Brain Topography, Volume 13, Number 3, 2001

Brain Electrical Activity Evoked by Mental Formation of Auditory Expectations and Images

Petr Janata*

Summary: Evidence for the brain's derivation of explicit expectancies in an ongoing sensory context has been well established by studies of the P300 and processing negativity (PN) components of the event-related potential (ERP). "Emitted potentials" generated in the absence of sensory input by unexpected stimulus omissions also exhibit a P300 component and provide another perspective on patterns of brain activity related to the processing of expectancies. The studies described herein extend earlier emitted potential findings in several aspects. First, high-density (128-channel) EEG recordings are used for topographical mapping of emitted potentials. Second, the primary focus is on emitted potential components preceding the P300, i.e. those components that are more likely to resemble ERP components associated with sensory processing. Third, the dependence of emitted potentials on attention is assessed. Fourth, subjects' knowledge of the structure of an auditory stimulus sequence is modulated so that emitted potentials can be compared between conditions that are identical in physical aspects but differ in terms of subjects' expectations regarding the sequence structure. Finally, a novel task is used to elicit emitted potentials, in which subjects explicitly imagine the continuations of simple melodies. In this task, subjects mentally complete melodic fragments in the appropriate tempo, even though they know with absolute certainty that no sensory stimulus will occur. Emitted potentials were elicited only when subjects actively formed expectations or images. The topographies of the initial portion of the emitted potentials were significantly correlated with the N100 topography elicited by corresponding acoustic stimuli, but uncorrelated with the topographies of corresponding silence control periods.

Key words: Emitted potential; Imagery; ERP; N100; High-density EEG; Superior temporal gyrus.

Introduction

Mental images and expectancies shape our perception of the world. Auditory expectations play an important role in the way we perceive incoming auditory information - in other words, what we expect to hear affects what we actually hear. For instance, auditory expectations allow us to restore the percept of a portion of a speech sound that has been masked partially by noise (Samuel 1996; Samuel 1981a; Samuel 1981b; Warren 1970; Warren 1984), and they assist us in tracking a single speaker's voice at a cocktail party (Cherry 1953). From a theoretical perspective, "adaptive resonance" (Grossberg 1980) and "reentrant" (Edelman 1989) models propose that interactions of incoming sensory information and experientially derived expectations generate dynamic activity patterns within networks of reciprocal cortico-cortical, cortico-thalamic, and cortico-limbic connections. Unfortunately, very little is known about the actual neurophysiological mechanisms underlying the interactions of internally generated images and expectations with representations of sensory input.

Event-related brain potentials (ERPs) have been used extensively to study the perceptual and cognitive processes associated with the processing of discrete external (exogenous) stimulus events and the modulation of this processing by endogenous factors such as attention and various forms of memory (Näätänen 1992). One hallmark of the auditory evoked potential is the N100, a vertex-negative peak in the ERP waveform occurring approximately 80-120 ms following stimulus onset. The N100 is of particular interest because it is influenced by both exogenous and endogenous factors (Näätänen and Picton 1987) and is believed to arise in the auditory cortex (Liegeois-Chauvel et al. 1994;

---

* Institute of Neuroscience, University of Oregon, Eugene, OR, USA.

Accepted for publication: December 20, 2000.

Dr. Ramesh Srinivasan kindly provided programs for interpolating topographical images. I thank Dr. Michael Posner and Tom Winn for helpful discussions, and Drs. Michael Posner, Ramesh Srinivasan, Mireille Besson, and several anonymous reviewers for comments on earlier versions of this manuscript. The research was supported by the following NIH grants: Systems Neuroscience Training Grant, GM07257, at the University of Oregon , NRSA (NS10395) to P.J., MH42669 to Dr. Don Tucker, and Program Project Grant (NS17778-18) at Dartmouth College. Additional support was provided by the McDonnell/Pew Center for the Cognitive Neuroscience of Attention at the University of Oregon.

Correspondence and reprint requests should be addressed to Petr Janata, Dept. of Psychological and Brain Sciences, Dartmouth College, 6207 Moore Hall, Hanover, NH 03755, USA.

E-mail: petr.janata@dartmouth.edu

Pantev et al. 1995; Verkindt et al. 1995). Thus, neural events surrounding the N100 are particularly relevant to studying interactions of internally derived expectations or images, and representations of the external environment.

Attempts to separate endogenous components of the auditory ERP waveform from exogenous components usually involve a comparison of potentials evoked by sensory stimuli that have been influenced by endogenous factors to varying degrees. In all such cases, the potentials measured reflect an interaction of a sensory stimulus with endogenous factors. If the interaction begins as early as 50 ms (Woldorff and Hillyard 1991), the early interactions presumably influence later interactions, thereby making it very difficult to separate those aspects of the response that are due purely to top-down influences from those that are due purely to sensory influences. Very few attempts have been made to measure potentials associated with endogenous processes absent the sensory input. Thus, one goal of the experiments described in the present paper is to help dissociate the exogenous and endogenous components of evoked potentials as well as provide another perspective on how they interact as the response to an event evolves.

The primary difficulty in measuring responses to internally generated, i.e. imagined or expected, events is due to the uncertainty in determining the onset times of such events and the need to average responses to many such events in order to obtain an ERP. Two attempts by Weinberg and colleagues (1970, 1974) to overcome the problem of estimating when subjects' expectations for sensory events occurred resulted in the successful recording of "emitted" potentials in the absence of an actual sensory event. In one case, potentials were emitted when subjects developed a temporal expectancy for when one stimulus would follow another. In the other, subjects spontaneously signaled whether or not they expected an auditory event to occur. On trials in which they expected to hear a sound but none arrived, an emitted potential was generated, whereas none was present in the absence of an expectation.

Further work on emitted potentials focussed on whether the late positivity observed in response to unexpected stimulus omissions was the same as the P300 observed in response to rare or novel stimuli. Ruchkin et al. (1975) demonstrated a comparable dependence of evoked and emitted P300 amplitude on stimulus probability, and subsequently demonstrated that emitted potentials contain temporal expectations about the time of stimulus occurrence (Ruchkin and Sutton 1978). Emitted potentials to unexpected stimulus omissions in musical and linguistic contexts also include a large P300, consistent with the idea that expectancies are formed and then violated by the omission (Besson and Faïta 1995; Besson et al. 1997).

In principle, expectations of stimuli could consist of not only when the stimulus is expected to occur but also some information about its physical properties, e.g. sensory modality. If there is even a partial dependence of emitted potentials on the sensory modality of the expected stimulus, the topographical distributions of emitted potential features should show a difference across sensory modality and may even show similarities to exogenous components of potentials evoked by stimuli from different modalities. Simson et al. (1976) reported that topographical distributions of early portions of emitted responses resembled those of their evoked potential counterparts in trains of auditory or visual stimuli. With musical stimuli, the large early components of emitted potentials resemble the AEP recorded at midline sites, though the peaks are somewhat delayed relative to peaks in response to audible stimuli (Besson and Faïta 1995; Besson et al. 1997). One impetus for the present study were the results of an earlier ERP study of musical expectancy which showed N1/P2 and sustained positive waveforms at central sites during a period of silence in which participants were to imagine the best possible resolution to a sequence of chords prior to the occurrence of the actual resolution (Janata 1995). In this earlier study, however, the possibility of an N1/P2 waveform indexing auditory imagery was confounded with a possible offset response to the chord immediately preceding the imagery period. Thus, the present experiments were designed to eliminate contamination by offset responses.

A prevalent, yet untested, assumption is that emitted potentials are automatically elicited by stimulus omissions, and that they can be generated only by unexpected omissions. In the auditory ERP literature, a very strong distinction has been drawn between attentionally maintained expectations, e.g. for targets in a target detection task, and automatic expectations formed on the basis of repeated sensory input (Näätänen 1990; Näätänen 1992). The latter type of expectation is not influenced by attention and is indexed by the mismatch negativity (MMN). The MMN appears to be generated by stimulus omissions only when the stimulus onset asynchrony (SOA) is less than 150 ms (Tervaniemi et al. 1994). Thus the same automatic processes that are postulated for the MMN cannot readily explain the bulk of the emitted potential data. Omissions of regularly spaced auditory stimuli do, however, elicit responses during the non-attentive condition of silently reading a book (Joutsiniemi and Hari 1989; Raij et al. 1997). Although the book reading task is used as a condition of auditory inattention in MMN studies, the authors of these papers argue that the emitted responses do not represent the same generators as the MMN because the amplitude of the emitted responses depends on attention.

Overall, the question of whether emitted potentials reflect an automatic response remains in need of testing. Thus, one goal of the two experiments described in this paper was to test whether emitted potentials really represent responses to violations of expectancies generated automatically over the course of a sequence of repeating sounds, or whether they depend on voluntarily formed mental images. More importantly, the present studies seek to establish whether emitted potentials are generated in a situation that does not involve unexpected stimulus omissions. Because the stimulus omissions in the studies described above were unexpected, the emitted potentials all reflect a violation of an expectancy. Will potentials be emitted following the end of a note sequence even when the subject knows that no stimulus will occur? Finally, the second experiment compared emitted responses generated under task conditions that resembled the traditional emitted potential paradigm with potentials emitted as subjects explicitly imagined a continuation of a simple melodic phrase. As such, the experiments represent a preliminary comparison of imagery and expectancy formation. In both studies, the primary focus of the analysis was on the early stages (0-200 ms) of the emitted potential response rather than the P300. Specific comparisons were made between the topographies associated with N100 responses to heard notes, and the topographies to corresponding imagined notes and expected and unexpected silences during the same time windows.

Experiment 1

Methods

Participants & Stimuli

Seven subjects engaged in the experiment, and were compensated $5/hr. The age range of participants was 18-42 years (mean = 26.4). The range of experience playing a musical instrument was 10–37 years (mean = 20.7 years). All subjects provided informed consent and procedures were approved by the Institutional Review Board at the University of Oregon.

Individual notes used for ascending and descending melodic phrases were constructed in Matlab (The Mathworks, Natick, MA) using a sampling rate of 22.255 kHz. Fundamental frequencies corresponding to the notes of a D-major scale starting at 246.94 Hz (D2) were used. Each note was a harmonic sound consisting of harmonics 1–6 of its fundamental frequency. The relative amplitudes of the harmonics were set at 1.0, 0.8, 0.7, 0.6, 0.5, and 0.4. Individual notes were 250 ms in duration and ramped on and off with 2.5 ms linear ramps. The resulting timbre resembled smooth notes created on a brass instrument. Individual notes were converted to 16-bit AIFF format using the "sfconvert" utility on an Indigo 2 workstation (Silicon Graphics, Mountain View, CA) and arranged into melodies using SoundEdit16 on an Apple Macintosh (Apple, Cupertino, CA). The stimuli were stored as 'snd' resources for use by the stimulus presentation/data acquisition program (EGIS, Univ. of Oregon). Stimuli were presented (~65 dB SPL) from an Apple Macintosh Power PC 8100 via a mixer (Radio Shack, Ft. Worth, TX), amplifier (Marantz, Roselle, IL) and two speakers (Miller & Kreisel, Culver City, CA) situated 1 m apart on either side of a computer monitor about 2 m in front of the subject.

Procedures

One ascending and one descending melodic phrase (figure 1A) were used in the experiment. Each melody was presented an equal number of times (50 trials/melody). Assignment of melodies across the trials was random. The experiment consisted of four blocks of 50 trials each. The 1st and 3rd blocks were "active" blocks in which subjects were required to attend to the melodies, make key presses, and generate images. The 2nd and 4th blocks did not require that images be generated, and subjects were free to ignore the melodies.

At the start of each trial, a number appeared on the computer monitor in front of the subject indicating the number of trials remaining in the block. The subject initiated the trial at his/her leisure by pressing a lever. The number disappeared and a fixation point appeared at the center of the screen. The melody commenced 500 ms after the appearance of the fixation point. Each trial in an active block consisted of three sub-sections. In the first sub-section of the trial, subjects heard the entire 8-note melody and were asked to press the lever synchronously with the onset of the last note, as though they were playing an instrument and were to join in with the melody at the time of the last note ["All Heard" (AH) condition, figure 1B]. The SOA between notes in the melody was 500 ms. The fixation point disappeared 750 ms following the offset of the 8th note in the sequence. A 2 s pause preceded the appearance of the next fixation point which signaled the beginning of the second sub-section. Five hundred ms after the appearance of the fixation point, subjects heard the initial five notes of the same melody and were asked to continue imagining the final three notes, pressing the lever at the time they thought the onset of the last note would have occurred ["3-Imagined" (3I) condition, figure 1B]. The fixation point disappeared 750 ms after the time when the last note would have ceased. After a 2 s pause and subsequent appearance of the fixation point, subjects heard the initial three notes of the melody and were to continue imagining the remaining five, pressing the lever when they thought the last note would have occurred ["5-Imagined" (5I) condition, figure 1B].

