

Research report

Cortical processing of complex tone stimuli: mismatch negativity at the end of a period of rapid pitch modulation

Maria Vaz Pato, Stephen J. Jones *

Department of Clinical Neurophysiology, The National Hospital for Neurology and Neurosurgery, Queen Square, London WC1N 3BG, UK

Accepted 7 July 1998

Abstract

In this study, synthesised instrumental tones were used to examine human auditory cortical processes engaged at the end of a period of rapid pitch modulation. It was previously [S.J. Jones, O. Longe, M. Vaz Pato, Auditory evoked potentials to abrupt pitch and timbre change of complex tones: electrophysiological evidence of ‘streaming’?, *Electroencephalogr. Clin. Neurophysiol.*, 108 (1998) 131–142] suggested that the ‘change-N1’ produced by infrequent changes in pitch or timbre of a continuous complex tone represents the activity of a neuronal population topographically distinct from that responsible for the ‘onset-N1’ at the beginning of the tone. In the present study a superficially similar negativity was produced when the tone came to rest on a steady pitch after a period of rapid (8–16 changes/s) modulation; its scalp maximum was anterior to that of the two previously identified potentials but similar to that of the mismatch negativity elicited by discontinuous tones. By varying the modulation rate the latency was shown to be relatively constant with respect to the time the next pitch change was expected but failed to occur. The largest responses averaging c. 7 μ V were evoked at the end of modulation sequences which were both rhythmic and repetitive, but a potential was still produced when there was no rhythmic pattern or repetition of individual notes. This response to non-occurrence of an expected but not necessarily specified change implies an automatic process for comparing the incoming sound with an extrapolated template of the preceding pattern in which timing as well as pitch information is accurately represented. We suggest this technique offers a robust method for eliciting the mismatch negativity, which may extend the opportunities for electrophysiological investigation of higher auditory processes. © 1999 Elsevier Science B.V. All rights reserved.

Keywords: Audition; Auditory evoked potential; Auditory cortex; Complex tone; Echoic memory; Mismatch negativity; Music perception

1. Introduction

Long-latency human auditory evoked potentials (AEPs) have been studied for more than 25 years and it is a decade since the extensive literature on the N1 component peaking approximately 100 ms after the onset of a tone was reviewed [32]. In recent years much attention has been paid to one of the contributory potentials identified by these authors, the ‘mismatch negativity’ (MMN). The latter is evoked by infrequent tones differing perceptibly (in terms of their pitch, loudness, duration, source location, etc.) from a regular sequence [e.g., Refs. [2,31,41]] and is believed to be generated as a consequence of automatic (extra-attentional) comparison of the incoming sound with an image retained for a few sounds in the ‘echoic memory’

[35] or ‘long auditory store’ [11]. The MMN therefore, offers a potentially powerful tool for investigating auditory processes dependent on sensory memory, although in practice its low signal/noise ratio has proved to be a significant obstacle.

The brief pure tone bursts usually employed to evoke the MMN are very different from the natural sounds of the environment which tend to be more-or-less continuous, spectrally complex and in a constant state of modulation. Single unit studies of the auditory cortex of animals have generally moved away from the use of pure tones to frequency and amplitude-modulated sounds [12,13,15, 29,30,39,45,51,52]. Some early workers in the AEP field examined the responses to frequency ‘glides’ [4,8,20] and there have been a few more recent studies of responses to pitch change of continuous pure tones [e.g., Refs. [26,27,46]], but the bulk of human research still concentrates on disconnected sounds whose physiological and behavioral relevance is questionable.

* Corresponding author. Fax: +44-71-713-7743; E-mail: s.j.jones@ion.ucl.ac.uk

In a previous study [18] we used synthesised instrumental tones as a model for the kind of sounds occurring naturally, quantifying their pitch and duration in units of western music. The N1 to pitch or timbre change of a continuous complex tone was found to be more posteriorly distributed on the scalp than the superficially similar N1 to onset of the tone from silence, suggesting that the neuronal populations responsible must be at least partially distinct from one another. A later ‘T-complex’ [54,55] was recorded at lateral temporal electrodes; this was consistently larger over the right hemisphere of right-handed subjects and showed the converse laterality in about 25% of left-handers [17]. The differential interfering effect of background tones of similar or different timbre [18] suggested that these responses reflect processing of the whole sound rather than its individual frequency components, and therefore must be generated subsequent to the grouping of frequencies to form auditory objects (‘primitive stream segregation’ [6]). Since generation of these potentials requires no active cooperation from the subjects, they offer a promising tool for investigating higher auditory processes.

In an earlier study [16] a large N1 potential was recorded at the end of periods of regular, repetitive pitch modulation, sufficiently rapid for there to be little or no measurable response to each individual pitch change. This potential could not have been generated by the neurones responsive to pitch change at a lower rate (for the generator to have become dishabituated would imply foreknowledge that this was to be the last change), but rather must reflect a process for comparing the instantaneous sound with the preceding modulation pattern. In the present study we presented synthesised instrumental tones which oscillated between two pitches at various rates, to compare the potential evoked at the end of the sequence with the MMN to discontinuous tones, and with the N1 to pitch change of a steady tone. We also looked at different modulation patterns to try to discover the conditions necessary for a response to be generated at the end of the sequence. The findings suggest a prominent role for the auditory sensory store in processing rapidly modulated tone sequences, and a more robust method for evoking the MMN than the conventional one using discontinuous tones.

2. Material and methods

The stimuli were tones synthesised by a Soundblaster 16 sound card (Creative Labs, Milpitas, CA, USA) controlled by an IBM-compatible PC. The software used to construct and play stimulus sequences was Midisoft Recording Studio (Diamond Computer Systems). The left and right output channels were split in order to use one channel for the stimulus (presented to both ears) and the other to trigger the recording apparatus. The sounds all had a smooth onset and offset with a rise time of approximately 40 ms (established by examination of their wave-

forms; this was also done to establish the precise temporal relationship between stimulus and trigger) and a similar period of overlap between elided tones. The ‘clarinet’ timbre was chosen on account of its unexaggerated onset and subsequently steady intensity and pitch, its spectrum analysed using a fast A/D converter and associated software (Pico Technology, Cambridge, UK) [18]. The baseline frequency was A_4 (440 Hz) and the intensity was 45–50 dB above threshold.

The subjects were 11 normally hearing right handed volunteers, five males and six females aged 26–46 years, who gave their informed consent according to the Declaration of Helsinki. The subjects sat in an armchair and read a book for the duration of each recording which was approximately 1 h. No instructions were given apart from to continue reading and no verbal responses were required.

Up to eight recording electrodes were attached to the head with collodion or paste, on the midline at Fpz, Fz, Cz and Pz and in a coronal plane across the vertex at T4, C4, T3 and C3 locations of the 10–20 system. The reference electrode was on the dorsum of the neck at the base of the skull. The recording bandwidth was less than 3 dB down between 1 Hz and 5 kHz, the average epoch was 320 or 500 ms starting 15 ms before each stimulus and the A/D conversion rate was 800 or 512 Hz. The averaged waveforms were subsequently digitally filtered with a low-pass corner frequency of 150 Hz. Peak latencies and amplitudes (measured from the pre-stimulus baseline) were obtained for each subject and compared between conditions using the Wilcoxon test. Group mean waveforms and difference waveforms were computed for illustration.

