < .001$), but the difference between large and no discrepancies was not significant (see Figure 2, right panel). The interaction between discrepancy (small and large) and loudness (hearing soft and loud sounds) was also significant, $F(1,22) = 10.40$, $p < .01$. Inspection of Figure 3 (bottom right corner) suggests that the interaction occurred because the difference between small and large discrepancies was greater when participants heard soft sound than loud sound.

The scalp distribution of the small and large discrepancy effect when participants heard soft and loud sounds in Experiment 2 was generally consistent with those obtained in Experiment 1, except for the large discrepancy effect when participants heard loud sound. The interaction between three factors (loudness: hearing soft and loud sounds; discrepancy: no, small, and large; and anterior–posterior electrodes) was significant, $F(12,264) = 5.00$, $p < .01$, $\varepsilon = .30$. Post hoc tests revealed that for hearing both soft and loud sounds, the small discrepancy effect (small vs. no discrepancy) was broadly distributed along the anterior–posterior dimension but was relatively small at frontal and occipital areas. The large discrepancy effect, however, showed a quite different pattern for hearing soft and loud sounds. When hearing loud sound, large discrepancy showed a significantly larger amplitude of P3 than no discrepancy at frontal scalp areas, $F(1,22) = 6.17$, $p < .05$); when hearing soft sound, however, this large discrepancy effect did not occur at any scalp areas.

Discussion

The present study investigated the effects of discrepancies between auditory imagery and auditory perception on the N2

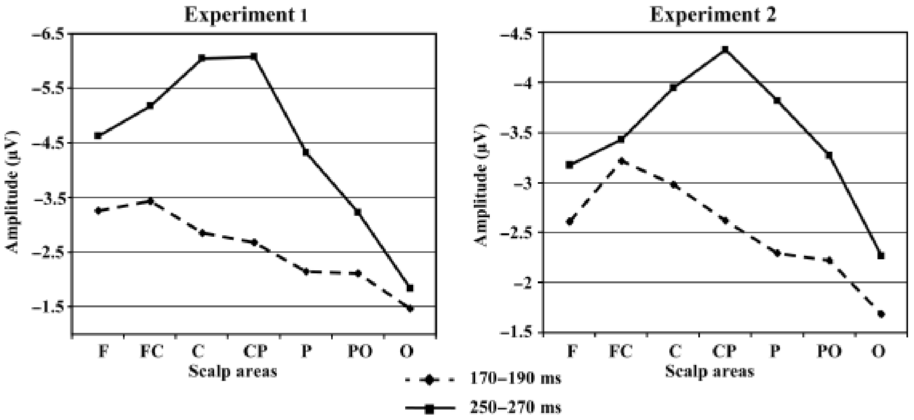

Figure 5. Mean amplitudes of large–no discrepancy difference waves measured at seven scalp areas during the 170–190 ms and 250–270 ms time windows after the onset of high-pitch sound in Experiment 1 and loud sound in Experiment 2.

component of the ERP. The results of the imagery vividness questionnaire suggest that the subjects successfully executed the auditory imagery task in both experiments. Most of the subjects reported that they had experienced subvocalization while imagining sounds, which is consistent with our previous study (Wu, Mai, Chan, Zheng, & Luo, 2006) and other studies showing the role of subvocalization in auditory imagery (e.g., Smith, Reisberg, & Wilson, 1992; Smith, Wilson, & Reisberg, 1995), providing further evidence for the execution of the auditory imagery task as required in the present study.

The results of our two experiments showed that the discrepancy in pitch and loudness between the perceived and imagined tones produced a very similar pattern of behavioral data and ERPs. The main findings of the present study can be summarized as follows. First, an N2 was reliably elicited when the perceived S2 was different from the imagined S1. Second, the amplitude of the N2 increased with the degree of discrepancy when hearing high-pitch sound in Experiment 1 and loud sound in Experiment 2, but this pattern did not appear when the subjects heard low-pitch sound in Experiment 1 and soft sound in Experiment 2.

