Spectro-temporal analysis of complex tones: two cortical processes dependent on retention of sounds in the long auditory store

S.J. Jones*, M. Vaz Pato, L. Sprague

The National Hospital for Neurology and Neurosurgery, London, UK

Abstract

Objectives: To examine whether two cortical processes concerned with spectro-temporal analysis of complex tones, a 'C-process' generating CN1 and CP2 potentials at cf. 100 and 180 ms after sudden change of pitch or timbre, and an 'M-process' generating MN1 and MP2 potentials of similar latency at the sudden cessation of repeated changes, are dependent on accumulation of a sound image in the long auditory store.

Methods: The durations of steady (440 Hz) and rapidly oscillating (440–494 Hz, 16 changes/s) pitch of a synthesized 'clarinet' tone were reciprocally varied between 0.5 and 4.5 s within a duty cycle of 5 s. Potentials were recorded at the beginning and end of the period of oscillation in 10 non-attending normal subjects.

Results: The CN1 at the beginning of pitch oscillation and the MN1 at the end were both strongly influenced by the duration of the immediately preceding stimulus pattern, mean amplitudes being 3–4 times larger after 4.5 s as compared with 0.5 s.

Conclusions: The processes responsible for both CN1 and MN1 are influenced by the duration of the preceding sound pattern over a period comparable to that of the 'echoic memory' or long auditory store. The store therefore appears to occupy a key position in spectro-temporal sound analysis. The C-process is concerned with the spectral structure of complex sounds, and may therefore reflect the 'grouping' of frequency components underlying auditory stream segregation. The M-process (mismatch negativity) is concerned with the temporal sound structure, and may play an important role in the extraction of information from sequential sounds. © 2000 Elsevier Science Ireland Ltd. All rights reserved.

Keywords: Auditory evoked potential; Mismatch negativity; Complex tones; Spectro-temporal sound analysis; Grouping; Stream segregation

1. Introduction

In the neurophysiological literature of the auditory cortex, considerable uncertainty still exists as to what are the most appropriate stimuli to apply. The tonotopic organization of the cortex is well-known, but many single-unit workers have moved away from the use of pure tones to frequency- and amplitude-modulated sounds (e.g. Whitfield and Evans, 1965; Mendelson and Cynader, 1985; McKenna et al., 1989; Heil et al., 1992; Eggermont, 1994; Phillips et al., 1994; Gaese and Ostwald, 1995), or even species-specific vocalizations (Wang et al., 1995). In human auditory evoked potential (AEP) research, the mismatch negativity (MMN) is recognized as reflecting processes concerned with discrimination of sequentially presented sounds, and

appears to be generated within or adjacent to the auditory cortex on the supratemporal plane (reviewed by Alho, 1995; Näätänen, 1995; Ritter et al., 1995). Although most MMN studies have employed disconnected pure tones occasionally changing in frequency, duration, intensity, etc., there is also literature on the MMN to complex and rapidly modulated tones (e.g. Schröger et al., 1994, 1996; Alho et al., 1996; Tervaniemi et al., 1997) and speech sounds (e.g. Kraus et al., 1995). It is generally agreed that the MMN depends on accumulation of the sound pattern in the 'echoic memory' or 'long auditory store', a pre-perceptual memory which retains the spectro-temporal structure of sounds over a period of several seconds (Cowan, 1984) and permits unattended sounds to be retrospectively perceived and analyzed.

In an earlier study of AEPs to synthesized musical instrument tones (Jones et al., 1998) it was found that the N1 peaking approximately 100 ms after a sudden change in the spectral composition (pitch or timbre) of a continuous tone was more posteriorly distributed on the scalp than the

* Corresponding author. Department of Clinical Neurophysiology, The National Hospital for Neurology and Neurosurgery, Queen Square, London WC1N 3BG, UK. Tel.: +44-207-837-3611 ext. 4109; fax: +44-207-713-7743.

E-mail address: sjjones@ion.ucl.ac.uk (S.J. Jones).

superficially similar N1 evoked at the onset of the tone from silence. The 'change-N1' and following large P2 were attenuated in the presence of continuous or intermittent background tones, to a greater degree when the latter were of similar timbre to the test tone. The effect of introducing a silent interval between tones alternating between two different timbres suggested that the change process only operates when the gaps are less than about 200 ms. This might be explained by the overlapping of consecutive tones in the 'short auditory store' (Cowan, 1984), a more elementary form of sensory memory which causes the perceptual image of very brief sounds to persist for at least 150 ms (Efron, 1970a,b,c) and whose other manifestations include loudness integration and backward recognition masking (Scharf and Houtsma, 1986).