In the first experiment potentials were recorded during and at the end of a period of rapid oscillation between two pitches. Responses of eight subjects were obtained to connected tones oscillating between A_4 and B_4 before coming to rest on B_4 . The identical sequence was presented at three different speeds, such that the rate of oscillation was 8, 12 or 16 changes/s with durations of 5, 3.3 or 2.5 s, respectively, and the duration of the steady note was 0.5, 0.33 or 0.25 s. The triggers were delivered to the averager 15 ms before the 3rd, 7th, 11th, 15th and 19th repetitions of the note A_4 (these responses being averaged together) and at the end of the oscillatory period, 15 ms before the next presentation of A_4 should have occurred. Responses were averaged to 220 oscillatory changes and 45 ‘non-events’ at the end of the sequence, and three repetitions performed for each condition.

The second experiment was designed to record the MMN to brief complex tones, to compare it with the responses obtained when the same tones were connected, and with those occurring at the end of pitch oscillation in the first experiment. Responses of eight subjects were recorded to the notes A_4 and B_4 100 ms in duration (including rise/fall time) with silent gaps of 250 ms. The deviant B_4 tones occurred regularly after every five presentations of the standard A_4 . In each run the responses to

100 deviants and 300 standards (the three preceding each deviant) were averaged and three repetitions made. In the second condition the duration of the tones was extended so that the sound became continuous, the ‘standard’ being present for 1750 ms and the ‘deviant’ for 350 ms. The intensity was reduced by about 5 dB in order for the loudness to be subjectively equal to that of the discontinuous tones. One hundred responses were averaged to each transition and three repetitions made.

In the third experiment, responses of five subjects were recorded at the beginning and end of 16/s sequences of connected tones to try to establish the conditions necessary for a response to be generated when a steady pitch was

resumed. Each sequence was 2 s in duration and finished on the note B_4 which was held constant for a further 2 s. The seven sequences (whose musical notation is shown in Figs. 5 and 6) are mentioned below.

(1) Oscillation between A_4 and B_4 as in the first experiment.

(2) A repeating cell of 3 adjacent notes A_4 , B_4 and C_5 .

(3) A repeating cell of 5 adjacent notes in a rising pattern G_4 , A_4 , B_4 , C_5 and D_5 (the constraint of having a total of 32 tones in each sequence necessitated omission of the first G_4).

(4) The same 5 notes as in condition 3 but in pseudo-random order such that each occurred with the same

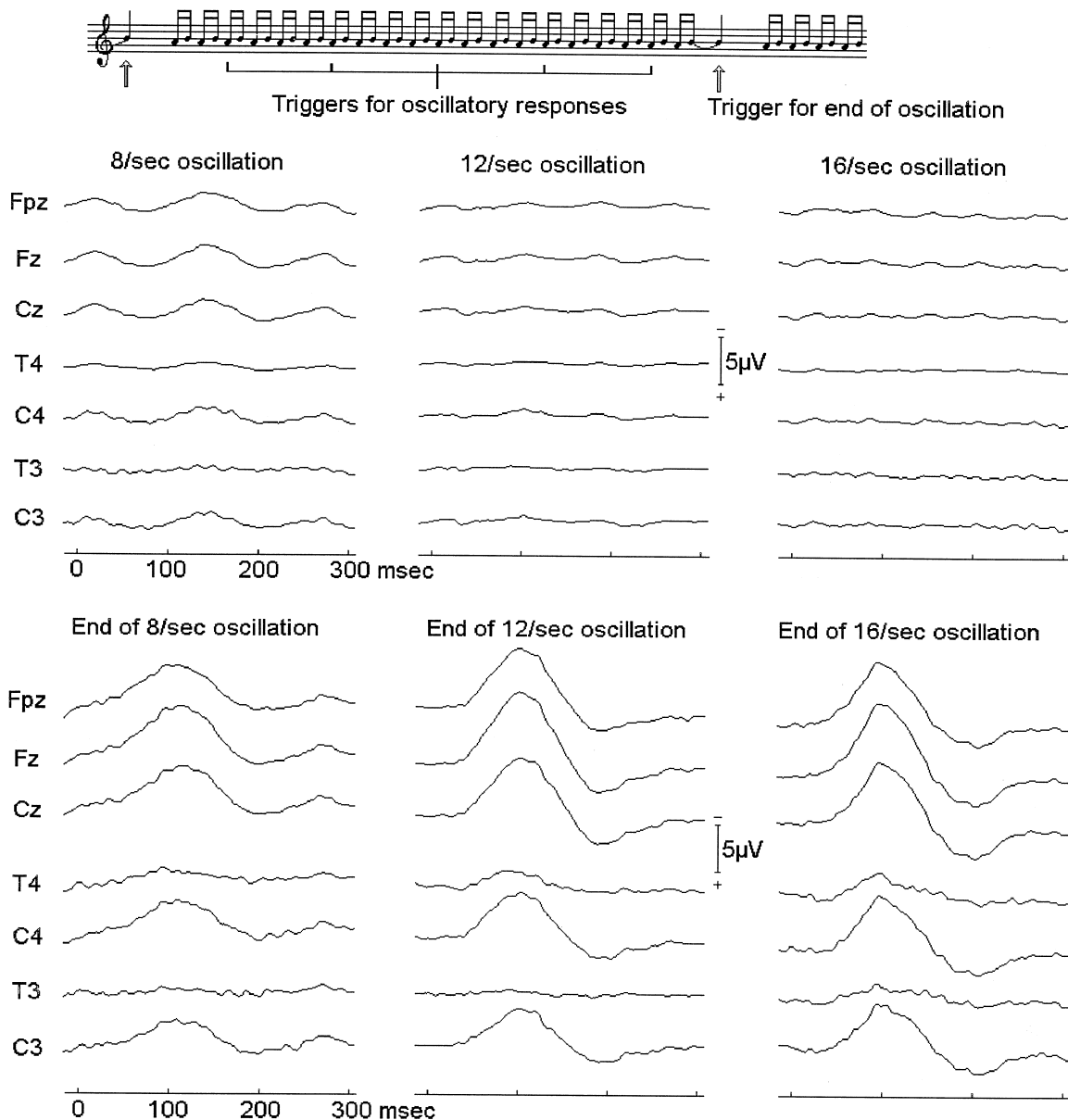


Fig. 1. Experiment 1. Group mean responses of eight subjects to connected tones oscillating in pitch between A_4 and B_4 at three different rates. Whereas, the responses to individual oscillations diminished as the rate increased, the negative peak at the end of the oscillation was larger at the faster rates; its latency of *c.* 100 ms was relatively constant with respect to the expected time of occurrence of the next pitch change.

probability but no pattern of 3 or more notes was repeated (this had the effect of eliminating the rhythmic pattern of condition 3).

(5) A pseudorandom sequence of 32 notes in which every note of the chromatic scale between G_3 and D_6 occurred only once.

(6) A down/up/down sweep of 32 notes in which every note between G_3 and D_6 occurred only once. The sweep commenced on A_4 , descended to G_3 in intervals of a tone, ascended from $G\#_3$ to D_6 also in intervals of a tone and finally descended from $C\#_6$ to B_4 . This sequence was therefore non-repetitive and non-rhythmic but had a distinct overall pattern.

(7) A sequence comprising all 32 notes between G_3 and D_6 presented in 8 groups of 4 descending notes of the chromatic scale. The groups were widely spaced starting on Bb_4 , Bb_5 , D_4 , D_6 , $F\#_4$, $F\#_5$, Bb_3 and D_5 . This introduced a rhythmic pattern without repetition of any individual note.