The analysis of the main effect of discrepancy on behavioral performance in both experiments revealed that RTs were shortest for no discrepancy conditions, intermediate for large discrepancy, and longest for small discrepancy, and accuracies were higher for the large discrepancy than small discrepancy conditions. The long RTs and low accuracies to the small discrepancy condition indicate that it is more difficult to discriminate the perceived sound from the previously imagined one. This result is consistent with previous findings that used the oddball paradigm, in which RTs are reduced or accuracies are increased as the degree of mismatch increases (e.g., Novitski, Tervaniemi, Huotilainen, & Näätänen, 2004; Pakarinen, Takegata, Rinne, Huotilainen, & Näätänen, 2007; Tiitinen, May, Reinikainen, & Näätänen, 1994).

The ERP results in both experiments showed that large and small discrepancy conditions elicited an N2 when compared to the no discrepancy condition, and this N2 effect (across the whole N2 time window) had a similarly broad scalp distribution for all discrepancy conditions. The N2 has been generally observed in response to visual discrepancy in previous studies that utilized the S1-S2 matching paradigm and considered as reflecting mismatch or conflict processing (Cui et al., 2000; Mao & Wang, 2007; Wang et al., 2001, 2003, 2004; Yang & Wang, 2002; Zhang et al., 2001, 2005). The present study demonstrates this effect in the auditory modality, supporting the argument proposed by Wang, Wang, Cui, Tian, and Zhang (2002) that the N2 reflects a supramodal conflict process.

More significantly, in our experiments the elicitation of the N2 was achieved by the discrepancy between perceived sound and imagined sound, providing new insight into the concept of mismatch processing. In the S1-S2 matching task, the information from S1 is maintained in the working memory of participants and then is compared with S2; the detected discrepancy generates the mismatch signal, which leads to the scalp-recorded N2 component. In the previous studies applying the S1-S2 paradigm, the S1 was presented immediately prior to the presentation of the S2. Näätänen (1986) has proposed that it is "expectancy mismatch" that determines the N2, rather than "physical stimulus deviation from the preceding stimulus." The results from studies by Gehring, Gratton, Coles, and Donchin (1992) and Breton, Ritter, Simson, and Vaughan (1988) also support this proposal. In Widmann, Kujala, Tervaniemi, Kujala, and Schröger's (2004) study with a "Symbol-to-Sound Matching" paradigm, an early negativity was elicited by a violation of visually induced auditory expectation. The authors suggested that this effect was induced by the discrepancy between the representation of the current sound and "visually induced representation of the expected forthcoming sound." Sams, Alho, and Näätänen (1983) reported that a task-irrelevant stimulus elicited a larger N2 when it was preceded by a longer sequence of task-relevant stimuli, suggesting that the N2 reflects "mismatching with the mental image of the target stimulus voluntarily held by the subject." The results from our study provide direct evidence that a top-down originated S1 (i.e., the imagined S1 in the present study) can also form a memory trace for comparison to the following S2. Therefore, it seems quite reasonable to suggest that a mismatch with the information in working memory, no matter from where the information originated, is a major determinant of the N2.

We also examined the relationship between N2 amplitude and the degree of discrepancy. The results differed according to whether the high-pitch/loud sound or the low-pitch/soft sound was presented. When hearing the high-pitch sound in Experiment 1 and the loud sound in Experiment 2, the large discrepancy elicited a more negative N2 than the small discrepancy. This result supports our prediction that N2 amplitude is directly proportional to the degree of discrepancy and provides further evidence in favor of the mismatch interpretation of the N2 elicited in the S1-S2 paradigm. However, when participants heard the low-pitch sound in Experiment 1 and the soft sound in Experiment 2, the differences in N2 amplitude between small and large discrepancies were not significant. One possible interpretation for this incongruence would be that N2 amplitude is affected by other factors in addition to mismatch. The RTs showed a greater reduction going from the small to large discrepancy condition when participants heard low-pitch/soft sounds than high-pitch/loud sounds, suggesting that the differences in discriminability are greater between the small and large discrepancies for low-pitch/soft sounds than for high-pitch/loud sounds. We thus propose that the greater differences in discriminability reflected by RTs might contribute to the lack of difference in N2 amplitude between small and large discrepancies when the subjects heard low-pitch/soft sounds.