When pitch changes are made at rates greater than about 10/s the associated N1 and P2 potentials are markedly attenuated or absent, but on resumption of a steady pitch a large (up to 10 μV) negative potential with a more anterior distribution is evoked (Vaz Pato and Jones, 1999). We have argued that this is probably an MMN in a relatively pure form, undistorted by overlapping N1 components. Its latency appears to be determined by the expected time of occurrence of the next pitch change, suggesting a process of comparison between the incoming sound and an extrapolated template of the preceding temporal modulation pattern. The potential was most strongly evoked after patterns which were both rhythmic and repetitive, but a smaller response was still present after sequences in which there was no rhythmic pattern or repetition of individual notes.

We therefore distinguish two processes concerned with the spectro-temporal analysis of continuous, complex sounds, i.e. a 'C-process' activated by change in the spectral composition and an 'M-process' involved with analysis of the temporal pattern. The former may be analogous to the 'grouping' or 'streaming' of simultaneous spectral components, such that a sound emanating from a single source but comprising many constituent frequencies is perceived as a whole (Bregman, 1990). The latter, on the other hand, may be an important stage in the extraction of temporally encoded information such as speech. In order to determine the dependence of these processes on the auditory stores, it is necessary to distinguish the effects of memory accumulation from those of refractoriness. Refractoriness will affect the amplitude of neuronal potentials when the generating process requires additional time to recover (as, for example, with restoration of the resting membrane potential after a nerve action potential), but when memory is involved the recorded signal may also depend on the strength of the memory trace on which the process operates. The fact that the C-process is entrained at the beginning of a period of oscillating pitch changes and the M-process at the end suggested a means of studying the memory-relatedness of both while holding the factor of refractoriness constant, by altering the durations of steady and oscillating pitch reci-

procally within a fixed duty cycle (constant stimulus onset asynchrony).

2. Material and methods

The subject group comprised 10 volunteers, 5 male and 5 female aged 21–49 years, who had no history of hearing impairment and who gave their informed consent according to the Declaration of Helsinki. The subjects sat in a reclining chair and were instructed to stay awake and read a book or magazine for the duration of the recording. Eight recording electrodes were attached to the scalp with paste at locations Fpz, Fz, Cz, Pz, F4, C4, F3 and C3 of the 10–20 system. The reference electrode was on the dorsum of the neck at the base of the skull. AEPs were recorded with a bandpass of 1–200 Hz and a sampling rate of 1 kHz for an epoch of 450 ms.

Artificial musical instrument tones were synthesized by a Yamaha MU10 tone generator controlled by an IBM-compatible notebook PC and presented to both ears via headphones at a moderate intensity (45–50 dB above threshold). The software used to construct and play stimulus sequences was Cubasis (Steinberg Soft & Hardware GmbH). The left and right output channels were split in order to use one channel for stimulation (presented to both ears) and the other to trigger the recording apparatus. By examination of the envelope the rise time of each tone was determined to be approximately 10 ms (latencies were measured from the mid-point of the rise time) and the period of overlap between contiguous tones was approximately 20 ms. The 'clarinet' timbre was chosen on account of its unexaggerated onset and subsequently steady intensity and pitch. The frequency spectrum was examined by a fast A/D converter and associated software (Pico Technology Limited, Cambridge, UK; Fig. 1).

There were 6 stimulus conditions, each involving a duty cycle (stimulus onset asynchrony) of 5 s within which the period of oscillating and steady pitch varied reciprocally (Fig. 2). In condition 'Osc1' the pitch oscillated between A_4 (440 Hz fundamental) and B_4 (494 Hz) at 16 changes/s for 0.5 s, followed by 4.5 s of steady A_4 pitch. In 'Osc2' the corresponding durations were 1 s (oscillation) and 4 s (steady), in 'Osc3' they were 2 and 3 s, respectively, in 'Osc4' 3 and 2 s, in 'Osc5' 4 and 1 s, and in 'Osc6' 4.5 and 0.5 s. Two averager stores were triggered, one at the beginning and the other at the end of oscillation (at the end of oscillation the trigger in fact coincided with the time of non-occurrence of the next expected pitch change). Responses to 30 stimuli were averaged and 3 repetitions were made of each average in a balanced design.