3. Results

3.1. Experiment 1

Potentials associated with oscillation between the notes A_4 and B_4 were consistently present at the slowest rate of 8 changes/s (negativity peaking about 150 ms after each change) but of very low amplitude or absent at the two faster rates (Fig. 1). At the end of the oscillatory period, in every subject a prominent negative/positive response was recorded in which the mean negative peak latency was relatively constant (104–114 ms) with respect to the expected time of the next pitch change (Table 1). Latencies with respect to the last actual pitch change were approximately 240, 190 and 175 ms. The negativity at the end of the fastest modulation rate was of largest amplitude in spite of the fact that the cycle period was shortest for this condition. As compared with the N1 recorded to tone onset or infrequent (once every 2 s) spectral changes [18], the negativity at the end of the oscillation sequence appeared to be more anteriorly distributed on the scalp, being maximal at Fz or Fpz, whereas, the other potentials were both

Table 1

Experiment 1: mean, standard deviation and range (eight subjects) of negative peak amplitude and latency values (Fpz) at the end of a period of oscillation between 2 pitches

Oscillation rate (Hz)	Amplitude (μ V)		Latency (ms)	
	Mean \pm S.D.	Range	Mean \pm S.D.	Range
8	6.4 \pm 2.5	3.0–9.7	114 \pm 15	96–137
12	8.2 \pm 3.0	2.3–12.3	104 \pm 9	93–119
16	9.2 \pm 4.3	4.7–18.0	104 \pm 9	91–113

Latency values are with respect to the first expected pitch change which failed to occur.

on average largest at Cz. In ANOVA of the six subjects who participated in both studies the interaction between stimulus condition (onset, spectral change and end of oscillation) and electrode location in the midline sagittal plane did not reach statistical significance, but the ratio of negative peak amplitudes at Fpz to Cz was significantly larger for the responses at the end of pitch oscillation as compared with onset or spectral change ($p < 0.05$), suggesting the end-of-oscillation negativity to be the most anteriorly distributed.

3.2. Experiment 2

In the MMN study, responses to the standard A_4 tones presented five times in every cycle consisted of a P1 at 60–70 ms followed by a shallow N1 and P2. Responses to the deviant B_4 presented once in every cycle consisted of a similar P1 and P2 but a larger N1. In the deviant–standard difference waveforms an MMN peaking between 110 and 140 ms was evident in seven out of eight subjects, being largest in amplitude at Fz (mean 3.0 μ V) and on average slightly larger at Fpz than Cz (Fig. 2). The MMN was therefore similar in distribution to the negativity at the end of pitch oscillation although lower in amplitude. Fig. 3 shows the responses of one subject who had a very indistinct MMN but a clear negativity at the end of 16/s pitch oscillation.

When the tones were extended to eliminate the silent intervals, the responses were clearly altered in morphology and distribution (Fig. 2). A P1/N1/P2 complex was now maximal at Cz and the P2 was particularly prominent, being more than twice the amplitude of the N1. The response to the ‘standard to deviant’ (long to short tone) pitch change was markedly larger in amplitude than the converse response but similar in morphology and distribution, as was the corresponding difference waveform. The Fpz/Cz amplitude ratio of the latter was significantly lower than that of the MMN in the difference between responses to deviant and standard discontinuous tones ($p < 0.05$).

Fig. 4 shows the sagittal profile of mean negative peak amplitudes for the 6 subjects who were common to Experiments 1 and 2 and also the previous study [18]. Responses to infrequent spectral changes of continuous tones were maximal at Cz and more than 50% smaller at Fpz, whereas responses at the end of pitch oscillation (Experiment 1) and the MMN (Experiment 2) were maximal at Fz and less than 20% smaller at Fpz. Responses to the onset of discontinuous tones had an intermediate distribution.

3.3. Experiment 3

Starting from a steady tone, the responses evoked at the beginning of the 16/s sequences of changes were similar to those to infrequently occurring pitch changes, consisting

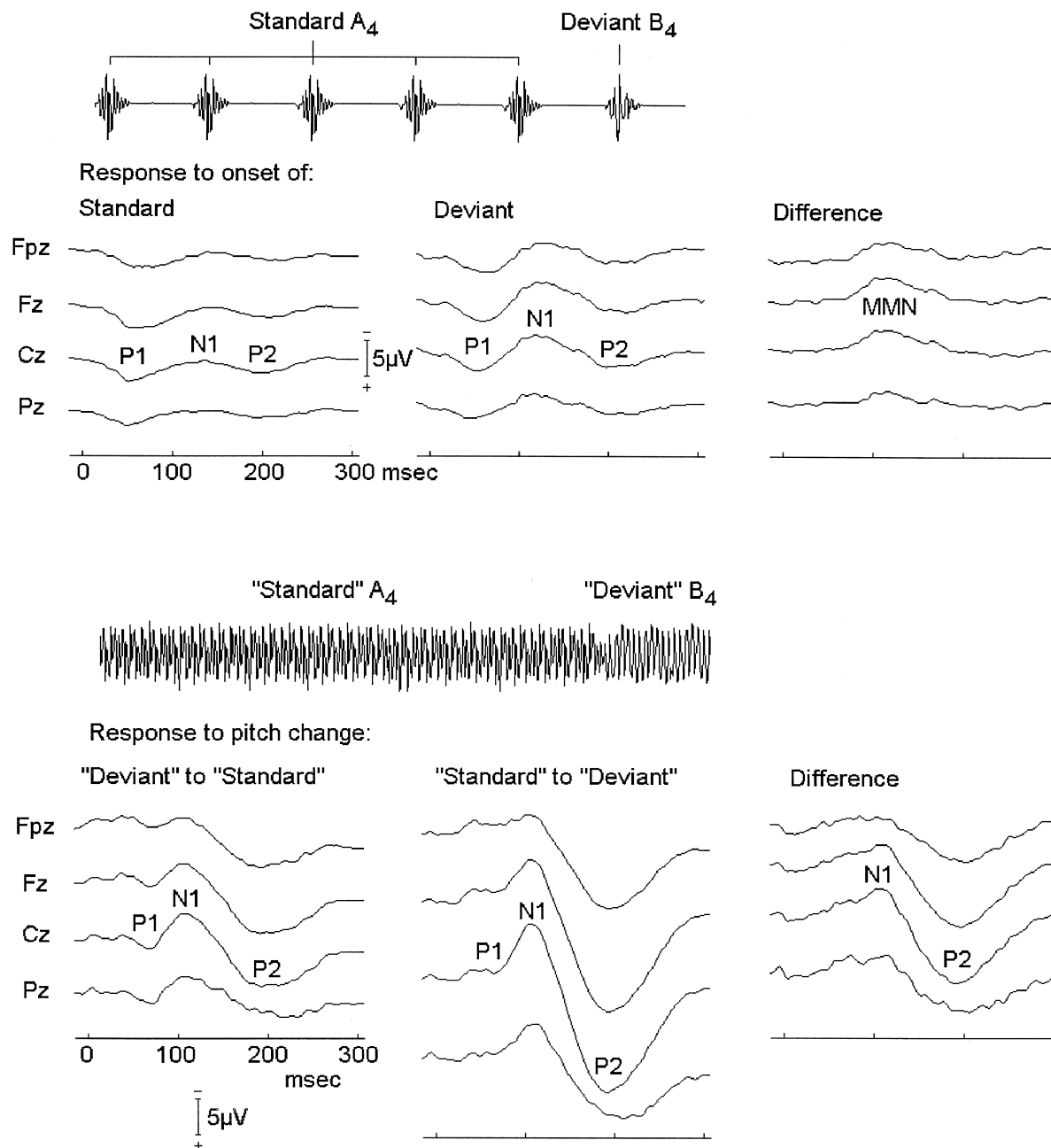


Fig. 2. Experiment 2. The upper panel shows the group mean responses of eight subjects to standard A_4 tones (duration 100 ms, gaps 250 ms) with deviant B_4 tones occurring regularly after every five standards, plus the difference waveforms showing a mismatch negativity maximal at Fz. The lower panel shows group mean responses to the same tones extended to the point that the gaps were eliminated. In addition to a much larger P2, the N1 was more posteriorly distributed, being maximal at Cz. The response to transition from the longer 'standard' to the shorter 'deviant' tone was of larger amplitude than that to the converse change but the difference waveforms show no evidence of an MMN which would be expected to have a more anterior distribution.