In contrast to the "active" blocks in which subjects were imagining notes and making timing judgements, trials in "passive" blocks consisted of only one sub-section per trial. The fixation point and timing pa-

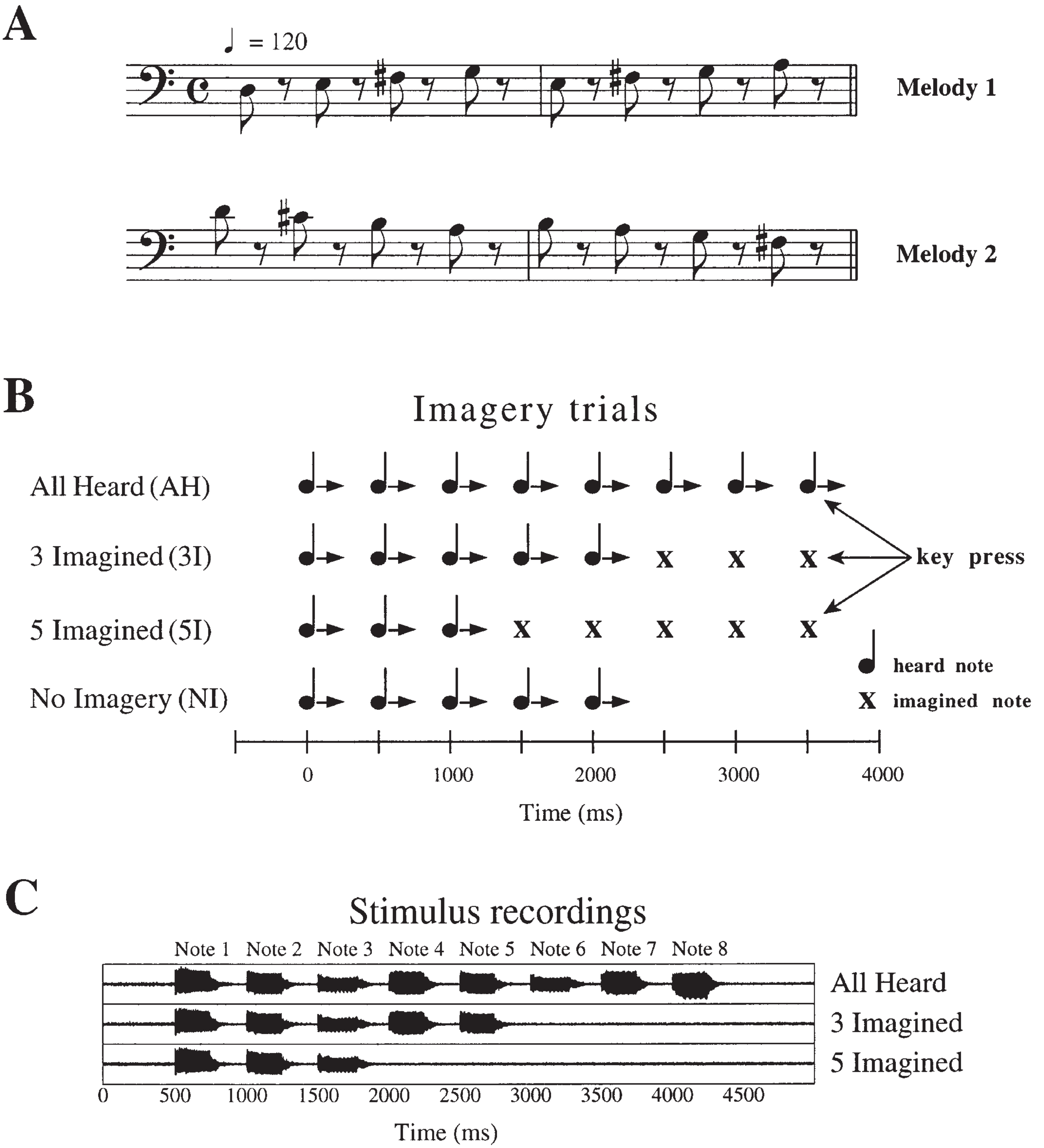

Figure 1. Melodies and experimental conditions used in Experiment 1. A) The two simple melodies in musical notation. B) Schematic diagram of the imagery trials. In the "All Heard" condition subjects heard the entire melodic phrase. In the "3 Imagined condition," the initial five notes were heard and the remaining three imagined. In the "5 Imagined" condition, subjects heard the initial three notes and imagined the remaining five. On each trial, these three conditions appeared in immediate succession. In separate blocks of "No Imagery" trials, subjects heard the initial five notes but did not imagine the remaining three. C) Microphone recordings of the stimulus sequences during one trial taken from the position of the subject's head. Note that the sound from each note decayed fully prior to the onset of the following note.

rameters were the same as the 3I condition in the "active" blocks. However, on any given trial, subjects heard only the initial five notes of the melody. There were no response requirements and subjects were asked to simply ignore the melodies and let them pass by ["No-Imagery" (NI) condition, figure 1B]. It was suggested to them that if they did find themselves listening to the melodies, that they should hear them as a complete phrase and not to imagine their continuation. The melodies were composed such that the five note fragments could be heard as complete motifs. In general, subjects reported no difficulty in either ignoring the melodies altogether or hearing them as complete motifs, although 4/7 of the subjects reported imagining the continuation or repeating the last note on some of the trials. The NI trials were presented in immediate succession, rather than interleaved with the imagery trials. The purpose of these trials was to test whether emitted potentials are elicited automatically when sequences of notes are not actively attended to. We felt the likelihood they would be ignored was greatest if they were presented in blocks.

Prior to the collection of electrophysiological data, subjects received practice trials until their performance on the timed judgements was consistent and they felt comfortable with the task (typically 5–15 trials). Subjects were encouraged to pause and make whatever movements they needed to make between trials in the "active" blocks. It was emphasized to them that their imaging of the melodies should be completely internal with no external movements (finger or toe-tapping) or vocalizations (including sub-vocalizations) to aid in keeping time or imagining the pitches.

The electrophysiological recording session typically lasted 90 minutes. Following the experiment, subjects filled out a short questionnaire with questions about their musical background and subjective evaluation of their performance on the task. Subjects were asked to rate on a scale of 1-7 the difficulty of imagining the notes (1=very difficult; 7=very easy), and the ease with which they could maintain the tempo/rhythm while imagining the notes (1=very difficult; 7=very easy).

Electrophysiological recording and data analysis

The EEG, recorded from 128 scalp electrodes referenced to an electrode at the vertex (Cz) using a Geodesic Electrode Net (Electrical Geodesics, Inc., Eugene, OR), was low-pass filtered (50 Hz) and digitized (12-bit precision) at 125 samples/second. Electrode impedances were in the range of 50 kOhms. All voltage values are expressed relative to the average voltage across all sites (average-reference)(Tucker 1993). In each sub-trial, EEG recording began 500 ms prior to the onset of the first note in the melody and continued for 5 s.

All analyses of the data were performed using Matlab (Mathworks, Natick, MA). Data were analyzed offline for movement artifacts. For any given trial, a channel's data were rejected if any voltage value in a specified time window exceeded 50 μV, the standard deviation of the channel's values exceeded 20 μV or any single value was 6 times the standard deviation of the values for that channel on that trial. If more than 12 channels (9.3% of all channels) were rejected for the particular trial, the entire trial was rejected from further analysis. The time window selected for artifact rejection bracketed the events of most interest and began with the onset of the 3rd heard note and ended 250 ms after the heard/expected offset of the 8th note. Each channel's values were baseline corrected on each epoch by removing the linear trend from the entire epoch.

For each subject, the mean key-press times and standard deviations relative to the onset of the last note in the melody or the predicted onset of the last imagined note were determined for the AH condition and each of the imagined melody conditions, respectively. A single factor repeated-measures analysis of variance (ANOVA) was performed on these values.

Time-domain averages were constructed for each of the experimental conditions by averaging all artifact-free trials for each subject. Grand-average waveforms were computed by averaging the individual subjects' average waveforms. In order to illustrate how the patterns of the potential distributions evolved through time within each of the tasks, interpolated images were created based on the data from 111 of the 129 electrodes using a 3-D spline algorithm (Srinivasan et al. 1996). The interpolations were truncated slightly before the outer ring of electrodes because interpolations at the outer edge were more likely to be inaccurate due to lack of empirical data beyond the outer edge of electrodes. In addition, recordings from the outer ring of electrodes, particularly at posterior sites, tended to be noisy due to poor electrode net fits. The outer ring of electrodes on the Geodesic Nets extends approximately 3 cm beyond the outer ring of the 10-20 electrode placement standard.

The goal of this experiment was to see if the patterns of brain electrical activity, as manifested in the scalp voltage distributions, were similar under active perceptive (AH) and imaginative (3I, 5I) task conditions and different from those obtained in the passive perceptive (NI) condition. Topography similarity was assessed by correlating the topographies observed during corresponding epochs in the different task conditions, and by a permutation bootstrap procedure developed for comparing EEG topographical maps (Galan et al. 1997). Briefly, to compare two topographical maps using this procedure, the voltage values in each map are rank ordered, keeping track of the associated electrode positions. The Euclidean distance between electrode pairs of equivalent rank is computed for each electrode pair. The sum of distances provides a statistic, *D*. The likelihood of obtaining the observed value of *D* given the topographical maps is assessed by comparison with a distribution of the *D* statistic derived from the measured data. A reference sample for the distribution is created by switching voltage values between maps at a random selection of 50% of the electrode sites and computing *D* for the mixed maps. We created *D* distributions using 1000 iterations of this procedure.

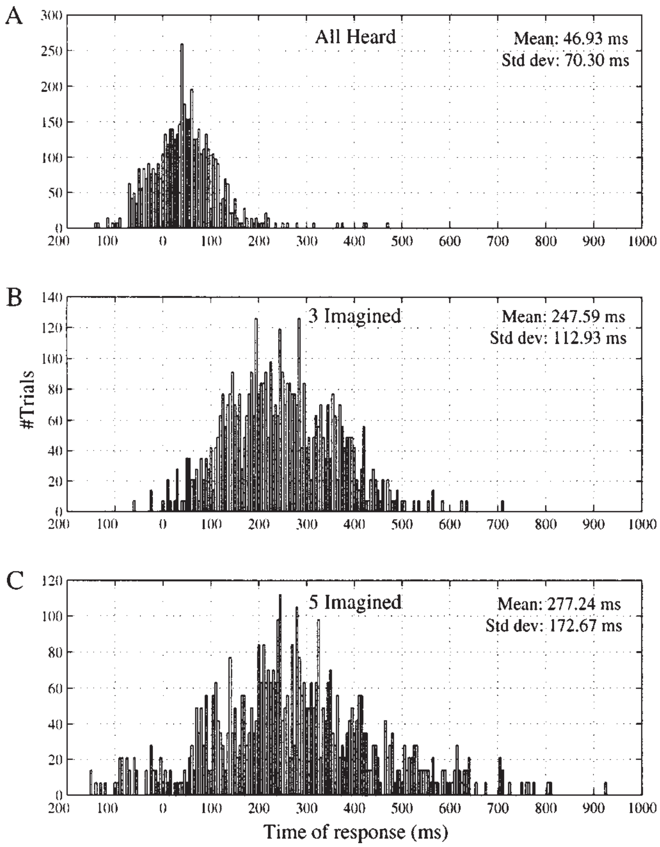

Figure 2. Cumulative distribution of response times from the 7 subjects. The abscissa is expressed in ms relative to the scheduled onset of the 8th note. Responses were collapsed into 5 ms bins. The means and standard deviations of the plotted distributions are indicated in the upper right-hand corner of each plot. A) AH condition in which subjects heard 8 notes and were to press a lever synchronously with the onset of the 8th note. B) 3I condition in which subjects heard 5 notes and proceeded to imagine 3, making a lever press at the time they would have expected the 8th note to have its onset. C) Same as B, except subjects heard 3 notes and imagined 5 (5I condition).

Results

Behavior

Participants reported having no trouble imagining the notes in the melody. The mean rating for the ease of imagining the notes was 6.3 ± 0.8, and 6 ± 1 for ease of maintaining the tempo/rhythm. Four of the subjects reported inadvertently continuing to imagine the melody on some of the trials in the NI condition. Four of the subjects reported a mixed strategy for imagining the notes. On some trials they imagined pure pitches whereas on other trials, they mentally sang a syllable with each pitch in the melody. One subject reported using primarily labels and two reported forming mostly pure images.

In order to estimate subjects' mental timing of their musical images, they were asked to press a lever synchronously with the onset of the last note in the complete melody condition and synchronously with the moment when they expected the onsets of the last notes in the two imagined melody conditions. The cumulative response time distributions across subjects are shown as histograms in figure 2. Times are given relative to the onset of the 8th note in the melody. An ANOVA was performed to test differences in the mean key press times and standard deviations of these times across the three experimental conditions in which subjects were required to make a key press. A highly significant effect was found for both the mean and std. dev. [$F(12,18)_{\text{mean}} = 25.34$, $p < 0.0001$; $F(2,18)_{\text{std.dev.}} = 36.27$, $p < 0.0001$]. Post-hoc Scheffé comparisons ($\alpha = 0.05$) revealed that the mean delay in the AH condition of 47 ms was significantly shorter than the mean delays in the 3I and 5I conditions (248 ms and 277 ms, respectively). The mean delays in the 3I and 5I conditions were not significantly different from each other. The post-hoc comparisons indicated that the standard deviations in the response distributions differed significantly across the experimental conditions. The mean standard deviation increased from 70 ms in the AH condition to 113 ms and 173 ms in the 3I and 5I conditions, respectively.

Not suprisingly, the variability in subjects' responses increased as the time between external timing cues provided by the melody and the target response time increased. The lever-press in the AH condition provided a baseline measure of the variability in the execution of a timed motor act following relative quiescence of motor effector systems but in full presence of external timing cues. Subjects were explicitly instructed to make no movements that would assist them in keeping time. Two subjects reported that making the key press synchronously with the final note interfered with their ability to imagine it, and three subjects commented that making the key press was somewhat of a distraction.

Electrophysiology

The principal aim of this study was to compare the topographies of evoked potentials generated by heard musical events in simple musical phrases with their imagined counterparts in the same, but shortened, musical phrases which the subjects had to complete mentally. The acoustical conditions in each of the task conditions are illustrated in figure 1C by microphone recordings of the melodies made from the position at which subjects were seated. The recordings illustrate that each note was clearly separated from the previous one and that there were no sharp offsets or clicks as the note decayed that might have resulted in evoked potentials induced by acoustic artifacts during times when subjects were required to imagine notes in the melody.