Here, it is worthwhile to mention the ERP results from the active oddball paradigm. In contrast with the passive or classic oddball paradigm, the active oddball paradigm requires subjects to voluntarily detect the infrequent stimuli. There are two sequential but highly overlapping negativities elicited by the task-relevant deviant stimuli (or task-irrelevant but highly deviant stimuli with respect to the ongoing standard; e.g., Näätänen, Paavilainen, Titinen, Jiang, & Alho, 1993; Näätänen, Simpson, & Loveless, 1982; Novak, Ritter, Vaughan, & Wiznitzer, 1990; Sams, Paavilainen, Alho, & Näätänen, 1985). The first is the mismatch negativity (MMN), which is typically elicited by unattended deviant auditory stimuli compared with standard sounds in the passive oddball paradigm and has been proposed to reflect an automatic neural mismatch process (for a review, see Näätänen, 1990). The second is the N2b, which has been considered to be "a sign of detection of stimulus deviance" based upon the previous mismatching process (Näätänen et al., 1982; Sams et al., 1985). Schröger (1997) also proposed that "the conscious perception of infrequent deviant sounds . . . may in part be based on the output of an obligatorily operating deviance detection system." Many studies have shown that the amplitude of the MMN is directly proportional to the degree of deviation (Berti, Roeber, & Schröger, 2004; Näätänen, Paavilainen, Alho, Reinikainen, & Sams, 1989; Novitski et al., 2004; Pakarinen et al., 2007; Tiitinen et al., 1994). The amplitude of the N2b, however, shows an opposite trend: Its amplitude is greater the more difficult it is to discriminate between the standard and the deviants (as, for example, when there is less deviation) in the active oddball paradigm (Fitzgerald & Picton, 1983) and between two simultaneously presented visual stimuli (Senkowski & Herrmann, 2002).

The stimuli are usually task relevant in the S1-S2 paradigm because a same–different comparison task is required. Thus, similar to the finding in the active oddball paradigm, there might also be two immediately sequential and highly overlapping cognitive processes that oppositely affect the amplitude of the N2 observed in the S1-S2 paradigm. The first is automatic or preattentive mismatch processing as indexed by the early part of the N2 (N2a), with its amplitude directly proportional to the degree of mismatch; the second is controlled or conscious detection processing as indexed by the late part of the N2 (N2b), the amplitude of which reflects the amount of effort allocated to the detection of the stimulus that the initial mismatch process determines as possible discrepancy; that is, the N2b amplitude is inversely related to the discriminability (proportional to the degree of mismatch between the two stimuli). Therefore, the actual N2 amplitude might depend on the balance between these two processes: The large discrepancies would produce higher amplitude of the N2 than small discrepancies when the mismatch effect dominated, and this effect of discrepancy degree would be attenuated when it is modulated by the conscious detection effect to a greater extent, as reflected by greater differences in discriminability when participants heard low-pitch and soft sounds in the present study.

In the literature related to the active oddball paradigm, the MMN reportedly has a fronto-central scalp distribution. The separation of MMN and N2b was based on the longer latency of N2b and its relatively posterior scalp distribution (Näätänen & Gaillard, 1983; Novak et al., 1990; Sams et al., 1985). Along these lines, the dynamic scalp topographies observed during the N2 time window in the present study provide spatiotemporal evidence for such a two-stage composition of the N2: The earlier phase of the N2 consisted of a fronto-centrally distributed automatic mismatch process, whereas the later phase represented the centro-parietally distributed conscious detection process. This trend disappeared in the small discrepancy conditions and in the large discrepancy conditions when participants heard low-pitch or soft sounds, however, which might be explained by the greater reciprocal influences between the automatic and controlled-related subcomponents relative to the large discrepancy conditions when participants heard high-pitch or loud sounds.