of a small P1, an N1 maximal at Cz and a prominent P2. At the end of each sequence the responses were similar to those of Experiment 1, the negative peak being more anteriorly distributed than at the beginning of the sequence. Whereas the potentials at the beginning of each sequence were similar to one another, those at the end were consistently largest for the regular 2-, 3- or 5-note patterns, the negativity measuring on average about 7 μV

(Table 2). At the end of the pseudorandom (non-rhythmic) 5-note sequence the response was on average slightly lower in amplitude but not significantly so (Fig. 5). Responses at the end of the non-repeating 32-note sequences were smaller and less consistent, but present in every subject except one (with the notes patterned in rhythmic groups) with a mean amplitude of about 4 μV (Fig. 6). No significant differences were observed according to whether

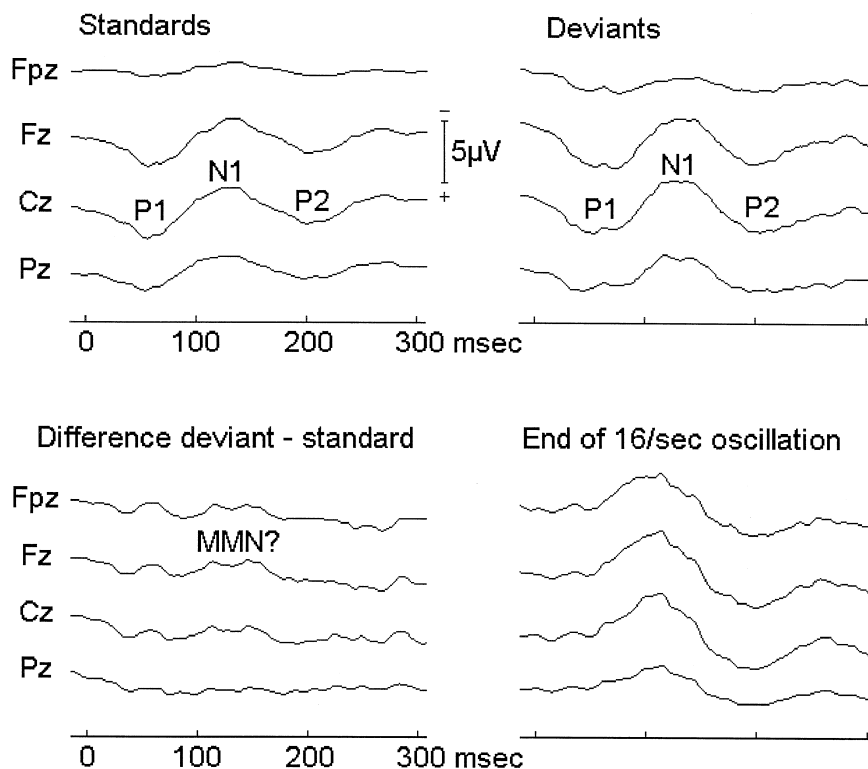


Fig. 3. Responses of one subject to standard and deviant tones of Experiment 2 plus the difference waveforms showing an indistinct MMN; however, at the end of 16/s note sequences of Experiment 1 a large negativity was evoked, of maximal amplitude at Fz.

the 32 notes were presented in pseudorandom order, as a rising and falling sweep or in rhythmic groups of four.

4. Discussion

The main finding of this study is that a negative potential probably equivalent to the mismatch negativity is evoked when a complex tone comes to rest on a steady pitch at the end of a period of rapid modulation. The fact that the generator of the N1 to individual changes of pitch was almost completely refractory at the higher modulation rates indicates that the end of modulation must have been signalled by a different process, by which the immediate sound is compared with a template of the preceding modulation pattern, extrapolated so as to predict the time that the next pitch change should occur. This implies the participation of a sensory memory store with a duration of a few seconds and an internal clock whose precision is in the order of milliseconds, as well as a comparator. Although 'automatic' in so far as it appears not to require the active attention of the subject, this is clearly a more complex process than that necessary to signal simple onset and offset of a tone, or a change in its spectral content. Our pilot studies indicate that similar potentials may be generated when an oscillating sequence ends in silence, but not at the end of a steady tone [18]. Detailed mapping will be necessary to confirm the topographic separation of the

onset-N1, change-N1 and mismatch negativity suggested but not absolutely proven by these studies.

The very extensive literature on long latency AEPs includes rather few studies in which spectrally complex tones [e.g., Refs. [3,48–50,53]] or 'musical' stimuli [e.g., Refs. [5,9,10,38]] were used. Apart from the fact that they were spectrally complex, the discontinuous tones employed in the second experiment were fairly typical of those used to produce an MMN [36], the deviants occurring once every 2.1 s with one-fifth the probability of the standards. For comparability with the condition in which the gaps between the tones were eliminated, the deviants were presented at regular intervals; one study [44] has reported no significant difference in the MMN according to whether the deviants occurred regularly or randomly with the same overall frequency, and this was also our experience (unpublished). The MMN seen in the deviant–standard difference waveform was fairly typical in its amplitude and distribution, measuring about 3 μV at Fz. Its latency of approximately 120 ms was shorter than in most previous studies, but this has been noted by others when complex tones are used (Tervaniemi, personal communication).

It is now generally accepted [e.g., Ref. [14]], that the scalp distribution of the MMN is consistent with origin in the supratemporal plane of one or both hemispheres, the frontal maximum suggesting a locus somewhat anterior to that of the principal generator of the onset-N1. We previ-