After having been examined for consistency, the repeated runs were averaged together for each subject. Peak latencies of the P1 (when present), N1 and P2 potentials were measured at Fz where amplitudes were usually largest, although in some individuals the C-potentials were slightly larger at Cz. At the same latency, the amplitudes at 8 scalp
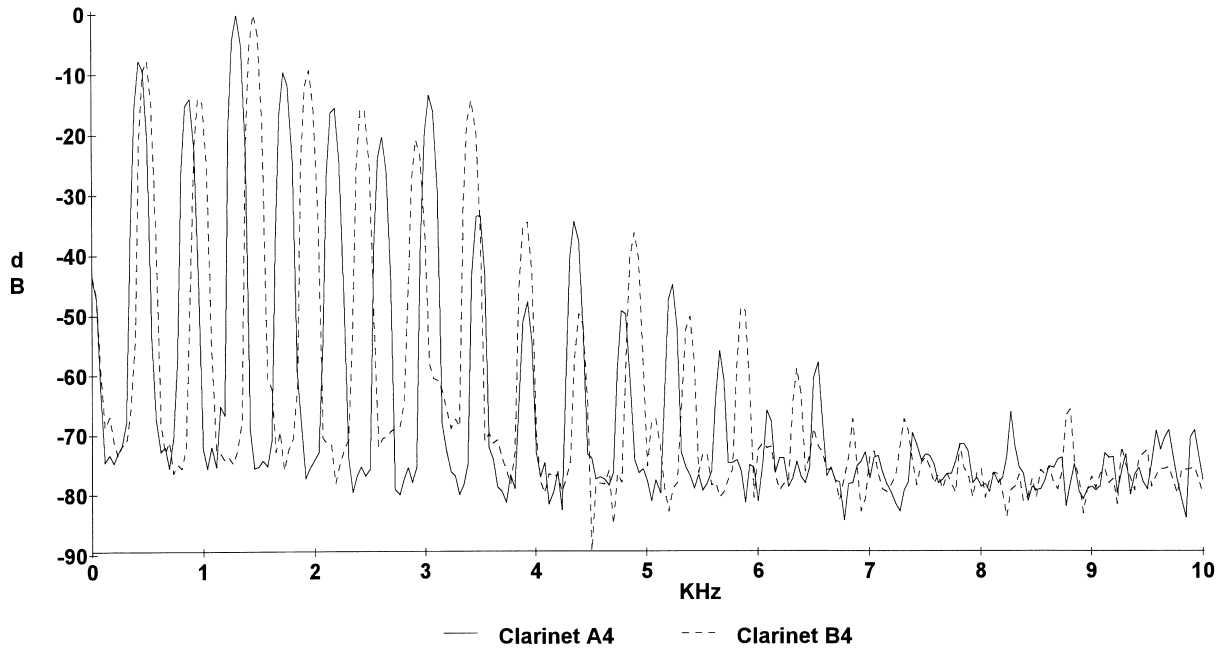
Fig. 1. Frequency spectra of synthesized 'clarinet' tones at the pitches used in the study (A_4 fundamental, 440 Hz; B_4 fundamental, 494 Hz).

locations were measured relative to the baseline potential (a representative point chosen within the first 20 ms of the recording epoch). Amplitude and latency values were analyzed by ANOVA and paired t test. Mean waveforms for each subject were superaveraged into group mean waveforms for illustration.

3. Results

3.1. Effect of stimulus duration on P1, N1 and P2 amplitude

Consistent responses consisting of a negativity peaking at approximately 100 ms (N1) and a positivity at 180 ms (P2) were evoked at the beginning and end of each oscillatory period, except after the shortest (0.5 s) duration of oscillating or steady pitch when the responses were sometimes ill-defined. The CN1 at the beginning of the oscillatory period was preceded by a CP1 peaking at about 55 ms, but no positivity was identifiable preceding the MN1 at the end of oscillation. All potentials were maximal in the fronto-central region and symmetrically distributed, there being no significant amplitude differences between F3 and F4 or C3 and C4. Values of mean amplitude and latency at Fz are listed in Table 1, and shown plotted against the duration of the preceding stimulus pattern (steady for CP1, CN1 and CP2, oscillating for MN1 and MP2) in Fig. 3. Fig. 4 shows the group mean waveforms in each condition.