While the topography time-series is rich in information, the display of many voltage maps in many conditions becomes rather unwieldy. For this reason, most of the topographical maps depicted in the figures are topographies of voltages averaged over time windows as long as 96 ms. Figure 3 illustrates that the averaging procedure effectively reduces the number of maps that need to be displayed without eliminating or blurring prominent topographical states. The voltages recorded in response to the first of five imagined notes are shown averaged over 16 ms (2 sample), 64 ms (8 sample), and 128 ms (16 sample) windows (figure 3A-C, respectively). Each colored circle represents a view down onto the top of the head (the nose would be at the top of the circle). The primary topography during the initial 256 ms is characterized by a negative focus centrally and positive ring around inferior peripheral sites. Similarly, both the positive centro-parietal focus between 384 and 576 ms and the frontal positive focus from 705 to 832 ms are readily captured across all scales. What suffer most in the 128 ms averages are the transitions among topographical states, e.g. 512-768 ms in figure 3A. As the parietal-positive focus decays, the frontal positive focus gains in strength. The dashed lines between figure 3A and 3B, and 3B and 3C bracket the corresponding segments of maps. In figure 3C, the entire transition is summarized in the two maps between 512 and 768 ms. Presented only with these two maps, it is obvious that a transition occurs, but it is not possible to discern how smooth or abrupt it is. Nonetheless, the more sustained topographical distributions are well preserved. As these form the principal focus of this paper, topographies averaged over durations of up to 100 ms are considered an appropriate amount of data compression for reducing the complexity of the figures while retaining the prevalent topographical states and transitions.

Figure 4 illustrates the dynamics in the grand-average scalp topographies in each of the task conditions. Because the events in each trial spanned five seconds, the series of topographies is divided into five panels. Within each panel (figure 4A-E), the topographical maps from all four experimental conditions are represented for the purpose of comparison. Every topographical image in the figure represents the average of 12 consecutive time points (96 ms window), with no overlap between successive windows. Each panel is divided by a vertical dashed line into two, 500 ms segments, and each segment corresponds to a single stimulus event. The stimulus event labels are to the left of each row of images.

In addition to showing the topography differences among conditions (rows within each panel) at analogous positions within the melodies, figure 4 illustrates the scalp topographies associated with different ERP components, such as the N100, P200, and P300, and the transitions between them. First, the topographies associated with the stimulus events of most interest are described qualitatively. A statistical comparison of the topographies in the different conditions follows in the section, "Correlations of the potential distributions between task conditions" and figure 6.

Responses to heard notes

The acoustic stimulation was identical across task conditions in the epochs represented in figure 4A and 4B. Not surprisingly, during the 500 ms prior to the onset of the melody, the topographies observed in the four conditions were very similar. This similarity persisted throughout the 1st and 2nd heard notes. The responses to the 1st heard note were characterized by a topography in which the central sites became more negative and the perimeter became more positive. The responses to the 2nd and 3rd heard notes better illustrate the topographical distribution that is characteristic of the N100 component of the ERP, namely a centro-frontal negative focus and positive voltage distribution at inferior sites. The positive peak in response to the 1st heard note had a distinct topography that was characteristic of the P200: positive at central and frontal sites and negative about the perimeter (figure 4A, 2nd segment, window centered on 750 ms, 250 ms after the onset of the 1st heard note). Together, these windows illustrate the topography of the negative peak (N100) to positive peak (P200) transition of the auditory evoked potential (AEP) (see also figure 8 - inset). The topographies shown in response to other heard notes across the conditions show that the N100 epoch was typically characterized by a centro-frontal negative distribution and positive perimeter, while the P200 tended to be represented by a frontal and/or more diffuse positivity. Particularly in the imagery conditions (2nd and 3rd rows within each panel), the topography following the P200 returned to that of the N100 topogra-

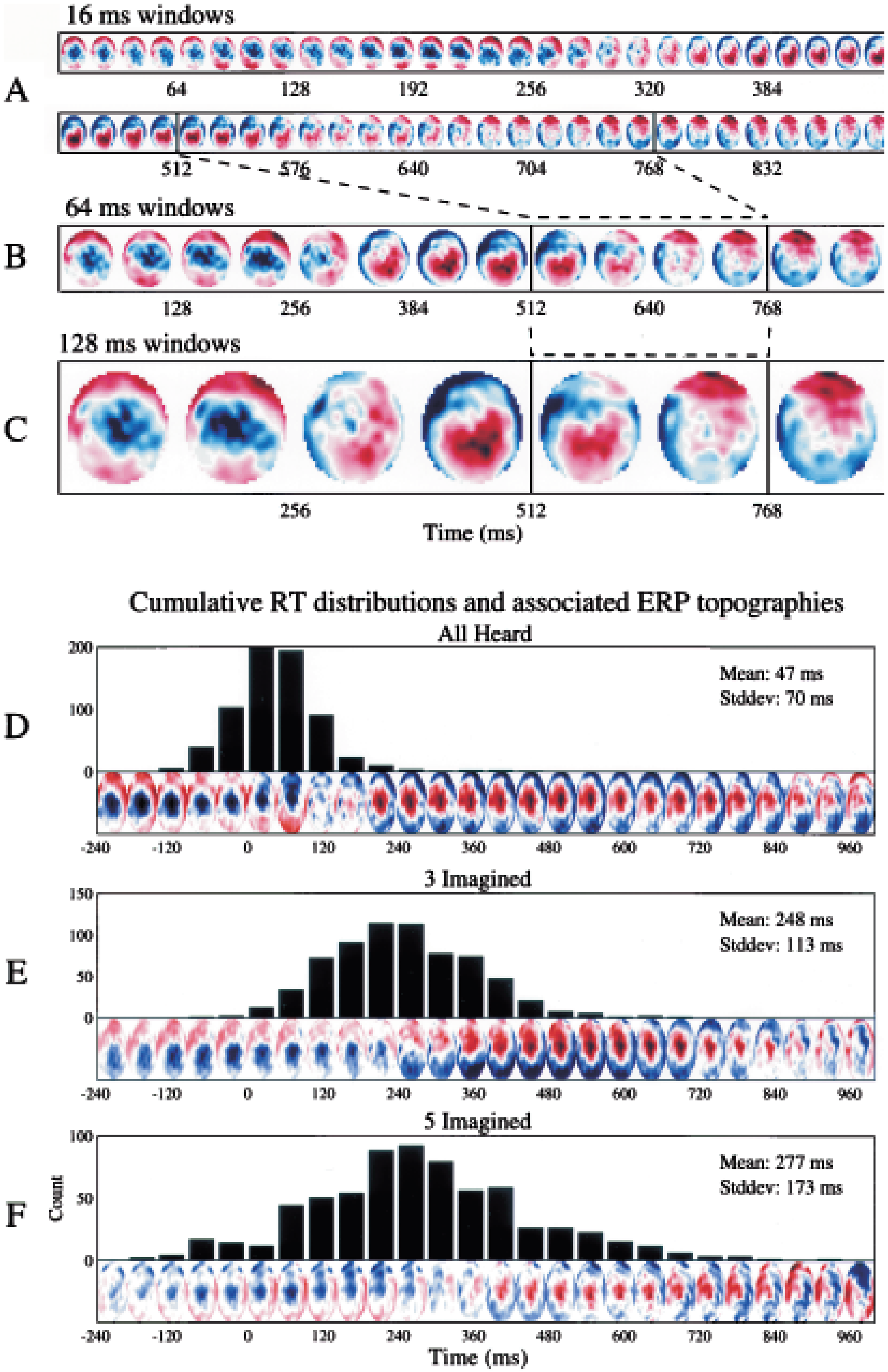

Figure 3. See caption at end of paper.

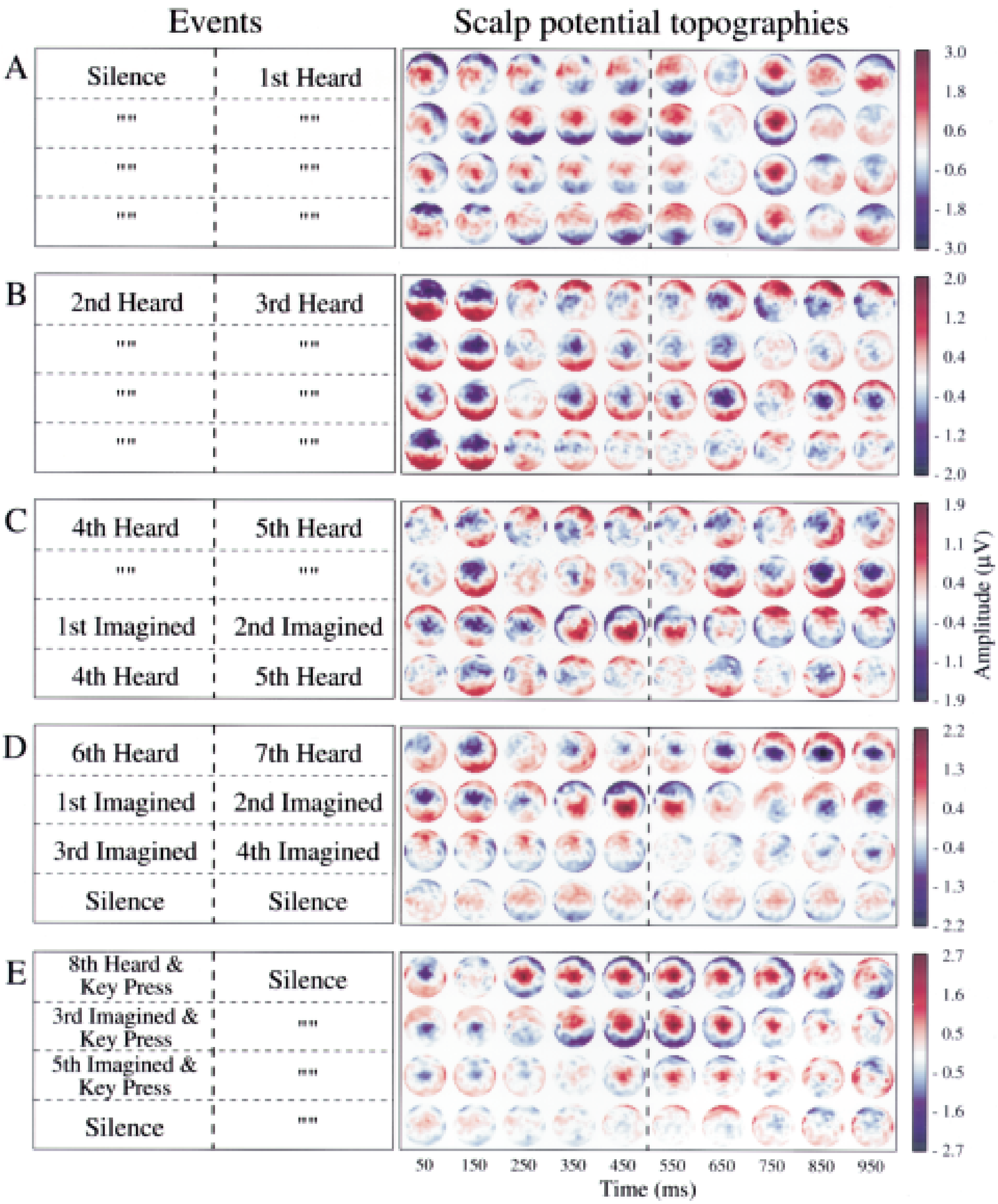

Figure 4. Time-series of the scalp topographies recorded in the four conditions in Experiment 1. The color scales in each panel are scaled to 85% of the largest absolute value occurring during the epoch represented in each panel. Each interpolation is averaged across 12 data samples (96 ms total) and seven subjects. The five panels of interpolations are divided into two 500 ms segments by a dashed line. Each segment corresponds to an event in the task. The events corresponding to each row and segment are specified in the panels to the left. Within each panel, from top to bottom, the rows of circles corresponds to the AH, 3I, 5I, and NI conditions, respectively. Stimulus conditions are identical across the four conditions in panels A & B, but they begin to differ in panel C.

phy, especially in response to the last note preceding the imagination of the remainder of the melody (figure 4B, 2nd segment, 3rd row from the top; figure 4C, 2nd segment, 2nd row from the top).

Responses to the 1st and 2nd imagined note

Of primary interest in this study were the responses to imagined notes. The responses to the first two imagined notes were similar in the two imagery conditions (figure 4C, 3rd row from top; 4D, 2nd row from top, also summarized in figure 5). The topography established during the last 200 ms of the heard note epoch preceding the first imagined note was sustained for 300 ms into the 1st imagined note epoch. This topography was similar to the N100 distribution for heard events, but with an enhanced frontal positivity (figure 4C & D, window centered on 150 ms). When the task did not involve imagery, the voltage distribution generated under identical acoustical conditions was markedly different (cf. figure 4D, rows 2 & 4).

In the imagery conditions, the potential distribution transitioned to a state characterized by a strong centro-parietal positivity at around 350 ms after the predicted onset of the 1st imagined note (figure 4C, row 3; figure 4D, row 2). Between 150 and 250 ms after the predicted onset of the 2nd imagined note, there was a transition in the topography of the potential distribution from the centro-parietal positive distribution to a centro-frontal positive distribution (figure 4C, row 3; figure 4D, row 2 at 650-750 ms). This transition occurred in both of the imagery conditions, but was entirely absent from the NI and AH melody conditions.