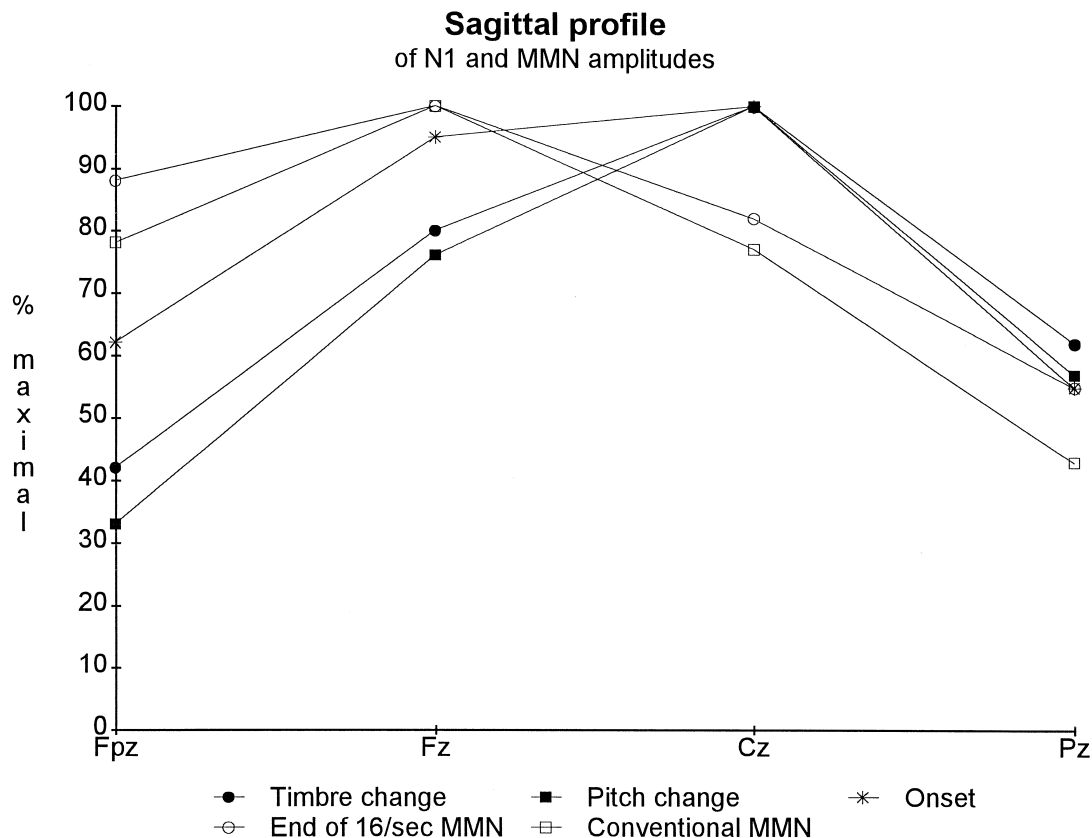


Fig. 4. Profile of negative peak amplitudes in Experiments 1 and 2, plus data from the previous study [18]. Mean amplitudes from the six subjects common to all three experiments is presented. The N1's to timbre change [18] and to pitch change of connected tones were both largest at Cz and very much smaller at Fpz. The mismatch negativity recorded conventionally (Experiment 2) and at the end of 16/s oscillation (Experiment 1) were both largest at Fz and only a little smaller at Fpz. The onset-N1 to disconnected tones [18] was of intermediate distribution.

ously [18], on the other hand, presented evidence that the change-N1 has a more posterior distribution. Our previous study also suggested that a transition from onset-N1 to change-N1 starts to occur when the gaps between discontinuous tones of different spectral composition become less than about 200 ms. In addition to these three topographically distinct potentials consistent with multiple generators on the supratemporal plane, the T-complex which was a

Table 2

Experiment 3: mean, standard deviation and range (five subjects) of negative peak amplitude and latency values (Fz) at the end of 16/s note sequences

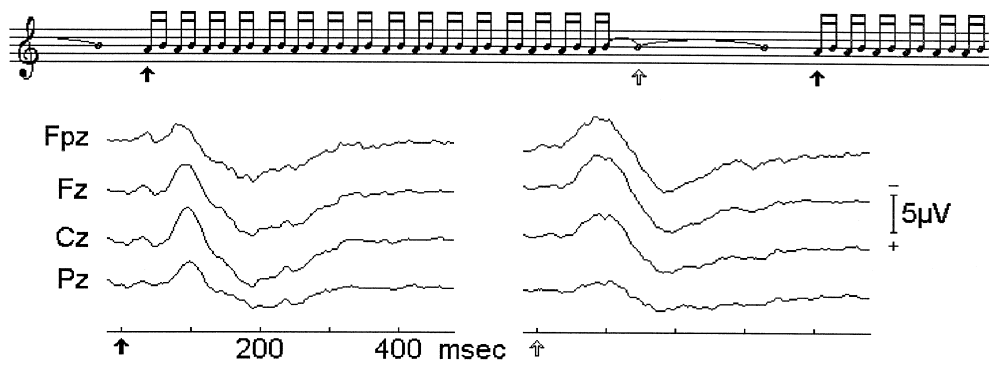
Sequence	Amplitude (μ V)		Latency (ms)	
	Mean \pm S.D.	Range	Mean \pm S.D.	Range
2-notes	6.6 \pm 3.2	3.6–11.6	105 \pm 14	90–126
3-notes	6.7 \pm 2.5	3.8–9.2	106 \pm 11	98–126
5-notes regular	7.7 \pm 3.4	3.6–11.3	117 \pm 12	104–135
5-notes pseudorandom	6.4 \pm 2.7	3.6–9.6	104 \pm 12	96–124
32-notes pseudorandom	3.6 \pm 2.6	0.7–7.0	101 \pm 14	90–122
32-notes up/down sweep	4.7 \pm 2.1	2.4–7.4	102 \pm 21	85–135
32-notes in groups of four	4.2 \pm 2.4	0.0–7.6	109 \pm 4	104–112

Latency values are with respect to the first expected pitch change which failed to occur.

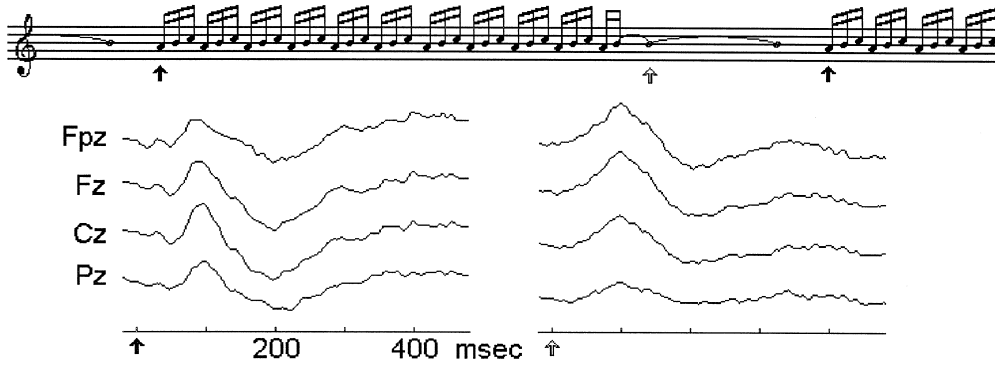
consistent feature of the previous study suggests a fourth generator probably located in the lateral gyri of the temporal lobes. We have already [17] reported the consistently larger amplitude of the T-complex over the right hemisphere of right-handed subjects and the reversed asymmetry seen in about 25% of left-handers. The functional significance of this potential is unknown, but a particular role for the right temporal cortex in the processing of 'musical' sounds has been suggested by PET studies and the effect of clinical lesions [7,42,56,57].

Näätänen et al. [33] examined the effect on the MMN of increasing the overall rate of stimulation and reducing the duration of the silent intervals, to the point that in the limiting case the tones became continuous. A large negativity was elicited when the deviant stimulus followed the standard without a break, but the authors were uncertain whether this was in fact an MMN. In a more recent study of auditory evoked magnetic fields to brief (100 ms) frequency changes in a continuous tone [26], two processes were distinguished contributing to the field at about 100 and 130 ms, respectively in six out of nine subjects. On the basis of source location modelling the generators were identified as those of the onset-N1 and the MMN, respectively. This experiment can be criticised in that only the