The effect of preceding stimulus duration on CP1 amplitude was statistically significant ($F(5, 432) = 3.099$, $P = 0.009$). Paired t tests revealed no significant differences between individual conditions, but the overall tendency, unexpectedly, was for smaller CP1 amplitudes to be recorded

after longer periods of pitch constancy. The amplitudes of both CN1 and MN1, on the other hand, increased monotonically with the duration of the preceding stimulus, being 3–4 times larger after 4.5 s as compared with 0.5 s. After collapsing the N1 amplitude measures across electrode locations the main effect of stimulus duration was highly significant ($F(5, 108) = 21.730$, $P < 0.0001$) but there was no significant effect of C- versus M-response type and no significant
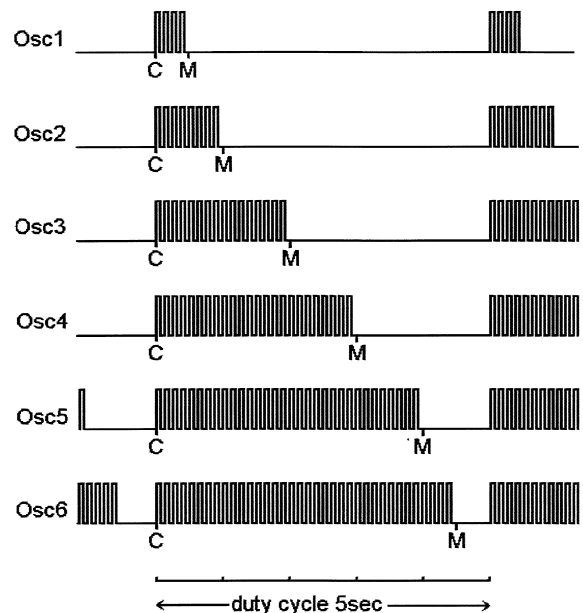

Fig. 2. Stimulus sequences used in the study. The trigger for C-responses coincided with the first change in pitch, while that for M-responses coincided with the first non-occurrence of pitch change, expected on the basis of extrapolation of the oscillatory pattern.

Table 1
P1, N1 and P2 amplitude and latency at Fz (mean ± SD)

		Preceding stimulus duration (s)					
		0.5	1	2	3	4	4.5
C-potentials at beginning of pitch oscillation							
		Osc1	Osc2	Osc3	Osc4	Osc5	Osc6
CP1	Amplitude (μV)	2.8 ± 0.9	1.9 ± 1.2	2.3 ± 1.8	2.3 ± 2.5	2.0 ± 1.4	1.5 ± 1.7
	Latency (ms)	60 ± 10	60 ± 7	57 ± 6	54 ± 7	54 ± 6	54 ± 9
CN1	Amplitude (μV)	−1.9 ± 2.1	−2.7 ± 1.7	−4.9 ± 1.5	−6.4 ± 2.9	−7.0 ± 3.0	−7.7 ± 3.5
	Latency (ms)	95 ± 9	92 ± 7	90 ± 8	91 ± 7	91 ± 8	91 ± 7
CP2	Amplitude (μV)	3.0 ± 1.8	4.9 ± 2.4	5.6 ± 2.3	5.4 ± 2.3	6.4 ± 2.0	6.0 ± 2.9
	Latency (ms)	167 ± 17	159 ± 8	166 ± 11	168 ± 15	173 ± 11	173 ± 10
M-potentials at end of pitch oscillation							
		Osc6	Osc5	Osc4	Osc3	Osc2	Osc1
MN1	Amplitude (μV)	−1.5 ± 1.6	−3.3 ± 1.5	−4.9 ± 1.3	−6.4 ± 1.5	−6.6 ± 1.7	−7.2 ± 2.2
	Latency (ms)	95 ± 13	93 ± 13	101 ± 8	105 ± 9	104 ± 6	107 ± 8
MP2	Amplitude (μV)	4.5 ± 1.7	5.4 ± 2.6	5.9 ± 2.1	5.8 ± 1.5	5.0 ± 2.5	4.1 ± 2.8
	Latency (ms)	177 ± 17	177 ± 5	182 ± 10	181 ± 7	187 ± 13	188 ± 1

interaction between stimulus duration and response type. Post-hoc t tests confirmed significant amplitude enhancement of both CN1 and MN1 after a stimulus duration of 4.5 s as compared with 2 s or less ($P < 0.05$ or better).