The main topographical states assumed in response to the 1st imagined notes are summarized in figure 5. Also shown are the topographies from the equivalent time windows during which a subject was hearing a note in the AH condition and when the subject was not imagining anything in the NI condition. The topographical maps in the imagery conditions differ from the maps in the NI condition, both with regard to the central negativity in the 96–192 ms window (figure 5D, 5G) and the centro-parietal positivity in the 376–480 ms window (figure 5E, 5H). In contrast, the maps in the 96–192 ms window in the imagery conditions are similar to the map corresponding to subjects' hearing the 6th note (figure 5A). The map in figure 5C shows the response when subjects were hearing the 7th note in the melody and preparing to make a key press synchronously with the onset of the 8th note.

Responses to subsequent imagined notes

Following the 2nd imagined note, the tasks and consequently the responses in each of the two imagery conditions differed. In the 3I condition, the potential distributions that were observed during the 2nd imagined note resembled those in the AH condition prior to the key press (compares rows 1 and 2 in figure 4D segment 2, and figure 4E segment 1). In the 5I condition, the frontally positive potential distribution that was established during the 2nd imagined note persisted until the predicted onset of the 4th imagined note (row 3, 4C segment 2, and 4D segment 1), whereupon it changed and ultimately assumed the potential distribution that appeared to characterize preparation and execution of a key press synchronously with the final imagined note.

Responses to preparation and execution of a key press

The timed key presses that were required in the AH and imagery conditions gave rise to sequences of potential distributions which were entirely absent in the NI condition which did not require a key press (figure 4D, 2nd segment; figure 4E). A detail of the topographies is shown alongside the response time distributions in figure 3D–F. The sequence is best illustrated in the AH condition in which subjects were required to make a key press synchronously with the final (8^{th}) heard note. In the 300 ms preceding the onset of the 8^{th} note, a potential distribution was established that was strongly negative at central sites and strongly positive frontally and around the perimeter (figure 3D). Shortly after the mean key-press time there was an inversion in the polarity of this topography. The same general pattern of a central negativity and peripheral positivity that transitioned to a central positivity shortly after the average key press time was observed in both imagery conditions (figure 3E,3F). The delay in the topography transition in the imagery conditions compared to the complete melody condition followed the differences seen in the response time distributions.

Correlations of the potential distributions between task conditions

The similarities and differences across potential distributions shown in figure 4 were assessed statistically by computing the correlation coefficients between the average topographies in the 3I condition and every other condition for each 96 ms window. Correlation coefficients were computed individually for each subject, and the correlation coefficients at each time window were compared with a two factor (window X comparison) repeated-measures ANOVA, where the "comparison" refers to the correlation between the 3I condition and each of the other conditions. The ANOVA showed a significant main effect of window ($F[19,114]=4.53$, $p < 0.0001$), comparison ($F[2,12]=12.41$, $p < 0.005$), and a significant interaction between the two factors ($F[38,228]=2.17$, $p <0.001$). Figure 6 shows the changes in the average correlation coefficients across two different epochs in the task conditions. Windows in which there was a significant difference in correla-

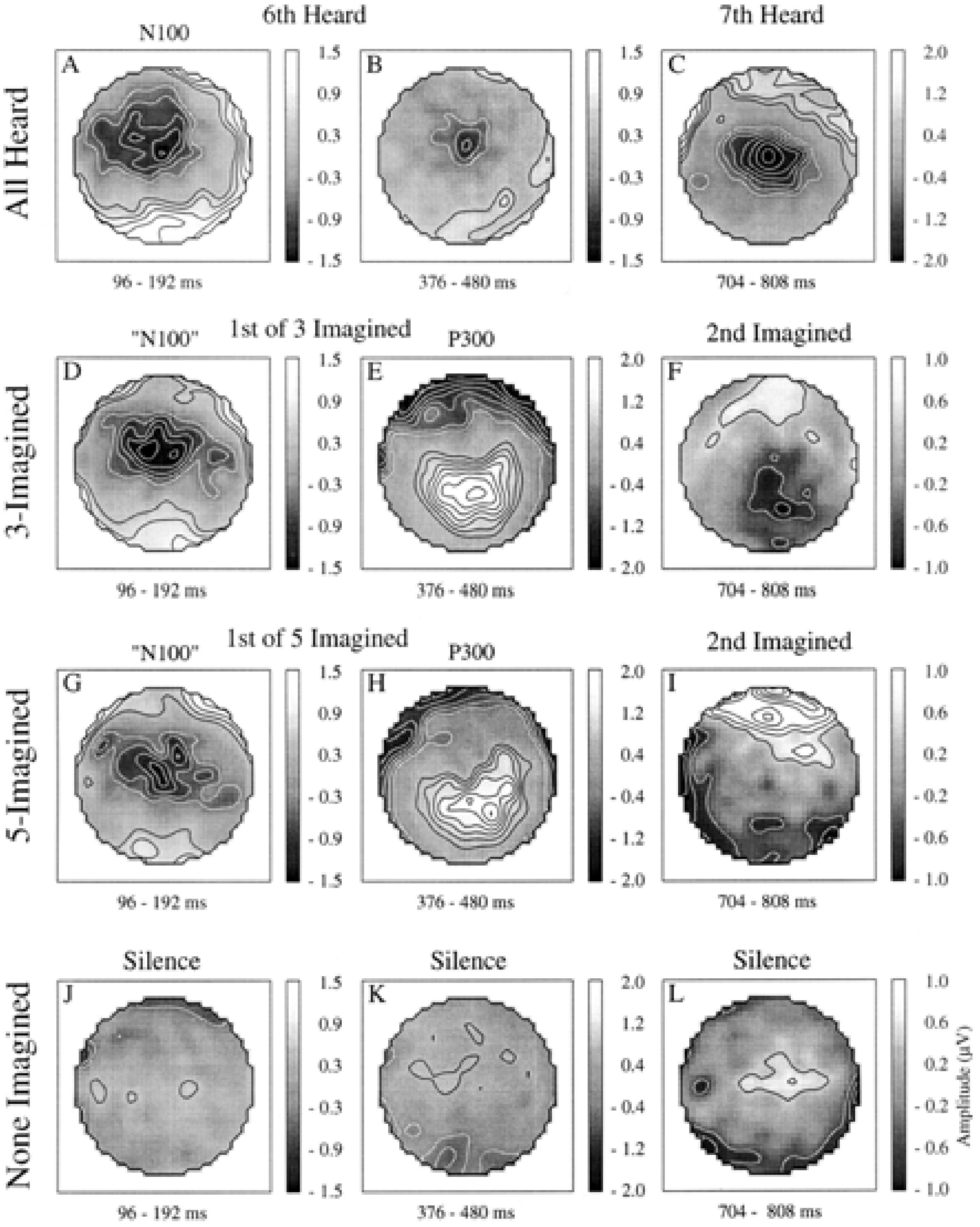

Figure 5. Summary of key topographical states in the imagery task. A) Average topography during the period of the N100 in response to the 6th note in the melody. B) Topography between the offset of the 6th heard note and onset of the 7th heard note. C) Topography after the onset of 7th heard note and at the beginning of preparations to make a key press. D, E) Average response to 1st imagined note in the 3I condition. F) Topography at the time the 2nd note was to be imagined and preparation for key press should have commenced. G, H) Average response to 1st imagined note in the 5I condition. I) Topography at the time the 2nd note (of 5) was to be imagined. In this condition, preparation for the key press was still 1 s away. J,K,L) Response in the no-imagery condition. This condition was acoustically identical to the 3I condition shown in the second row (D-F). For more detail about the transitions linking the topographies shown here, see figure 4.

tions among conditions are marked with an asterisk. The correlation coefficients among the topographies resulting from identical acoustical stimulation during the first and second heard notes of the melody are shown in figure 6A. As expected, a repeated-measures ANOVA on only these time windows showed that for none of the time windows did the correlations differ significantly among comparisons (F[2,12]=3.31, n.s.), indicating that the auditory responses to the initial two notes of the melody in the 3I condition were equally similar to the responses to the same acoustic events in each of the other conditions.

During the imagery portions of the trials (figure 6B), the correlations between the topographies evoked under the various task conditions differed significantly (F[2,12]=13.79, $p < 0.001$). The solid bars indicate that the correlation between topographies elicited in the imagery conditions during imagery of the 1st and 2nd imagined notes remained high. In contrast, the topographies from the acoustically identical 3I and NI conditions were uncorrelated. Correlations with the AH condition tended to assume intermediate values.

The similarity of topographies in the N100 time window (96-112 ms) was further assessed across conditions using the *D* statistic and permutation technique described above. All pairwise comparisons between the conditions were computed. The topography of the N100 for the heard condition did not differ significantly from the topographies during the corresponding time window in either of the imagery conditions, but it did differ significantly ($p<0.01$) from the topography in the NI condition. Similarly, the 3I and 5I condition topographies did not differ significantly from each other, but they did differ significantly from the NI topography ($p<0.0001$ in both cases).

Discussion

Across task conditions, sequences of stable states and transitions could be attributed to hearing a sequence of notes, imagining a sequence of notes, making a timed key press, and waiting quietly. Importantly, activation patterns were correlated between the two imagery conditions, but uncorrelated with the acoustically identical no-imagery condition. The significant similarity between the N100 topography to the sixth heard note and the topographies at corresponding time-points in the imagery, but not the NI conditions (figure 5) is consistent with the hypothesis that similar regions are recruited as the process of mentally completing a melody in one's mind continues in the absence of sensory input. However, this hypothesis must be restricted to the first imagined note following a sequence of audible notes, given that the same pattern was not observed for subsequent imagined notes.

The observed responses to the first imagined note fall into a class of brain responses known as "emitted potentials." Emitted potentials are potentials that are measurable in the average ERP, i.e. reliably evoked over multiple repetitions of a preceding stimulus, that occur in the absence of an external stimulus. Emitted potentials have been described for situations that require a subject to expect a stimulus, but none occurs. Under such conditions, a potential is emitted at the time when the stimulus was expected to occur but was unexpectedly omitted (Besson and Faïta 1995; Besson et al. 1997; Simson et al. 1976; Sutton et al. 1967; Weinberg et al. 1974; Weinberg et al. 1970). In the case of musical stimuli, the emitted potentials in ERP traces recorded at midline sites resemble the AEP, though the peaks are somewhat delayed relative to responses to audible stimuli (Besson and Faïta 1995). Emitted potential components that resemble sensory ERP components are followed by a P300 (Besson and Faïta 1995; Besson et al. 1997; Simson et al. 1976). The most common explanation offered for the presence of emitted potentials is that they represent a temporal expectation for a sensory event (Besson et al. 1997; Simson et al. 1976; Weinberg et al. 1974; Weinberg et al. 1970). Because these studies generated emitted potentials by omitting expected stimuli, the P300 elicited by an unexpected omission has been considered to represent the same process indexed by a P300 component elicited by infrequent stimuli (Besson et al. 1997). Emitted potentials to the first imagined notes in this study also exhibited a response that had the centro-parietal distribution and occurred in the time window that is characteristic of the P300 (figure 5E,H). Notably, this component was elicited when there was explicit knowledge that no stimulus would occur. Thus, the P300 component of the emitted potential cannot be construed purely in terms of a response to an unexpected omission.

As mentioned in the Introduction, the dependence of emitted potentials on attention has not been tested, nor has the requirement for voluntary expectancy formation been fully explored. The emitted potentials in the studies cited above could be interpreted either as involuntary responses to the disruption of a regularly repeating sequence of sounds, as a response to an explicit temporal expectation that is not met, as a manifestation of the expectancy formation process itself, or some combination of these factors.