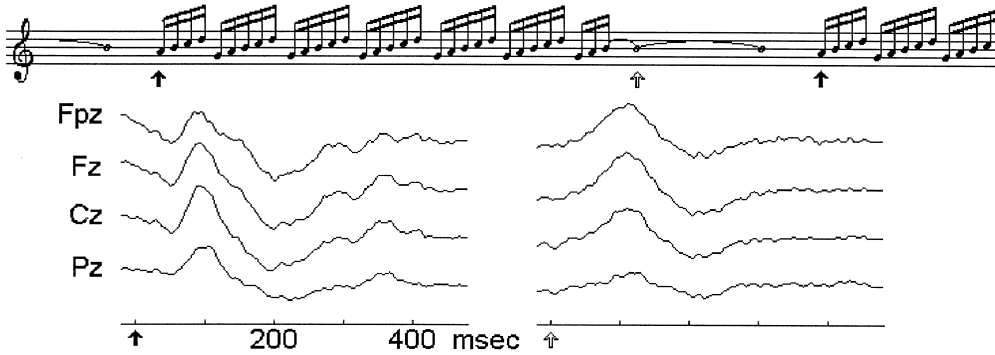
Repeating 2-note pattern



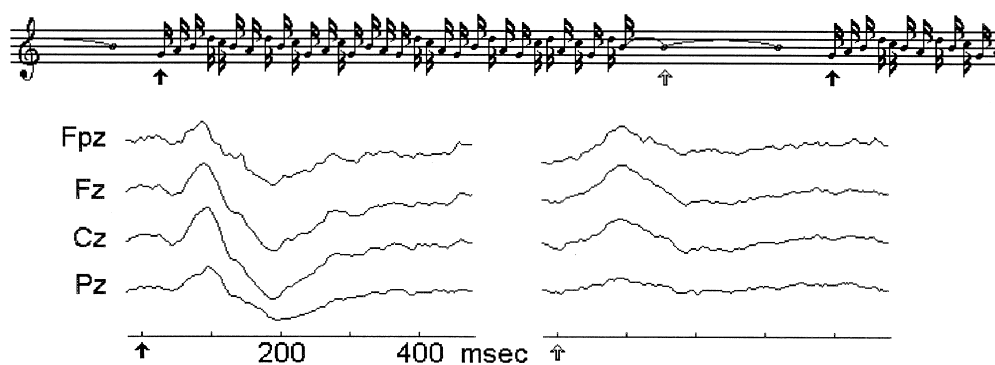
Repeating 3-note pattern



Repeating 5-note pattern



Pseudorandom (non-rhythmic) 5-note pattern



32-note patterns with no repetition of individual notes

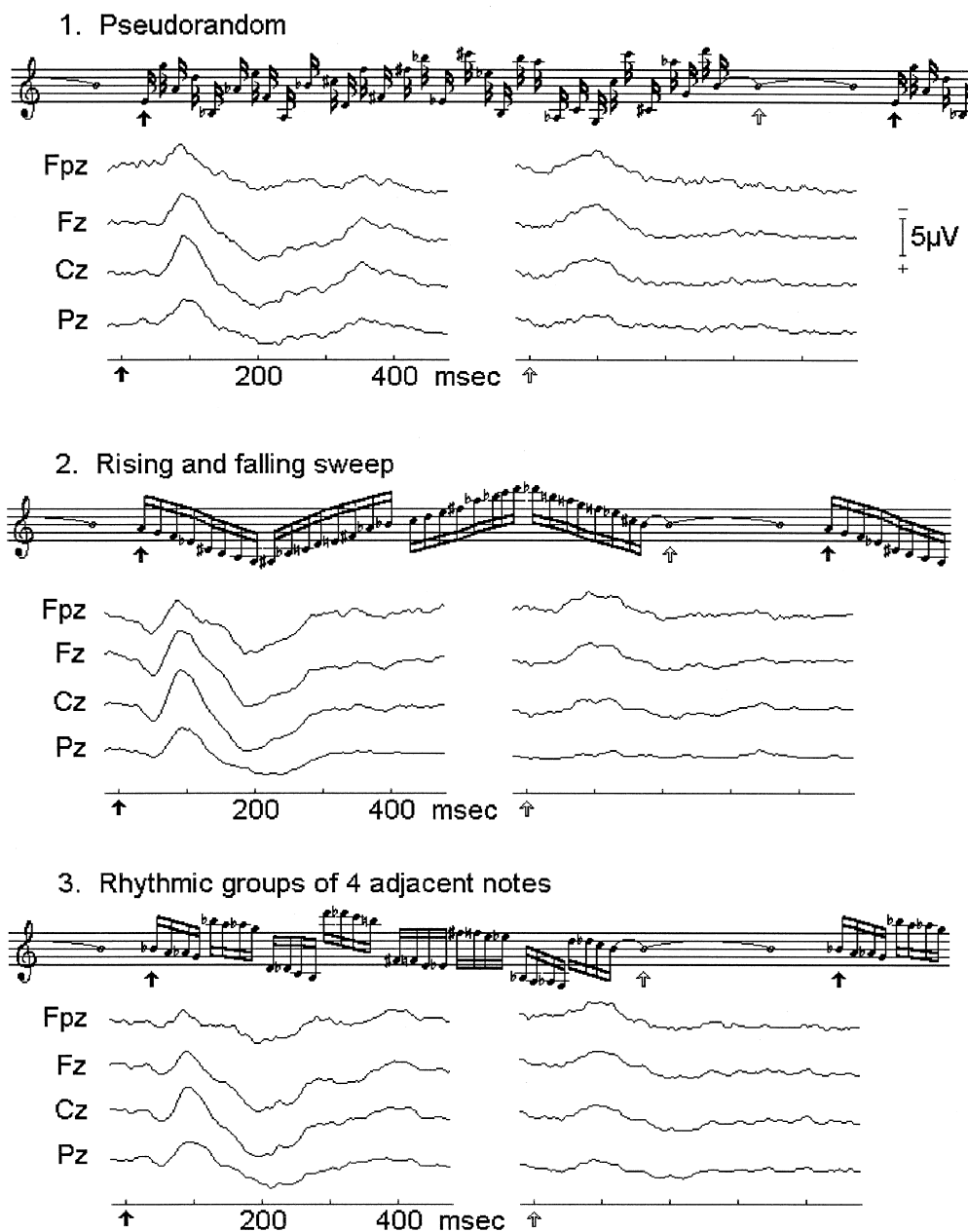


Fig. 6. Experiment 3. Group mean responses of five subjects at the beginning and end of 16/s sequences of connected tones, duration 2 s, followed by 2 s of a steady tone (B_4). The sequences comprised 32 notes in different patterns but with no repetition. A small mismatch negativity is evoked at the end of each sequence, its latency *c.* 100 ms after the next expected change in pitch with a more anterior distribution than the P1/N1/P2 complex to the first pitch change.

subjects who produced a distinct 2-peaked waveform were selected for dipole modelling, and the mean goodness of fit of only 90.7% for the first and 86% for the second peak suggests that other generators may have made a significant contribution. It is also hard to reconcile these authors'

conclusions with our [18] observation that the N1 to pitch change of a continuous complex tone has a distribution slightly posterior to that of the onset-N1; addition of the MMN to the onset-N1 should conversely result in an anterior shift. Our data, in other words, provide no evi-

Fig. 5. Experiment 3. Group mean responses of five subjects at the beginning and end of 16/s sequences of connected tones, duration 2 s, followed by 2 s of a steady tone (B_4). The patterns comprised 2, 3 and 5 notes in repeating sequences and 5 notes in a pseudorandom (non-rhythmic) sequence. In each case a typical P1/N1/P2 complex was evoked by the first change in pitch and a more anteriorly distributed mismatch negativity at the end, *c.* 100 ms after the next expected pitch change. The latter was of slightly lower amplitude for the pseudorandom five-note pattern.

dence that an MMN is generated when the tones are connected (or the gaps are less than about 200 ms) and pitch changes occur relatively infrequently. Under these conditions the mechanism for change detection based on the long auditory store appears to give way to a more direct process, possibly involving the short auditory store [11] which is believed to underly such phenomena as backward masking and the subjective increase in loudness when the duration of a tone is extended up to about 300 ms [43].

Experiments 1 and 3, however, suggest that a change-detection mechanism based on a longer duration sensory store does come into operation at the end of a rapid sequence of tones, signalling a change in the pitch modulation pattern. The main evidence for this was that when the modulation rate was varied the latency of the MMN was relatively constant with respect to the non-occurrence of the next pitch change, implying comparison of the instantaneous sound with a template of the preceding pitch modulation pattern and generation of a response when the expectancy that changes will continue to occur at the same rate is dissatisfied. We have not established the minimum duration of the sequence necessary for an MMN to be generated at the end, but pilot experiments suggest that at least 2 s may be necessary for maximal amplitude of the MMN to be achieved, and if at least three changes are necessary in order to establish an expectancy that further changes will follow at the same intervals, at a rate of 8/s a period of about 750 ms would seem to be the practical minimum. This clearly exceeds the duration of the short auditory store, and the similar distribution of the negativity at the end of pitch modulation to that of the conventionally elicited MMN (with silent gaps of 250 ms) suggests that these potentials probably reflect identical processes involving the long auditory store.