In the literature, the N2 elicited in the S1-S2 matching task has been generally considered as a mismatch process. Zhang, Wang, Li, and Wang (2003) further hypothesized that this component was related to endogenous processing, that is, the identification of discrepancy information. One of these results by Wang and his colleagues (2001), however, revealed that the N2 was elicited by task-relevant as well as task-irrelevant S1-S2 discrepancies, with the N2 amplitude being larger for the task-relevant discrepancy. The authors concluded that the discrepancy effect on the brain can "be initiated independently of the task, but is enhanced in task relevant conflict." Both the dynamic topography within the N2 window and the relationship between the N2 amplitude and reaction time in the present study support the two-stage hypothesis of stimulus discrimination between S1 and S2, that is, once a stimulus has been found to mismatch a template (imagined S1 here), effort is allocated and a conscious detection process is initiated to confirm or disconfirm its classification as a discrepancy. The output of the second stage leads to a more endogenous stage of information processing, namely, decision making, as reflected by the P3 component. The P3 amplitude has been generally considered to reflect decision confidence and task difficulty; more highly confident decisions (e.g., Hillyard, Squires, Bauer, & Lindsay, 1971; Sommer & Matt, 1990) or less difficult tasks (e.g., Ford, Roth, & Kopell, 1976; Palmer, Nasman, & Wilson, 1994; Polich, 1987) are associated with an increase of P3 amplitude or/and a shortening of its latency. Similar effects were obtained in both experiments of the present study, that is, both large and no discrepancies evoked more positive P3 amplitudes than small discrepancies.

One concern is that the N2 might be related not to an increased negativity on the discrepancy trials, but to an increased overlapping P3 on the no discrepancy trials, and this concern is heightened by the similarly broad distribution between the N2 effect (more negative N2 amplitudes for large/small discrepancies than no discrepancy) and the P3 effect (less positive P3 amplitudes for small discrepancy than no discrepancy). A similar concern exists for the comparison between large and small discrepancies. But the discrepancy-related modulations of the N2 cannot be accounted for solely by its overlap with the P3 component. The large discrepancy still elicited more negative N2 amplitudes than no discrepancy, although the P3 amplitude difference between large and no discrepancy is not significant when participants heard low- and high-pitch sounds in Experiment 1 and soft sounds in Experiment 2 or an even more positive P3 amplitude is elicited by a large discrepancy than by no discrepancy when participants heard loud sounds in Experiment 2. It is the same logic for the comparison between large and small discrepancies: a large discrepancy not only elicited more positive P3 amplitudes but also more negative N2 amplitudes than a small discrepancy when participants heard high-pitch or loud sounds.

In summary, the N2 can be reliably elicited when the perceived S2 is discrepant from the previously imagined S1. In addition, our results suggest that the N2 elicited in the S1-S2 paradigm might reflect two sequential but overlapping cognitive processes: the fronto-centrally distributed N2a subcomponent reflecting automatic mismatch processing and the centro-parietally distributed N2b subcomponent associated with conscious detection processing. The mismatch degree and discriminability are concurrently affected by the degree of discrepancy between S1 and S2, with a greater discrepancy producing both a stronger mismatch and an easier discriminability between S1 and S2. The N2 amplitude is augmented with increasing mismatch but concurrently reduced for easier discriminability between these two stimuli.

REFERENCES

Berti, S., Roeber, U., & Schröger, E. (2004). Bottom-up influences on working memory: Behavioral and electrophysiological distraction varies with distractor strength. *Experimental Psychology*, *51*, 249–257.

Breton, F., Ritter, W., Simson, R., & Vaughan, H. G. (1988). The N2 component elicited by stimulus matches and multiple targets. *Biological Psychology*, *27*, 23–44.

Cui, L., Wang, Y., Wang, H., Tian, S., & Kong, J. (2000). Human brain sub-systems for discrimination of visual shapes. *NeuroReport*, *11*, 2415–2418.