The NI condition in the present experiment was added specifically to assay whether potentials would be emitted automatically in the absence of auditory attention. The lack of an emitted potential in the NI condition and a sustained topography that was distinctly different from those in the imagery conditions speaks against the emitted potential representing an automatic response to a gap in, or termination of, a regularly repeating sequence. Instead, potentials were emitted when the task explicitly required the generation of a specific image at a

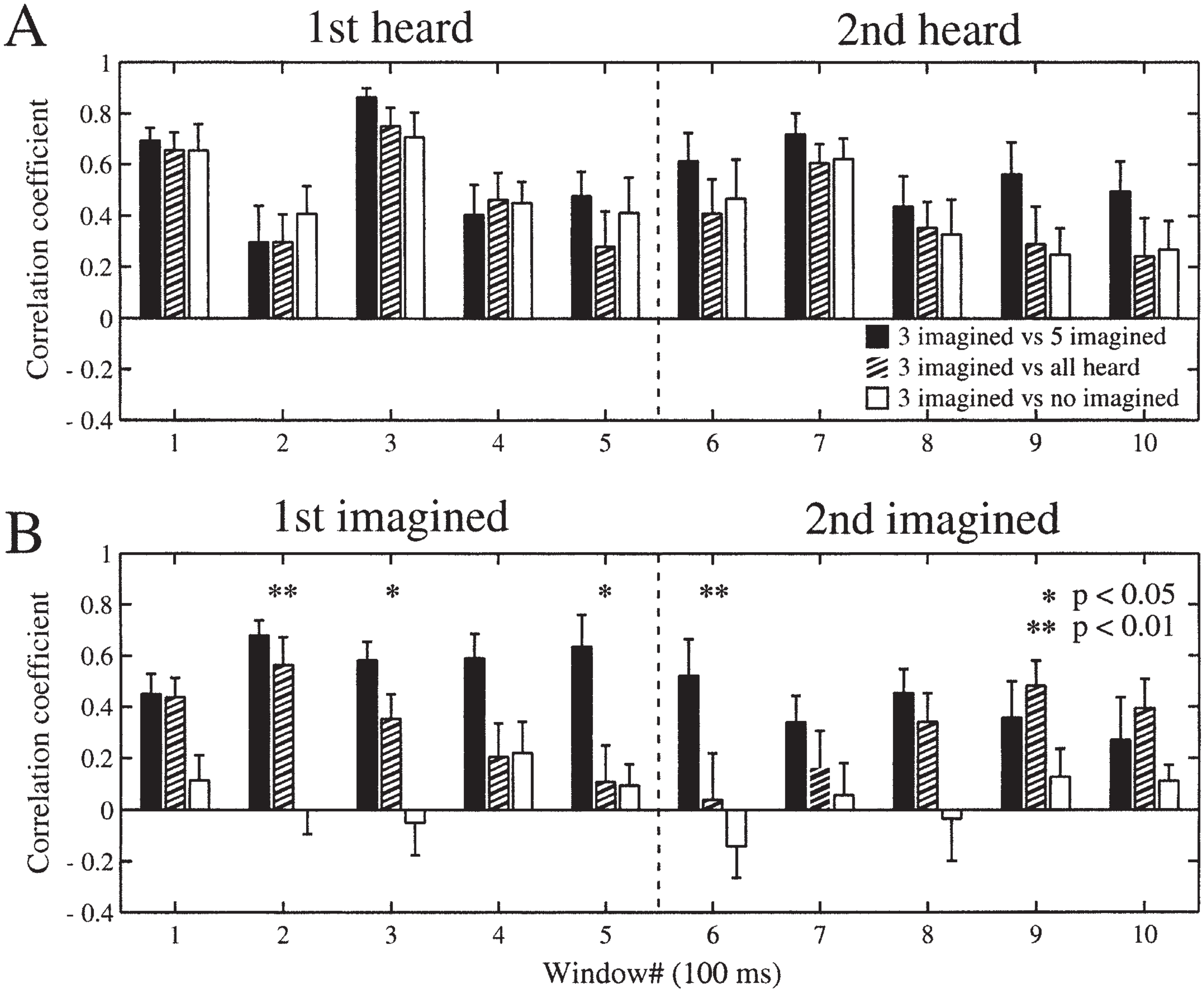

Figure 6. Time-series of the correlations among the experimental conditions. For each subject, the average scalp-topographies in successive 96 ms windows (such as those shown in figure 4) from the 3I condition were correlated with corresponding topographies in the other task conditions. The correlations between conditions (averaged across subjects) are plotted in A & B. Error bars indicate the standard error of the mean. A) Black bars show the correlation of evoked responses during the 1st and 2nd heard notes in the 3I and 5I task conditions. The hatched bars show the correlation between the 3I and AH conditions during the 1st and 2nd heard notes. The white bars shows the correlation between the 3I and NI conditions during the 1st and 2nd heard notes. The acoustic stimulation was identical across task conditions during this 1 s epoch. The correlations are comparisons of the topographical images from the 2nd segment in figure 4A and 1st segment in figure 4B. B) The black bars show the correlation between the 3I and 5I conditions during the 1st and 2nd imagined notes in each condition. Thus, the correlations are a comparison of the topographical images in the 3rd row from the top in figure 4C and the images in the 2nd row from the top in figure 4D. The hatched bars show the correlation between the 1st and 2nd imagined notes in the 3I condition and the 6th and 7th heard notes in the AH condition (top 2 rows in figure 4D). The white bars show the comparison of the 1st and 2nd imagined notes in the 3I condition and corresponding silence in the NI condition (2nd and 4th rows from the top in figure 4D). Asterisks indicate windows in which there were significant differences in the correlation coefficients between conditions.

specific moment in time in the complete absence of an expectation that an actual external stimulus would occur. Thus, these results suggest that it is possible to use ERPs to isolate the brain activations underlying mental image formation as an act unto itself, absent from the expectation that another stimulus will immediately follow.

However, the absence of an emitted potential in the NI condition could, in principle, be attributed to the block design of the experiment, in which subjects could completely ignore the stimuli if they so desired. In other words, attentional load was not well-matched between the imagery and no-imagery conditions. While it was unlikely that they completely ignored the stimuli, given their reports that they would on occasion inadvertently imagine a note following the partial sequence, we decided to test additional subjects in an experiment that required attention and expectancy formation in both imagery and no-imagery conditions.

Experiment 2

The purpose of this experiment was to replicate the findings of Experiment 1 and to more fully explore the range of task conditions under which emitted potentials might be generated. In order to better match the attentional demands on the participants across the imagery and non-imagery trials, a task was devised in which subjects had to make a simple judgment about the same note sequences that were used in the imagery trials but the judgment did not rely on maintenance of auditory images in working memory. In the "Cue Validity" task (described in more detail below), participants received trials in which they 1) expected 3 notes and heard 3 notes, 2) expected 5 notes and heard 5 notes, 3) expected 5 notes but heard only 3, and 4) expected 3 notes but heard 5. Conditions 1 and 2 allow one to look for emitted potentials following the termination of a sequence when subjects are attentive but know the sequence will end at a given time. Condition 3 is analogous to the emitted potential studies above in which the sudden cessation of the sequence constitutes an unexpected omission. Condition 4 looks at responses to unexpected continuations. To our knowledge, this type of response has not been investigated previously.

Methods

Participants & Stimuli

Data were collected for three participants in this experiment. Participants received $5/hr. One of the participants had been a participant in Experiment 1 also. The range of experience playing a musical instrument was 22–30 years (mean = 25.3 years). The age range of the participants was 32-40 years (mean = 34). All subjects provided informed consent and procedures were approved by the Institutional Review Board at the University of Oregon. The acoustic materials were the same as those in Experiment 1. Stimuli were presented via Sony earphones positioned in the subjects' outer ears.

Procedures

This experiment consisted of two trial types: "5 Imagined Notes (5I)" and "Cue Validity" (CV) judgments (figure 7). In contrast to the NI trials of the first experiment, the no-imagery (CV) trials in this experiment required the subject to maintain attention and make a judgment about the number of notes they heard. The 5I and CV trials were randomly interleaved. The melodies were the same as those used in Experiment 1. In contrast to Experiment 1, subjects did not hear the complete melody condition (AH) on each trial. The AH condition was omitted from the present experiment given the relative ease with which participants performed the imagery task in the first experiment and the need to keep the overall duration of the experiment within reason. With only the necessary conditions described below, the overall duration of the experiment, including electrode application and post-experiment proceedings, was 2 hours in duration. Prior to the experiment, we played the two melodies to the subjects until they were confident they could reproduce them in their minds. Overall, the subjects reported more difficulty (mean ease of imagery rating of 4.5) in imagining the notes than did the subjects in the first experiment (mean ease of imagery rating of 6.3). For this reason, two of the three subjects were given the opportunity to refresh their memory of the melodies during breaks that occurred every 10 trials.

At the beginning of each trial, the words "Imagine" or "Do Not Imagine" appeared on the screen to indicate the trial type. In the Imagine condition (figure 7A), the first note was played 1250 ms following onset of the visual cue. As in Experiment 1, the SOA between notes was 500 ms. The subjects heard the initial 3 notes of one of two 8-note melodies and they had to continue imagining the remaining 5 notes. In contrast to Experiment 1, the subjects were not required to make a key-press synchronously with the last note, thereby allowing us to record brain activity uncontaminated with the preparatory phase of making a key press. However, 1 s following the expected onset of the last imagined note, the subjects heard one of six probe notes. Three of the probe notes were the same as the last three notes of the melody they were supposed to imagine. The other three probe notes fell outside of the melody, but were constrained to be one semitone distant from the within-melody probes. The subjects' task was to determine as quickly and accurately as possible whether the probe note was the same or different than one of the last

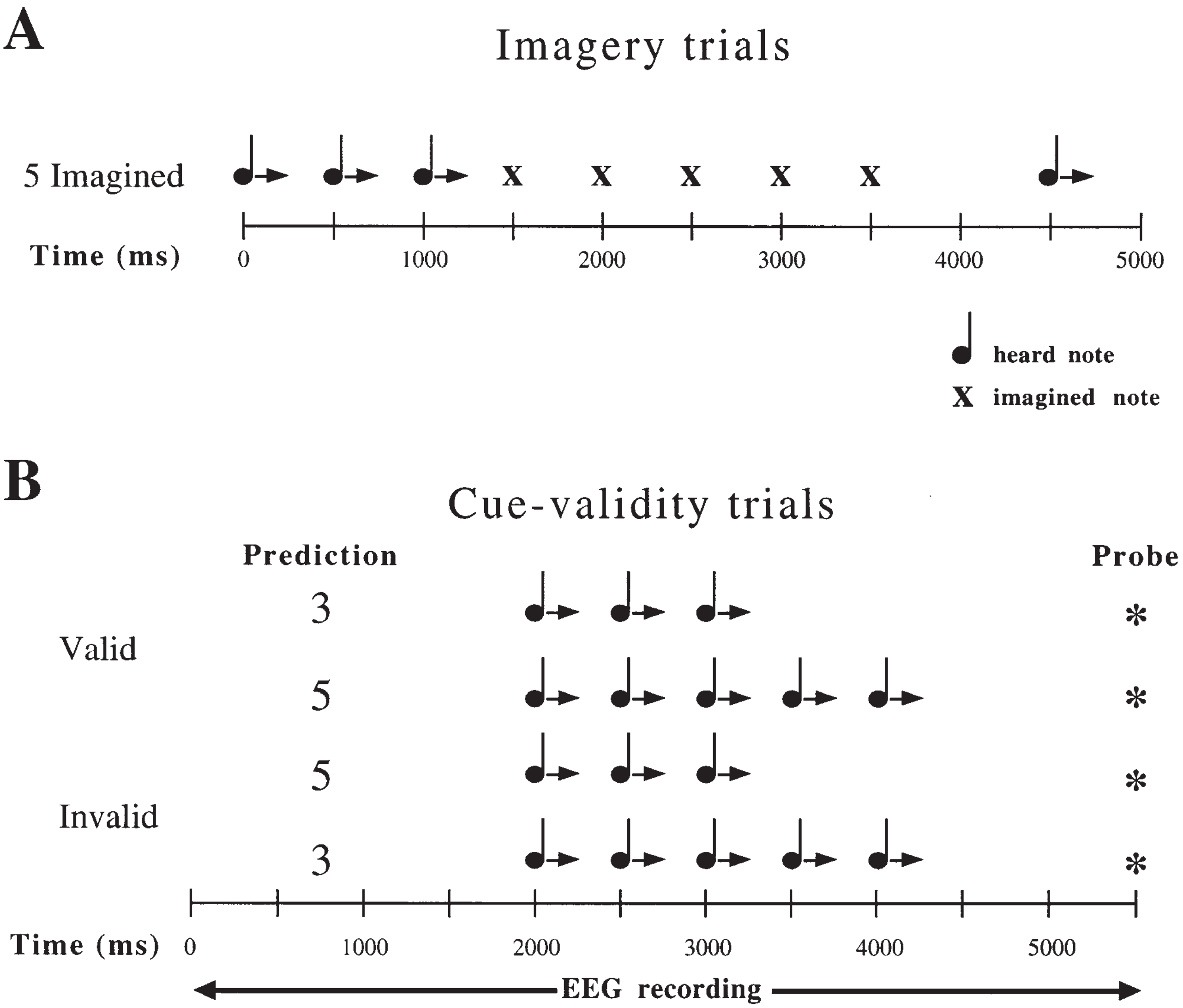

Figure 7. Experiment 2 methods. A) In all imagery trials, subjects heard only the initial three notes of the melody and imagined the remaining five. No key press was required. A probe note was presented 1000 ms following the scheduled onset of the last imagined note, and subjects responded whether or not this note was one of the last three notes they had heard. B) Cue-validity trials. These trials formed a set of control conditions whose purpose was to establish conditions under which emitted potentials would or would not be elicited. On each trial, a visual cue appeared on the screen predicting that either 3 or 5 notes would be heard. Either 3 or 5 notes were then presented. A visual stimulus was presented at fixation 1250 ms following the offset of the last note and subjects responded whether the prediction had been accurate. In 90% of the trials, the prediction was valid. In 10% of the trials the prediction was invalid. There were two types of invalid trials. Presentation of fewer notes than expected resulted in an unexpected omission. Presentation of more notes than expected resulted in an unexpected continuation.

three notes they had imagined. The next trial was initiated 1500 ms following the subjects' response.

In the CV conditions (figure 7B), a number cue appeared on the screen 750 ms after the trial type cue. This number predicted that they would hear either three notes or five notes. 1250 ms following the appearance of the number, subjects heard either three or five notes of the melody (500 ms SOA). On 10% of the trials, the prediction was violated. Note that the violations were of two types. Sometimes, subjects heard only three notes when they were expecting five. This condition is analogous to the traditional omitted stimulus paradigm in which a repeating stimulus is unexpectedly omitted some fraction of the time. In other CV trials, subjects heard five notes when only expecting three, thereby constituting an unexpected continuation. One second after the onset of the 5th note epoch, a visual cue appeared on the screen and subjects had to respond as quickly and accurately as they could whether or not the prediction matched the number of notes they actually heard. The CV conditions required that subjects maintain attention, but did not force them to retain or form an image in auditory working memory when the predicted number of notes had passed. Subjects performed 120 imagery trials and 400 no-imagery trials. Thus, data were collected for 20 of each of the expectancy violation trials in the no-imagery task.