The MMN at the end of pitch modulation was most robust when the modulation rate was relatively fast, 16/s changes producing larger and more consistent responses than 8/s. In the third experiment it was found that the largest negativity was generated at the end of sequences which were both repetitive and rhythmic, but that smaller responses were still produced at the end of a pseudo-random sequence in which each note occurred only once and there was no overlying rhythmic structure. We have not yet examined whether a response can be elicited at the end of sequences which are temporally irregular, but the present evidence would seem to imply that, whereas repetition of particular pitches and their rhythmic grouping are relatively unimportant, it is crucial that the precise *timing* of the next change can be anticipated.

Several recent MMN studies have looked at the effect of changing the temporal patterning of tones, the methods converging with those of the present study in their use of relatively short inter-tone intervals. Using pairs of tones of two distinct pitches [49] or intensities [47], an MMN was shown to be generated by infrequently occurring pairs

differing from the standards in that their temporal order was reversed. It has also been shown [34] that an MMN is produced when the time interval between repeated tones is shortened or lengthened. The magnetic counterpart of the MMN was shown to be produced when a single element of a repeating cell of five tones was changed [3], its distribution suggesting a generator situated slightly medially to that responsive to deviant simple tones. The element which changed was of different pitch to the other tones in the cell, and it would be interesting to establish whether a similar response is elicited when just the deviant pattern involves repetition of one of its notes within the repeating cell. All these studies reinforce the conclusion that the auditory sensory stores play an important role in defining and detecting change in the temporal modulation pattern of the sound, by accurate encoding of pitch, intensity and timing information. One report [37] also suggests that sensory memory processes may be used to abstract information of a higher order, such as whether pairs of tones whose absolute pitch level is allowed to vary, together constitute a consistently rising or falling pattern.

Several groups have examined the possibility of using the MMN as a clinical tool for testing auditory sensory memory processes in patients with developmental, peripheral hearing, linguistic, attentional and learning deficits as well as brain damage [1,19,21,22,24,25,40]. Others have begun to explore the utility of the MMN for investigating the neurophysiological processes underlying speech perception [23,28]. However, a sequence of identical, disconnected 'standard' phonemes followed by a 'deviant' is clearly far removed from actual speech patterns. The present study demonstrates that sensory memory processes can be shown to operate during periods of rapid, repetitive or non-repetitive pitch modulation, much more closely akin to real speech patterns. Perhaps the most significant practical implication is that it is possible to record an MMN of large amplitude (often approaching 10 μV) without the need for off-line waveform subtraction or statistical comparison. Some workers who have attempted to record the MMN in a clinical context [e.g., Ref. [24]] have been frustrated by its poor signal/noise ratio and have concluded that more robust methods are needed. We believe we now have a tool which is capable of clinical application in responsive and unresponsive patients; at present this may be confined to establishing that the spectral representation and chronological processes of the long auditory store are grossly intact, but in future it should be possible to examine the contents and processing capabilities of the store in greater detail.

References

- [1] O. Aaltonen, J. Tuomainen, M. Laine, P. Niemi, Discrimination of speech and non-speech sounds by brain-damaged patients; electrophysiological evidence for distinct sensory processes, *Brain Lang.* 44 (1993) 139–152.