Fitzgerald, P. G., & Picton, T. W. (1983). Event-related potentials recorded during the discrimination of improbable stimuli. *Biological Psychology*, *17*, 241–276.

Folstein, J. R., & Van Petten, C. (2008). Influence of cognitive control and mismatch on the N2 component of the ERP: A review. *Psychophysiology*, *45*, 152–170.

Ford, J. M., Roth, W. T., & Kopell, B. S. (1976). Auditory evoked potentials to unpredictable shifts in pitch. *Psychophysiology*, *13*, 32–39.

Gehring, W. J., Gratton, G., Coles, M. G., & Donchin, E. (1992). Probability effects on stimulus evaluation and response processes. *Journal of Experimental Psychology: Human Perception and Performance*, *18*, 198–216.

Hillyard, S. A., Squires, K. C., Bauer, J. W., & Lindsay, P. H. (1971). Evoked potential correlates of auditory signal detection. *Science*, *172*, 1357–1360.

Mao, W., & Wang, Y. (2007). Various conflicts from ventral and dorsal streams are sequentially processed in a common system. *Experimental Brain Research*, *177*, 113–121.

Näätänen, R. (1986). A classification of N2 kinds of ERP components. In W. C. McCallum, R. Zappoli, & F. Denoth (Eds.), *Cerebral psychophysiology: Studies in event-related potentials* (*EEG Suppl. 38*, pp. 169–172). Amsterdam: Elsevier.

Näätänen, R. (1990). The role of attention in auditory information processing as revealed by event-related potentials and other brain measures of cognitive Function. *Behavioral and Brain Sciences*, *13*, 201–288.

Näätänen, R., & Gaillard, A. W. K. (1983). The orienting reflex and the N2 deflection of the ERP. In A. W. K. Gaillard & W. Ritter (Eds.), *Tutorials in event-related potential research: Endogenous components* (pp. 119–141). Amsterdam: North-Holland.

Näätänen, R., Paavilainen, P., Alho, K., Reinikainen, K., & Sams, M. (1989). Do event-related potentials reveal the mechanism of the auditory sensory memory in the human brain? *Neuroscience Letters*, *98*, 217–221.

Näätänen, R., Paavilainen, P., Tiitinen, H., Jiang, D., & Alho, K. (1993). Attention and mismatch negativity. *Psychophysiology*, *30*, 436–450.

Näätänen, R., Simpson, M., & Loveless, N. E. (1982). Stimulus deviance and evoked potentials. *Biological Psychology*, *14*, 53–98.

Novak, G. P., Ritter, W., Vaughan, H. G. J, & Wiznitzer, M. L. (1990). Differentiation of negative event-related potentials in an auditory discrimination task. *Electroencephalography and Clinical Neurophysiology*, *75*, 255–275.

Novitski, N., Tervaniemi, M., Huotilainen, M., & Näätänen, R. (2004). Frequency discrimination at different frequency levels as indexed by electrophysiological and behavioral measures. *Cognitive Brain Research*, *20*, 26–36.

Pakarinen, S., Takegata, R., Rinne, T., Huotilainen, M., & Näätänen, R. (2007). Measurement of extensive auditory discrimination profiles using the mismatch negativity (MMN) of the auditory event-related potential (ERP). *Clinical Neurophysiology*, *118*, 177–185.

Palmer, B., Nasman, V. T., & Wilson, G. F. (1994). Task decision difficulty: Effects on ERPs in a same–different letter classification task. *Biological Psychology*, *38*, 199–214.

Polich, J. (1987). Task difficulty, probability, and inter-stimulus interval as determinants of P300 from auditory stimuli. *Electroencephalography and Clinical Neurophysiology*, *68*, 311–320.

Sams, M., Alho, K., & Näätänen, R. (1983). Sequential effects on the ERP in discriminating two stimuli. *Biological Psychology*, *17*, 41–58.