Electrophysiological recording and data analysis

The EEG was recorded from 129 electrodes and examined for artifacts as described above, with the sole difference that data were acquired at 250 samples/s and the low-pass filter was set to 100 Hz. Given the small number of subjects, each subject's average data were standardized. Z scores were computed based on the mean and standard deviation of average voltage values on all good channels across all time points in all task conditions. The standardized subject averages were then averaged together to form the grand-average waveforms. Topographical maps were constructed using spline interpolation as described above.

As in the first experiment, the point of the present experiment was to compare topographies during acoustically equivalent epochs experienced under different task conditions, and to compare topographies elicited by heard notes with topographies elicited by other task events such as imagined notes, unexpected omissions or continuations, and other heard notes. Grand average topographies in the time window from 92-108 ms after the onset of the 4th event epoch were compared using the permutation test of the *D* statistic described above. This window was specifically chosen so that the N100 topography for the 4th note in the 5 predicted, 5 heard CV condition could be compared with the topographies in the other conditions.

In addition to the permutation test on the *D* statistic, the topography (vector of voltage values across electrodes) at each time point was correlated with the topographies at all other time points. The correlation matrix **R** is given by,

$$\mathbf{R}_{\mathbf{ij}} = \frac{\sum_{\mathbf{k}} \mathbf{X}_{\mathbf{i},\mathbf{k}} \mathbf{X}^{\mathbf{T}}_{\mathbf{j},\mathbf{k}}}{\sqrt{\left(\mathbf{X}\mathbf{X}^{\mathbf{T}}\right)_{\mathbf{i},\mathbf{i}} \left(\mathbf{X}\mathbf{X}^{\mathbf{T}}\right)_{\mathbf{j},\mathbf{j}}}},$$

where **X** is the time (rows) by electrode (columns) matrix of voltage values. By specifying the correlation between topographies recorded at different time points, the correlation matrix **R** allowed us to identify temporally stable topographical patterns in the data, including those that correspond to the N100 and P200. Furthermore, the correlation matrix facilitated a statistical comparison of responses to the different heard notes in the melody as well as activity during imagery or silence.

The correlation values were transformed to t-values (Hays 1988, pg. 589). Given the large size of the correlation matrix, and thereby the large number of multiple comparisons, the significance of each t-value was corrected by Bonferroni correction for multiple comparisons. Because the correlation matrix is symmetrical and the correlation values along the diagonal are necessarily 1.0 and can therefore be ignored, the total number of multiple comparisons for the 1300 time points and five task conditions was ~4.22 x 10^6. Thus, the corrected p-value corresponding to a family-wise a=0.05 was ~1.19 x 10^{-8}. For df=127, this value corresponds to a correlation coefficient of 0.47.

Results

The grand-average activation time-course for all channels across the entire trial in each of the conditions is shown in figure 8. The data are presented as electrode X time voltage (ETV) matrices to highlight the patterns that are associated with the various events within and across different conditions. The matrices in figure 8C and 8D appear to be less smooth than the matrices in the other panels. This is due to a smaller number of trials (max of 20/subject) in these conditions compared to the other conditions (>100 trials/subject). The first note in the melody (onset at 0 ms) generates the same clear ETV pattern across the different conditions. At 100 ms this pattern appears as a beaded column of red (positive) and blue (negative) voltages which is distinct from the preceding topography in which channels 30-100 appear as mostly blue and the remaining channels as mostly red. When viewed as a topographic map (left inset above figure 8A), the beaded pattern is shown to correspond to the N100. The N100 pattern largely inverts polarity during the following 100 ms, and represents the P200 component of the AEP (middle inset above figure 8A). A distinct N100/P200 pattern is also observed in response to the probe tone in the imagery condition at 4500 ms (figure 8A).

The responses to successive heard notes in the melody appear in the ETV matrices as repetitions of the basic ETV pattern that is elicited by the first note in the melody (most clearly seen in figure 8A,B,E). The ETV response pattern to the first note is considered to encompass the portion of the matrix from just after 0 ms to shortly after 500 ms. Aside from the N100/P200 patterns already described, another feature of the ETV response to the heard

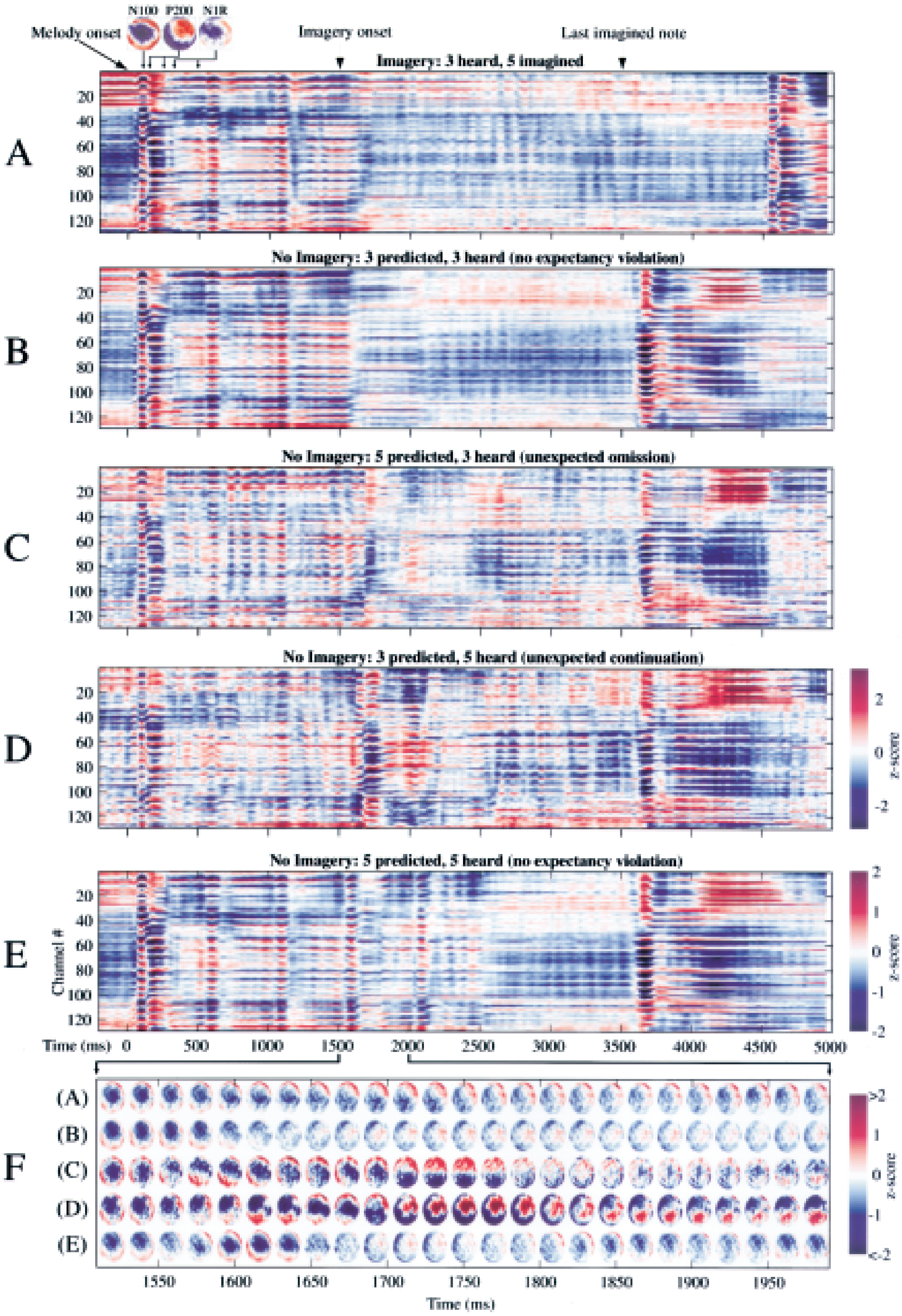

Figure 8. See caption at end of paper.

note is a re-emergence of the N100 distribution in the last 150 ms of the note epoch (right inset above figure 8A). Three repetitions of this basic pattern were observed when subjects heard three notes (figure 8A,B), and five repetitions are observed when they heard five notes (figure 8E). Note that the P200 pattern is strongest in response to the first note in the melody.

The prominent patterns that appear as vertical bands at ~3700 ms in figure 8B-E, represent the response to the visual stimulus prompting subjects to indicate whether or not the number of notes they heard was the number expected.

The ETV patterns of most interest are those that occur between 1500 and 2000 ms. In this time window, the tasks and sensory input across the conditions diverge. A detail of the ETV matrices during this epoch are shown as topographic maps in figure 8F. The response to the fourth heard note when five notes were expected is seen at 1600 ms in Row E (figure 8F). The top three rows in figure 8F show the responses in the three conditions in which only three notes were heard, but in which the tasks differed. The top row (A) shows the response when subjects imagined the first of five notes. The second row from the top (B) shows the response when subjects' expectancies of hearing only three notes was fulfilled. The third row from the top (C) shows the response when the subjects' expectancies of hearing five notes were violated. Each of these topographies should be compared with the topographical pattern in the bottom row (E) when a note was actually heard.

When five notes were expected but only three heard (figure 8F, Row C), a distinct topographical sequence emerged which was absent when the sequence terminated as expected (figure 8F, Row B). Thus, the topographical pattern in Row C of figure 8F represents an emitted potential to an unexpected auditory stimulus omission. As in Experiment 1, the topographies occurring around the peak of the N100 of the heard stimulus (92–108 ms from the onset of the epoch, or 1592–1608 ms on the timescale in figure 8) were compared using the D statistic. The difference in topographies to the unexpected omission in the 5-expected, 3-heard condition was significantly different from the topography in the expected termination in the 3-expected, 3-heard condition ($p < 0.05$) but not significantly different from the N100 topography to the fourth heard note in the 5-predicted, 5-heard condition ($p = 0.74$). In the imagery condition (figure 8F, Row A), the topography during the N100 window also differed significantly from the topography in the 3-expected, 3-heard condition ($p<0.05$) but did not differ significantly from either the emitted potential in figure 8F, Row C ($p=0.22$) or the N100 in figure 8F, Row E ($p=0.27$).

Correlation matrices

A more detailed statistical analysis of the similarities of topographies to different events within the task conditions is shown in figure 9. Shown at the left of figure 9 are the correlation matrices obtained by correlating the vector of voltage values at each time point with the voltage vectors at each other time point. Shown in black at the right of figure 9 are those correlations which exceeded the Bonferroni-corrected significance threshold of $|r| > 0.47$. Thus, the correlation between any pair of topographies can be considered statistically significant if the corresponding point in the correlation matrix on the right hand side of figure 9 is black. The triangles indicate the onsets of auditory, visual and imagined events. To determine how strongly the topography at any given reference time point in the trial was correlated with the topography at any other time point, look at the matrices on the left-hand side of figure 9 and locate the reference time point along the diagonal. Points directly above the point on the diagonal represent correlations with preceding topographies, while points extending to the right represent correlations with future topographies.

Several salient features emerge from the correlation maps. First, persistent topographical states appear as triangles with their base along the diagonal. For example, following the fifth heard note, there is a large triangle preceding the visual cue, indicating that the topographical pattern between ~2600 and 3500 ms was stable. Figure 8E shows the stable pattern. The same pattern following the expected termination of the three note sequence (figure 8B between 2100 and 3500 ms) shows up as a large dark triangle in the correlation matrix in figure 9B. Similarly, the repetitions of the ETV patterns described above for responses the individual heard notes, are seen as smaller dark triangles. Patches of increased correlations, appearing as squares off the diagonal, during the heard notes show that evoked responses to the different notes in the sequence elicited significantly similar topographies.

Two important observations can be made about the brain activity patterns that follow the last expected heard note in the imagery and no-imagery conditions. First, in the imagery condition (figure 9A), topographies throughout the 2 s period of imagery (1500-3500 ms) are strongly correlated not only with themselves, but also with topographies elicited during the preceding notes. The topographies during imagery are not as strongly correlated with the period between the end of imagery and onset of the auditory probe, suggesting a cessation in the imagery processes after the requisite number of notes had been imagined. This shift in topography between imagery and rest is also visible in figure 8A. Second, the correlation matrices differ between the imagery and the