- [2] K. Alho, Cerebral generators of mismatch negativity (MMN) and its magnetic counterpart (MMNm) elicited by sound changes, *Ear Hear.* 16 (1995) 38–50.
- [3] K. Alho, M. Tervaniemi, M. Huottilainen, J. Lavikainen, H. Tiitinen, R.J. Ilmoniemi, J. Knuutila, R. Näätänen, Processing of complex sounds in the human auditory cortex as revealed by magnetic brain responses, *Psychophysiology* 33 (1996) 369–375.
- [4] S.D. Arlinger, L.B. Jerlwall, T. Ahren, E.C. Holmgren, Slow evoked cortical responses to linear frequency ramps of a continuous pure tone, *Acta Physiol. Scand.* 98 (1976) 412–424.
- [5] M. Besson, F. Macar, An event-related potential analysis of incongruity in music and other non-linguistic contexts, *Psychophysiology* 24 (1987) 14–25.
- [6] A.S. Bregman, *Auditory Scene Analysis*, MIT Press, Cambridge, MA, 1990.
- [7] K.L. Chobor, J.W. Brown, Phoneme and timbre monitoring in left and right cerebrovascular accident patients, *Brain Lang.* 30 (1987) 278–284.
- [8] M. Clynes, Dynamics of vertex potentials, the R-M brain function, in: E. Donchin, L.B. Lindsley (Eds.), *Averaged Evoked Potentials: Methods, Results, Evaluation*, National Aeronautics and Space Administration (NASA SP-19), Washington, DC, 1969, pp. 363–374.
- [9] D. Cohen, A. Erez, Event-related-potential measurements of cognitive components in response to pitch patterns, *Music Percept.* 8 (1991) 405–430.
- [10] D. Cohen, R. Granot, H. Pratt, A. Barnea, Cognitive meaning of musical elements as disclosed by event-related potentials (ERP) and verbal experiments, *Music Percept.* 11 (1993) 153–184.
- [11] N. Cowan, On short and long auditory stores, *Psychol. Bull.* 96 (1984) 341–370.
- [12] J.J. Eggermont, Temporal modulation transfer functions for AM and FM stimuli in cat auditory cortex. Effects of carrier type, modulating waveform and intensity, *Hear. Res.* 74 (1994) 51–66.
- [13] B.H. Gaese, J. Ostwald, Temporal coding of amplitude and frequency modulation in the rat auditory cortex, *Eur. J. Neurosci.* 7 (1995) 438–450.
- [14] M. Giard, F. Perris, J. Pernier, P. Bouchet, Brain generators implicated in the processing of auditory stimulus deviance: a topographic event-related potential study, *Psychophysiology* 27 (1990) 627–640.
- [15] P. Heil, R. Rajan, D.R. Irvine, Sensitivity of neurons in cat primary auditory cortex to tones and frequency-modulated stimuli: II. Organization of response properties along the ‘isofrequency’ dimension, *Hear. Res.* 63 (1992) 135–156.
- [16] S.J. Jones, AEPs at the onset and offset of repetitive sound modulation, due to mismatch with the contents of an auditory sensory store, *Electroencephalogr. Clin. Neurophysiol.* 84 (1992) 149–156.
- [17] S.J. Jones, C. Byrne, The AEP T-complex to synthesised musical tones: left/right asymmetry in relation to handedness and hemisphere dominance, *Electroencephalogr. Clin. Neurophysiol.* 108 (1998) 355–360.
- [18] S.J. Jones, O. Longe, M. Vaz Pato, Auditory evoked potentials to abrupt pitch and timbre change of complex tones: electrophysiological evidence of ‘streaming’?, *Electroencephalogr. Clin. Neurophysiol.* 108 (1998) 131–142.
- [19] N.M. Kane, S.H. Curry, S.R. Butler, B.H. Cummins, Electrophysiological indicator of awakening from coma, *Lancet* 341 (1993) 688.
- [20] M. Kohn, F. Lifshitz, D. Litchfield, Averaged evoked potentials and frequency modulation, *Electroencephalogr. Clin. Neurophysiol.* 45 (1978) 236–243.
- [21] P. Korpilahti, H. Lang, Auditory ERP components and mismatch negativity in dysphasic children, *Electroencephalogr. Clin. Neurophysiol.* 91 (1995) 256–264.
- [22] N. Kraus, T. McGee, Mismatch negativity in the assessment of central auditory function, *Am. J. Audiol.* 3 (1994) 139–151.
- [23] N. Kraus, T. McGee, T.D. Carrell, A. Sharma, Neurophysiological bases of speech discrimination, *Ear Hear.* 16 (1995) 19–37.
- [24] D. Kurtzberg, H.G. Vaughan Jr., J.A. Kreuzer, K.Z. Flegler, Developmental studies and clinical application of mismatch negativity: problems and prospects, *Ear Hear.* 16 (1995) 105–117.
- [25] A.H. Lang, O. Eerola, P. Korpilahti, I. Holopainen, S. Salo, E. Uusipaikka, O. Aaltonen, Clinical applications of the mismatch negativity, *Ear Hear.* 16 (1995) 117–129.
- [26] J. Lavikainen, M. Huottilainen, R.J. Ilmoniemi, J.T. Simola, R. Näätänen, Pitch change of a continuous tone activates two distinct processes in human auditory cortex: a study with whole-head magnetometer, *Electroencephalogr. Clin. Neurophysiol.* 96 (1995) 93–96.
- [27] J.P. Makela, R. Hari, A. Linnankivi, Different analysis of frequency and amplitude modulations of a continuous tone in the human auditory cortex: a neuromagnetic study, *Hear. Res.* 27 (1987) 257–264.
- [28] A.C. Maiste, A.S. Wiens, M.J. Hunt, M. Scherg, T.W. Picton, Event-related potentials and the categorical perception of speech sounds, *Ear Hear.* 16 (1995) 68–90.
- [29] T.M. McKenna, N.M. Weinberger, D.M. Diamond, Response of single auditory cortical neurons to tone sequences, *Brain Res.* 481 (1989) 142–153.
- [30] J.P. Mendelson, M.S. Cynader, Sensitivity of cat primary auditory cortex (AI) neurons to the direction and rate of frequency modulation, *Brain Res.* 327 (1985) 331–335.
- [31] R. Näätänen, The mismatch negativity: a powerful tool for cognitive neuroscience, *Ear Hear.* 16 (1995) 6–18.
- [32] R. Näätänen, T. Picton, The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure, *Psychophysiology* 24 (1987) 375–425.
- [33] R. Näätänen, P. Paavilainen, K. Alho, K. Reinikainen, M. Sams, Interstimulus interval and the mismatch negativity, in: C. Barber, T. Blum (Eds.), *Evoked Potentials III*, Butterworth, Boston, 1987, 392–397.
- [34] R. Näätänen, D. Jiang, J. Lavikainen, K. Reinikainen, P. Paavilainen, Event-related potentials reveal a memory trace for temporal features, *NeuroReport* 5 (1994) 310–312.
- [35] U. Neisser, *Cognitive Psychology*, Appleton-Century-Crofts, New York, 1967.
- [36] G. Nyman, K. Alho, P. Laurinen, P. Paavilainen, T. Radil, K. Reinikainen, M. Sams, R. Näätänen, Mismatch negativity (MMN) for sequences of auditory and visual stimuli: evidence for a mechanism specific to the auditory modality, *Electroencephalogr. Clin. Neurophysiol.* 77 (1990) 436–444.
- [37] P. Paavilainen, J. Saarinen, M. Tervaniemi, R. Näätänen, Mismatch negativity to changes in abstract sound features during dichotic listening, *J. Psychophysiol.* 9 (1995) 243–249.
- [38] K.A. Paller, G. McCarthy, C.C. Wood, Event-related potentials elicited by deviant endings to melodies, *Psychophysiology* 29 (1992) 202–206.
- [39] D.P. Phillips, M.N. Semple, M.B. Calford, L.M. Kitzes, Level-dependent representation of stimulus frequency in cat primary auditory cortex, *Exp. Brain Res.* 102 (1994) 210–226.
- [40] C.W. Ponton, M. Don, The mismatch negativity in cochlear implant users, *Ear Hear.* 16 (1995) 130–143.
- [41] W. Ritter, D. Deacon, H. Gomes, D.C. Javitt, H.G. Vaughan Jr., The mismatch negativity of event-related potentials as a probe of transient auditory memory: a review, *Ear Hear.* 16 (1995) 51–66.
- [42] S. Samson, R.J. Zatorre, Contribution of the right temporal lobe to musical timbre discrimination, *Neuropsychologia* 32 (1994) 231–240.
- [43] B. Scharf, A.J. Houtsma, Audition II: Loudness, pitch, localization, aural distortion, pathology, in: K.R. Boff, L. Kaufman, J.P. Thomas (Eds.), *Handbook of Perception and Human Performance: Vol. 1. Sensory Processes and Perception*, Wiley, New York, 1986, pp. 151–160.
- [44] M. Scherg, J. Vajsar, T.W. Picton, A source analysis of the late human auditory vertex potentials, *J. Cogn. Neurosci.* 1 (1989) 336–355.
- [45] C.E. Schreiner, J.R. Mendelson, M.L. Sutter, Functional topography

- of cat primary auditory cortex: representation of tone intensity, *Exp. Brain Res.* 92 (1992) 105–122.
- [46] E. Schröger, P. Paavilainen, R. Näätänen, Mismatch negativity to changes in a continuous tone with regularly varying frequencies, *Electroencephalogr. Clin. Neurophysiol.* 92 (1994) 140–147.
- [47] E. Schröger, M. Tervaniemi, C. Wolff, R. Näätänen, Preattentive periodicity detection in auditory patterns as governed by time and intensity information, *Cogn. Brain Res.* 4 (1996) 145–148.
- [48] M. Tervaniemi, K. Alho, M. Paavilainen, M. Sams, R. Näätänen, Absolute pitch and event-related brain potentials, *Music Percept.* 10 (1993) 305–316.
- [49] M. Tervaniemi, S. Maury, R. Näätänen, Neural representation of abstract stimulus features in the human brain as represented by the mismatch negativity, *NeuroReport* 5 (1994) 844–846.
- [50] M. Tervaniemi, I. Winkler, R. Näätänen, Pre-attentive categorization of sounds as revealed by event-related potentials, *NeuroReport* 8 (1997) 2571–2574.
- [51] B. Tian, J.P. Rauschecker, Processing of frequency-modulated sounds in the cat's anterior auditory field, *J. Neurophysiol.* 71 (1994) 1959–1975.
- [52] I.C. Whitfield, E.F. Evans, Responses of auditory cortical neurons to changing frequency, *J. Neurophysiol.* 28 (1965) 655–672.
- [53] I. Winkler, M. Tervaniemi, M. Huotilainen, R. Ilmoniemi, A. Ahonen, O. Salonen, C.-G. Standerskjold-Nordenstam, R. Näätänen, From objective to subjective: pitch representation in the human auditory cortex, *NeuroReport* 6 (1995) 2317–2320.
- [54] J.R. Wolpaw, J.K. Penry, A temporal component of the auditory evoked response, *Electroencephalogr. Clin. Neurophysiol.* 39 (1975) 609–620.
- [55] J.R. Wolpaw, J.K. Penry, Hemispheric differences in the auditory evoked response, *Electroencephalogr. Clin. Neurophysiol.* 43 (1977) 99–102.
- [56] R.J. Zatorre, S. Samson, Role of the right temporal neocortex in retention of pitch in auditory short-term memory, *Brain* 114 (1991) 2403–2417.
- [57] R.J. Zatorre, A.C. Evans, E. Meyer, Neural mechanisms underlying melodic perception and memory for pitch, *J. Neurosci.* 14 (1994) 1908–1919.