Sams, M., Paavilainen, P., Alho, K., & Näätänen, R. (1985). Auditory frequency discrimination and event-related potentials. *Electroencephalography and Clinical Neurophysiology*, *62*, 437–448.

Schröger, E. (1997). On the detection of auditory deviations: A pre-attentive activation model. *Psychophysiology*, *34*, 245–257.

Semlitsch, H. V., Anderer, P., Schuster, P., & Presslich, O. (1986). A solution for reliable and valid reduction of ocular artifacts, applied to the P300 ERP. *Psychophysiology*, *23*, 695–703.

Senkowski, D., & Herrmann, C. S. (2002). Effects of task difficulty on evoked gamma activity and ERPs in a visual discrimination task. *Clinical Neurophysiology*, *113*, 1742–1753.

Smith, J. D., Reisberg, D., & Wilson, M. (1992). Subvocalization and auditory imagery: Interactions between the inner ear and inner voice. In D. Reisberg (Ed.), *Auditory imagery* (pp. 95–119). Hillsdale, NJ: Erlbaum.

Smith, J. D., Wilson, M., & Reisberg, D. (1995). The role of subvocalization in auditory imagery. *Neuropsychologia*, *33*, 1433–1454.

Sommer, W., & Matt, J. (1990). Awareness of P300-related cognitive processes: A signal detection approach. *Psychophysiology*, *27*, 575–585.

Tiitinen, H., May, P., Reinikainen, K., & Näätänen, R. (1994). Attentive novelty detection in humans is governed by pre-attentive sensory memory. *Nature*, *372*, 90–92.

Wang, H., Wang, Y., Kong, J., Cui, L., & Tian, S. (2001). Enhancement of conflict processing activity in human brain under task relevant condition. *Neuroscience Letters*, *298*, 155–158.

Wang, Y., Cui, L., Wang, H., Tian, S., & Zhang, X. (2004). The sequential processing of visual feature conjunction mismatches in the human brain. *Psychophysiology*, *41*, 21–29.

Wang, Y., Kong, J., Tang, X., Zhuang, D., & Li, S. (2000). Event-related potential N270 is elicited by mental conflict processing in human brain. *Neuroscience Letters*, *293*, 17–20.

Wang, Y., Tian, S., Wang, H., Cui, L., Zhang, Y., & Zhang, X. (2003). Event-related potentials evoked by multi-feature conflict under different attentive conditions. *Experimental Brain Research*, *148*, 451–457.

Wang, Y., Wang, H., Cui, L., Tian, S., & Zhang, Y. (2002). The N270 component of the event-related potential reflects supramodal conflict processing in humans. *Neuroscience Letters*, *332*, 25–28.

Widmann, A., Kujala, T., Tervaniemi, M., Kujala, A., & Schröger, E. (2004). From symbols to sounds: Visual symbolic information activates sound representations. *Psychophysiology*, *41*, 709–715.

Wu, J., Mai, X., Chan, C. C., Zheng, Y., & Luo, Y. (2006). Event-related potentials during mental imagery of animal sounds. *Psychophysiology*, *43*, 592–597.

Yang, J., & Wang, Y. (2002). Event-related potentials elicited by stimulus spatial discrepancy in humans. *Neuroscience Letters*, *326*, 73–76.

Zhang, X., Wang, Y., Li, S., & Wang, L. (2003). Event-related potential N270, a negative component to identification of conflicting information following memory retrieval. *Clinical Neurophysiology*, *114*, 2461–2468.

Zhang, X., Wang, Y., Li, S., Wang, L., & Tian, S. (2005). Distinctive conflict processes associated with different stimulus presentation patterns: An event-related potential study. *Experimental Brain Research*, *162*, 503–508.

Zhang, Y., Wang, Y., Wang, H., Cui, L., Tian, S., & Wang, D. (2001). Different processes are involved in human brain for shape and face comparisons. *Neuroscience Letters*, *303*, 157–160.

(Received November 6, 2008; Accepted April 29, 2009)