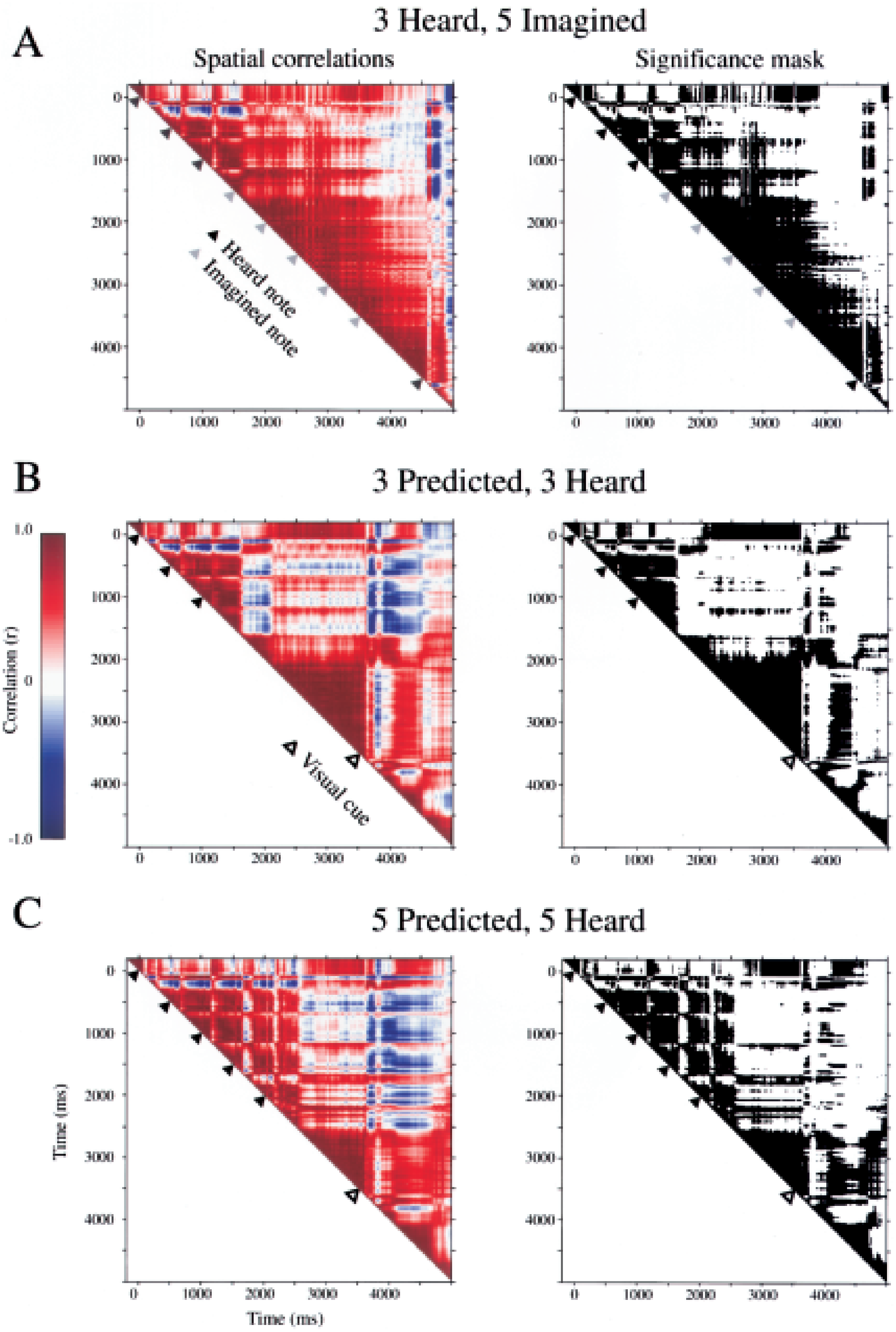

Figure 9. See caption at end of paper.

acoustically identical 3-expected, 3-heard condition. In particular, the period between 2200 and 3500 which showed a substantial number of correlations with the heard notes in the imagery condition, showed very few significant correlations with the heard notes in the no-imagery condition. However, during the silence immediately following the last heard note, the topography in the 3-expected, 3 heard condition was strongly correlated with the P200 topography, and therefore inversely correlated with the N100 topography. This effect shows up in figure 9B as the light blue squares above the dark red triangle along the diagonal between 1700 and 2200 ms. While the correlation with the P200 topography during this time was significant, the inverse correlation with the N100 topography was not. Surprisingly, the duration of this effect was equal to the duration of the inter-event-interval, even though there were no exogenous timing cues! Thus, there appears to be an emitted potential effect in the no-imagery condition, but it is a different process from that observed in the imagery condition, given the different topographies.

The mental state at the end of the 5-expected, 5-heard condition no-imagery condition was conceptually identical to that of the 3-expected, 3-heard condition described above. Subjects knew that there would be no continuation of the sequence beyond the fifth note. A similar, albeit shorter-lasting and less significant sequence termination effect following the last heard note was observed in the topographical pattern in this condition also. Thus, there appeared to be something particularly special about the expected termination of the sequence after three notes, which gave rise to a topography that was significantly correlated with the P200. One possibility is that subjects knew that their expectations would be violated on some number of trials (10% to be exact), either by an overshoot or undershoot of the number of predicted notes. Whether an over- or undershoot would occur was always determined uniquely during the 4th note epoch. In other words, the fourth note was a pivotal event in the trial, in that it would or would not be present on some trials. Support for the interpretation that the fourth note epoch was particularly salient, even when five notes were expected, comes from the topography correlations in this epoch (figure 9C). Specifically, there was a lengthening of the P200 topography, which appears as a vertical band of weak or negative correlations with the dominant N100-like topography at ~1700 ms along the diagonal. The P200 topography during the 4th heard note is significantly correlated with the P200 topography to the first heard note. The effect is seen most clearly in the significance mask plot on the right of figure 9C as a broader column of white separating the topography correlations among the individual heard notes.

Discussion

Experiment 2 replicated the findings from Experiment 1 and extended the comparison of potentials emitted under conditions of imagery with potentials following expected and unexpected terminations of note sequences during attentive listening. As in Experiment 1, imagining the continuation of a sequence of notes resulted in an emitted potential. The emitted potential topographies in the imagery condition, both immediately following the last of the heard notes and throughout the imagery portion of the trial were distinct from the topographies in the acoustically identical condition in which subjects expected to hear, and only heard, three notes. The imagery condition did not differ from the 5-heard note condition or the unexpected omission condition with regard to the topography in the N100 window of the fourth event epoch, whereas all three of these conditions differed significantly from the condition in which the sensory input terminated as expected.

The similarity of the topographies during the N100 time window among conditions in which subjects were actively forming expectations or images suggests that the processes of forming a mental image of an auditory event activates those brain regions actively involved in processing the sensory event. Note, however, that comparing between the N100 topography and the topography of the early portions of the emitted potential in response to an imagined event is not the same as saying that imagining an auditory stimulus generates an N100. For instance, a topography very similar to that of the N100 is observed to recur following the P200. Given that the duration of the sounds was 250 ms, the activity between 300 and 500 ms following stimulus onset may reflect an offset response (Hillyard and Picton1978; Picton et al. 1978). Nonetheless, the N100 is considered functionally distinct from the activity between 300 and 500 ms following stimulus onset. In summary, while the topographical similarity of the initial 200 ms of the emitted potential and the N100 to heard notes is used to argue that the imagery task involves activation of the auditory cortex, the emitted potential is not considered to be equivalent to the "true" N100 which is driven by exogenous stimuli (Näätänen and Picton 1987). However, the emitted potential may be related to the top-down modulators of the N100, because it appears in the time window relative to stimulus onset in which comparisons of sensory input and expectancies are believed to take place. Because these comparisons are believed to take place in the auditory cortex (Näätänen 1992; Näätänen and Winkler 1999), the emitted potential may also depend on generators in the auditory cortex.

The most direct evidence for activation of the auditory cortex in musical imagery tasks comes from functional neuroimaging experiments using PET by Zatorre and colleagues (Halpern and Zatorre 1999; Zatorre et al. 1996). In their tasks, imagining familiar melodies activates a network of brain areas, largely in the right hemisphere, including frontal, infero-frontopolar, as well as the auditory cortex along the superior temporal gyrus (STG). Patients with right temporal lobe excisions perform significantly worse on perceptual and imagery tasks involving pitch comparisons between lyrics in familiar melodies than do controls or patients with left-hemisphere lesions (Zatorre and Halpern 1993). Thus, the emitted potential results described here are consistent with the evidence from the functional neuroimaging and neuropsychology literature.

Both the imagery task context as well as the cue-validity judgement context, in which listeners expect to hear a specific note at a specific time but none occurs, generated emitted potentials whose early topographies were not significantly different from each other or from the N100 response to the corresponding heard event. Thus, it is possible, but not yet proven, that the neural mechanisms underlying the formation of auditory images in an "imagery" task may not be so different from the formation of auditory expectancies in other tasks requiring auditory attention. A similar equivalence has been proposed by Kosslyn for the case of visual imagery and visual perception (Kosslyn 1994, p. 287): "images are formed by the same processes that allow one to anticipate what one would see if a particular object or scene were present." Auditory images/expectancies appear to also activate sensory cortices as has been demonstrated for the visual system (Kosslyn et al. 1993; Le Bihan et al. 1993; Mellet et al. 1998). Carroll-Phelan and Hampson (1996) have argued for an approach to studying musical imagery with the main hypothesis that the cognitive sub-systems underlying the perception of melodies are the same as those underlying imagining of melodies.

General Discussion

Emitted potentials revisited

The results of the present pair of experiments indicate that emitted potentials are not a simple brain response elicited by only a single set of task characteristics. Rather, the presence and morphology (sequence of topographical states) of an emitted potential depends strongly on specific task requirements. While attention appears to be necessary to elicit an emitted potential, it may not be sufficient. The results of the second experiment suggest that in order to elicit an emitted response, it is necessary to form a mental image of an auditory event even if it is known that the event will not be heard. In other words, the emitted potential can be generated independently of an expectation that the event will actually occur.

Although emitted potentials have been identified in the past almost exclusively with expectancy violations and the P300 component, the appearance of emitted potentials in a mental imagery task should not discount the possibility that the emitted potentials actually reflect processes involved in mental imagery. Furthermore, the results from the present experiments raise the question of whether the explicit formation of auditory images in an imagery task is similar to the formation of auditory images described in countless studies on the PN and Nd components which index auditory attention during target detection.

Expectancy formation unmasked by unexpected omissions

Given the emphasis of earlier emitted potential studies on establishing the equivalence of the emitted-P300 and the P300 (Besson et al. 1997; Ruchkin and Sutton 1978; Ruchkin et al. 1980; Ruchkin et al. 1975), the earlier components of emitted potentials were not examined in sufficient detail to rule out involvement of neural generators associated with sensory processing. The data from the present experiments suggest that the early components of emitted potentials to unexpected omissions indeed activate neural substrates involved in processing incoming auditory information. This interpretation is supported by the extensive bodies of literature on the modulation of activity in the auditory cortex between 50 and 250 ms following stimulus onset by endogenous, top-down, processes (for a review see Näätänen 1992). Both the N100 and the MMN fall within this time window. The MMN has been interpreted recently as the stage at which "representations" of auditory objects are extracted from afferent sensory activation and are accessible to top-down influences, such as built up templates or expectations (Näätänen and Winkler 1999). Similarly, the N100 is considered to consist of both exogenous and endogenous factors (Näätänen and Picton 1987) and is believed to arise in the auditory cortex (Liegeois-Chauvel et al. 1994; Pantev et al. 1995; Verkindt et al. 1995). Failing to present an expected stimulus may therefore unmask the portion of the normal auditory evoked response which is due to top-down activation. If the locus of interaction of bottom-up and top-down information is within the auditory cortex, as has been proposed for both the N100 and MMN, it is reasonable to expect that top-down activations may also show an activation pattern that is characteristic of the N100.

In this view, the emitted potentials described in this study may be related to the "negative difference" (Nd) or "processing negativity" (PN) components of ERP waveforms. The Nd and PN appear as a negative displacement of the ERP when subjects attend to a particular feature of an auditory stimulus in the context of a target detection task, compared to when they passively listen to the same stimuli. Short-latency forms of the Nd and PN are believed to arise in the auditory cortex whereas longer-latency components of the PN may arise in the frontal lobes (Giard et al. 1988; for reviews of this extensive literature see Näätänen 1982; Näätänen 1990; Näätänen 1992). The topography of these components is consistent with that observed for the emitted potentials in the present study. The PN is believed to result from a comparison process between an "attentional trace" and an incoming sensory stimulus within the auditory cortex. The "attentional trace" is defined as a voluntary maintenance (involving both the frontal lobes and auditory cortex) of an expected sensory representation, i.e. mental image of the target to be detected, that has been built up over multiple presentations of the target (Näätänen 1982; Näätänen 1990). Thus, the cognitive processes underlying the "attentional trace" and voluntarily formed images may share the same functional neural architecture.

Expectations for the presence vs the absence of an event

It should be pointed out that expectations can exist both for the presence or the absence of an event. Thus, when subjects were told to only expect three or five notes, they presumably actively expected to not hear a fourth or sixth note, respectively. Indeed, an unexpected fourth note elicited a strong P300 (figure 8F, Row D). The expectation that a sequence would end after three notes resulted in a topographical pattern for 500 ms in the silent period that followed which was highly correlated with the P200 to the first note in the sequence and uncorrelated or weakly negatively correlated with the N100-like topography which dominated most of the heard note epochs (figure 9B). Quite surprisingly, the duration of this effect matched the SOA of the preceding notes! A similar effect, but of shorter duration was observed following the last of five expected notes. Given that subjects were instructed to not continue imagining when only three notes were predicted, this effect might reflect active inhibition of mental image formation. Further data are needed to fully address this issue.

Absence of the N100-like topography to subsequent imagined notes

One problem in arguing that the emitted potential observed in the imagery condition following sensory input reflects auditory image formation, is the absence of distinct modulations around the presumptive onsets of the remaining imagined notes. Before discussing potential reasons for this observation, it must be emphasized, however, that the lack of N100-like emitted potentials to subsequent imagined notes does not eliminate the fact that the emitted potential to the first note in the imagery condition is task dependent. Nor does it diminish the importance of the significantly autocorrelated state during the remainder of the imagery trial which disappears as imagery ends and the subject waits for the auditory probe. The topographical state during imagery is also significantly correlated with responses to the second and third heard notes (figure 9A). Thus, even though the focal amplitude modulations of an N100-like topography disappear, other indicators of sustained image formation remain.

One explanation for the absence of emitted potentials beyond the first one is rooted in the limitations of the ERP methodology. In Experiment 1, subjects displayed a significant degree of variability in the time at which they pressed the key as an indication of their expectancy for when the last imagined note should have occurred (112 ms s.d. in the 3I condition; 172 ms s.d. in the 5I condition). If the timing of subjects' mental imagery varied like their key presses, the temporal variability in any emitted potentials to imagined notes from trial to trial would have the effect of attenuating, broadening, and/or canceling the N100-like peaks in the averaged ERP. Originally it was thought that the key-press data might serve as a means of aligning individual trials prior to averaging. Several problems precluded this approach. When this technique was applied to the data from the AH condition, the variability in response times across trials altered the earlier auditory EPs beyond recognition (data not shown). In the imagery conditions, the problem of inherent variability in response timing was compounded by the inability to predict the onset of imagining each note. For example, on some trials, subjects might imagine one note a bit too early and another too late. On another trial they might imagine every note a bit late. Thus, it would be impossible to attribute differences in the latency-adjusted averages at each event position to a true difference between brain activations in listen and imagine conditions, or, instead, to an inability to properly align single-trial evoked responses.

Alternatively, the difference in emitted responses across the sequence of imagined notes may result from changes in other factors, such as the state of ongoing activity in the auditory cortex. Specifically, the emitted response during the first imagined note may reflect an interaction with sensory memory which is necessarily active following the third note. Auditory sensory memory appears to have a short component which lasts several hundred milliseconds from stimulus onset, and a longer component lasting 10-20 s (Cowan 1995). The interaction between a mentally formed image and the short form of sensory memory would be unique to the first imagined note. Interactions between the longer form of sensory memory and mentally formed images might be expected throughout the remainder of the imagery period, though it is uncertain whether the interaction of a mental image with sensory memory during the first imagined note would disrupt the sensory memory in such a way that it would no longer interact in the same way with subsequent images. In other words, would the patterns of neural activity maintaining the sensory memory remain the same in light of an interaction with a top-down expectancy/imagery process?

The mechanisms by which mentally formed images interact with sensory memory cannot be answered given the present data. Nonetheless, the present data illustrate that a variety of emitted potentials can be elicited in a task-dependent manner, and they form a basis for further studies of the relationship between mentally formed images and expectancies, and their interaction with sensory input.

Figure Captions

Figure 3. Temporal compression of scalp topography data (A–C) and ERP topographies and their associated response time distributions of key-presses made synchronously with the onset of the last heard or imagined note (D–F). A–C) ERP data from the same epoch are displayed on three different time-scales as topographical maps. Each circle represents a view down on top of the head. The nose would be at the top of each circle. The dashed lines between subplots bracket corresponding data segments. A) Each map is the average of topographies recorded at two adjacent samples. B) Each map is the average of 8 samples. C) Each map is the average 16 samples. Blue and red indicate negative and positive voltage values, respectively. Note that the topographical patterns that persist over extended periods of time in A are preserved in B and C. D–F) In each panel, histograms of all responses from all subjects are displayed above the grand average topographical maps. The color scale and orientation of the topographical maps is the same as in A-C. Each topographical map is the average of six time points (48 ms). D) Subjects heard 8 notes and were required to press the key synchronously with the onset of the final note. E) Subjects pressed a key at the time when they expected that the third imagined note would have occurred. F) Subjects pressed a key at the time when they expected that the fifth imagined note would have occurred.

Figure 8. Grand-average Electrode X Time X Voltage (ETV) matrices showing the spatio-temporal pattern across the entire trial in each of the five experimental conditions (A-E) and spatio-temporal patterns of activity during the fourth note epoch displayed as conventional topographical maps (F). As in earlier figures, voltage is encoded by the color scale. A-E) The pattern of voltage values across all electrodes (topography) at any single time point is represented as a column in the matrix. The sequence of notes began at 0 ms. In A, B, C, the acoustical conditions are identical. The last heard note lasted from 1000 ms until 1250 ms. The "onset" of the first imagined note (A) was at 1500 ms. B) Three notes were expected, three notes were heard. C) Five notes were expected, but only three were heard. D) Three notes were expected, but five were heard. E) Five notes were expected, five notes were heard. In A, an acoustic probe was presented at 4500 ms. In B-E, a visual probe was presented at 3500 ms. The color bar in E applies to A and B also. The color bar in D applies to C also. Approximately 100 trials/subject contributed to the image in A, 180 trials/subject to the images in B and E, and 20 trials/subject to the images in C and D. The average auditory evoked potential topographies recorded in response to the first note of the melody are shown in the insets above panel A. The N100 shows the typical negative focus at the vertex and positive ring at more inferior sites on the scalp. The P200 to the first note is characterized by a centro-frontal positivity that is somewhat right-lateralized. The average topography from 100 ms following the offset of the note until the onset of the next note (N1R) resembles that of the N100, though it is not as pronounced. F) The letters in parentheses associate each row of topographical maps with panels A–E shown above. The maps show the activation patterns in a time window encompassing the 4th event epoch (1500 – 2000 ms). In the case of the conditions shown in the top three rows, there was no acoustic input during this time window; only the task conditions differed. In the bottom two rows, a note was presented at 1500 ms.

Figure 9. Spatial correlation matrices as a function of task condition. The matrices on the left show how strongly the topography observed at any instant during the trial was correlated with topographies recorded at other times during the trial. Onsets of notes are marked by the tips of the black triangles. Predicted onsets of imagined notes are marked by gray triangles. Locate a time point of interest along the diagonal. The column of values above the point on the diagonal shows the strength of correlations with topographies preceding the event of interest, while points along the row to the right of the location on the diagonal indicate the strength of correlation with topographies following the event of interest. Periods during which there is very little change in the topography show up as triangular areas emanating from the diagonal. See figure 8 for comparison. Black points in the matrices on the right show the correlations which were statistically significant following Bonferroni correction for multiple comparisons.

References

Besson, M. and Faïta, F. An event-related potential (ERP) study of musical expectancy: Comparison of musicians with nonmusicians. JEP: Human Percept. and Perf., 1995, 21: 1278-1296.

Besson, M., Faita, F., Czternasty, C. and Kutas, M. What's in a pause: Event-related potential analysis of temporal disruptions in written and spoken sentences. Biological Psychology, 1997, 46: 3-23.

CarrollPhelan, B. and Hampson, P.J. Multiple components of the perception of musical sequences: A cognitive neuroscience analysis and some implications for auditory imagery. Music Perception, 1996, 13: 517-561.

Cowan, N. Attention and Memory. Oxford University Press, New York, 1995.

Edelman, G.M. The remembered present: a biological theory of consciousness. Basic Books, New York, 1989

Galan, L., Biscay, R., Rodriguez, J.L., Perez-Abalo, M.C. and Rodriguez, R. Testing topographic differences between event related brain potentials by using non-parametric combinations of permutation tests [published erratum appears in Electroencephalogr. Clin. Neurophysiol., 1998 Nov; 107(5): 380-381]. Electroencephalogr. Clin. Neurophysiol., 1997, 102: 240-247.

Giard, M.H., Perrin, F., Pernier, J. and Peronnet, F. Several attention-related wave forms in auditory areas: a topographic study. Electroencephalogr. Clin. Neurophysiol., 1988, 69: 371-384.

Grossberg, S. How does a brain build a cognitive code? Psychol. Rev., 1980, 87: 1-51.

Halpern, A.R. and Zatorre, R.J. When that tune runs through your head: a PET investigation of auditory imagery for familiar melodies. Cereb. Cortex, 1999, 9: 697-704.

Hays, W.L. Statistics. (4th ed.). Harcourt Brace Jovanovich, Orlando, 1988.

Hillyard, S.A. and Picton, T.W. On and off components in the auditory evoked potential. Percept. Psychophys., 1978, 24: 391-398.

Janata, P. ERP measures assay the degree of expectancy violation of harmonic contexts in music. J. Cogn. Neurosci., 1995, 7: 153-164.

Joutsiniemi, S.L. and Hari, R. Omissions of auditory-stimuli may activate frontal-cortex. European Journal of Neuroscience, 1989, 1: 524-528.

Kosslyn, S.M. Image and Brain. The MIT Press, Cambridge, 1994.

Kosslyn, S.M., Alpert, N.M., Thompson, W.L., Maljkovic, V., Weise, S.B., Chabris, C.F., Hamilton, S.E., Rauch, S.L. and Buonanno, F.S. Visual mental imagery activates topographically organized visual cortex: PET investigations. Journal of Cognitive Neuroscience, 1993, 5: 263-287.

Le Bihan, D., Turner, R., Zeffiro, T.A., Cuenod, C.A., Jezzard, P. and Bonnerot, V. Activation of human primary visual cortex during visual recall: a magnetic resonance imaging study. Proc. Natl. Acad. Sci. USA, 1993, 90: 11802-11805.

Liegeois-Chauvel, C., Musolino, A., Badier, J.M., Marquis, P. and Chauvel, P. Evoked potentials recorded from the auditory cortex in man: evaluation and topography of the middle latency components. Electroencephalogr. Clin. Neurophysiol., 1994, 92: 204-214.

Mellet, E., Petit, L., Mazoyer, B., Denis, M. and Tzourio, N. Reopening the mental imagery debate: lessons from functional anatomy. Neuroimage, 1998, 8: 129-139.

Näätänen, R. Processing negativity: an evoked-potential reflection of selective attention. Psychol. Bull., 1982, 92: 605-640.

Näätänen, R. The role of attention in auditory information processing as revealed by event-related potentials and other brain measures of cognitive function. Behav. Brain Sci., 1990, 13: 201-288.

Näätänen, R. Attention and Brain Function. Lawrence Erlbaum Associates, Hillsdale, 1992

Näätänen, R., and Picton, T. The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. Psychophysiology, 1987, 24: 375-425.

Näätänen, R. and Winkler, I. The concept of auditory stimulus representation in cognitive neuroscience. Psychological Bulletin, 1999, 125: 826-859.

Pantev, C., Bertrand, O., Eulitz, C., Verkindt, C., Hampson, S., Schuierer, G. and Elbert, T. Specific tonotopic organizations of different areas of the human auditory cortex revealed by simultaneous magnetic and electric recordings. Electroencephalogr. Clin. Neurophysiol., 1995, 94: 26-40.

Picton, T.W., Woods, D.L. and Proulx, G.B. Human auditory sustained potentials. I. The nature of the response. Electroencephalogr. Clin. Neurophysiol., 1978, 45: 186-197.

Raij, T., McEvoy, L., Makela, J.P. and Hari, R. Human auditory cortex is activated by omissions of auditory stimuli. Brain Research, 1997, 745: 134-143.

Ruchkin, D.S. and Sutton, S. Emitted P300 potentials and temporal uncertainty. Electroencephalography and Clinical Neurophysiology, 1978, 45: 268-277.

Ruchkin, D.S., Sutton, S. and Stega, M. Emitted P300 and slow-wave event-related potentials in guessing and detection tasks. Electroencephalography and Clinical Neurophysiology, 1980, 49: 1-14.

Ruchkin, D.S., Sutton, S. and Tueting, P. Emitted and evoked P300 potentials and variation in stimulus probability. Psychophysiology, 1975, 12: 591-595.

Samuel, A. Phoneme restoration. Language and Cognitive Processes, 1996, 11: 647-653.

Samuel, A.G. Phonemic restoration: insights from a new methodology. J. Exp. Psychol. Gen., 1981a, 110: 474-494.

Samuel, A.G. The role of bottom-up confirmation in the phonemic restoration illusion. J. Exp. Psychol. Hum. Percept. Perform., 1981b, 7: 1124-1131.

Simson, R., Vaughan, H.G. and Ritter, W. The scalp topography of potentials associated with missing visual or auditory stimuli. Electroencephalogr. Clin. Neurophysiol., 1976, 40: 33-42.

Srinivasan, R., Nunez, P.L., Tucker, D.M., Silberstein, R.B. and Cadusch, P.J. Spatial sampling and filtering of EEG with spline laplacians to estimate cortical potentials. Brain Topogr., 1996, 8: 355-366.

Sutton, S., Tueting, P., Zubin, J. and John, E.R. Information delivery and the sensory evoked potential. Science, 1967, 155: 1436-1439.

Tervaniemi, M., Saarinen, J., Paavilainen, P., Danilova, N. and Näätänen, R. Temporal integration of auditory information in sensory memory as reflected by the mismatch negativity. Biol. Psychol., 1994, 38: 157-167.

Tucker, D.M. Spatial sampling of head electrical fields: the geodesic sensor net. Electroencephalogr. Clin. Neurophysiol., 1993, 87: 154-163.

Verkindt, C., Bertrand, O., Perrin, F., Echallier, J.F. and Pernier, J. Tonotopic organization of the human auditory cortex: N100 topography and multiple dipole model analysis. Electroencephalogr. Clin. Neurophysiol., 1995, 96: 143-156.

Warren, R.M. Perceptual restoration of missing speech sounds. Science, 1970, 167: 392-393.

Warren, R.M. Perceptual restoration of obliterated sounds. Psychol. Bull., 1984, 96: 371-383.

Weinberg, H., Walter, W.G., Cooper, R. and Aldridge, V.J. Emitted cerebral events. Electroencephalogr. Clin. Neurophysiol., 1974, 36: 449-456.

Weinberg, H., Walter, W.G. and Crow, H.J. Intracerebral events in humans related to real and imaginary stimuli. Electroencephalogr. Clin. Neurophysiol., 1970, 29: 1-9.

Woldorff, M.G. and Hillyard, S.A. Modulation of early auditory processing during selective listening to rapidly presented tones. Electroencephalogr. Clin. Neurophysiol., 1991, 79: 170-191.

Zatorre, R.J. and Halpern, A.R. Effect of unilateral temporal-lobe excision on perception and imagery of songs. Neuropsychologia, 1993, 31: 221-232.

Zatorre, R.J., Halpern, A.R., Perry, D.W., Meyer, E. and Evans, A.C. Hearing in the mind's ear: A PET investigation of musical imagery and perception. J. Cogn. Neurosci., 1996, 8: 29-46.