Somewhat different behaviour was observed for the P2 potentials. The MP2 in fact reached its highest mean amplitude after 2 or 3 s of pitch oscillation, while the CP2 reached highest mean amplitude after 4 s of steady pitch (Table 1, Fig. 3). However, when collapsed across electrode locations, the main effects of stimulus duration and response type and their interaction were all non-significant. In contrast to CN1 and MN1, therefore, there was no clear evidence for stimulus duration-dependency of CP2 or MP2 amplitude between 0.5 and 4.5 s.
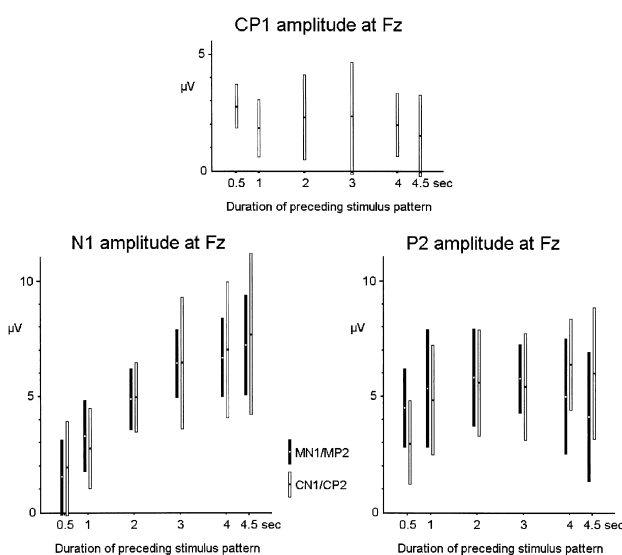

Fig. 3. Mean and standard deviation of CP1, CN1, CP2, MN1 and MP2 amplitude plotted against the duration of the preceding sound pattern (steady pitch for C-responses, oscillating at 16 changes/s for M-responses).

3.2. Effect of stimulus duration on P1, N1 and P2 latency

The main effect of stimulus duration on CP1 latency was not statistically significant. As regards the N1, the effect of response type was highly significant ($F(1, 108) = 30.080$, $P < 0.0001$), the latency of CN1 being up to 15 ms shorter than that of MN1 (Table 1). The effect of the duration of the preceding stimulus was not significant but the interaction between stimulus duration and response type was significant ($F(5, 108) = 2.788$, $P = 0.021$). Consequently, examining the two response types separately, the main effect of stimulus duration on MN1 latency was found to be significant ($F(5, 45) = 4.480$, $P = 0.002$), longer latencies being associated with longer stimulus durations. The tendency for CN1 was apparently in the opposite sense (longer latencies being associated with shorter stimulus durations) although this just failed to reach statistical significance ($F(5, 45) = 2.096$, $P = 0.083$). The effect of stimulus duration on the latency difference between CN1 and MN1 (comparing conditions with the same duration of the preceding stimulus, Osc1 with Osc6, Osc2 with Osc4, etc.) was highly significant ($F(5, 45) = 6.482$, $P = 0.0001$).

Examination of P2 latencies showed a significant main effect of stimulus duration ($F(5, 108) = 3.189$, $P = 0.010$), longer latencies being associated with longer durations, and of response type ($F(1, 108) = 40.504$, $P < 0.001$), MP2 latencies being on average longer than CP2 latencies, but no significant interaction.

3.3. Scalp distribution of P1, N1 and P2 potentials

As in the previous study (Vaz Pato and Jones, 1999), the MN1 and MP2 at the end of the period of pitch oscillation appeared to be more anteriorly distributed on the scalp than the CP1, CN1 and CP2 at the beginning of the oscillatory period. Whereas the CN1 was of comparable amplitude at Fz and Cz, MN1 was more evidently larger at Fz for all
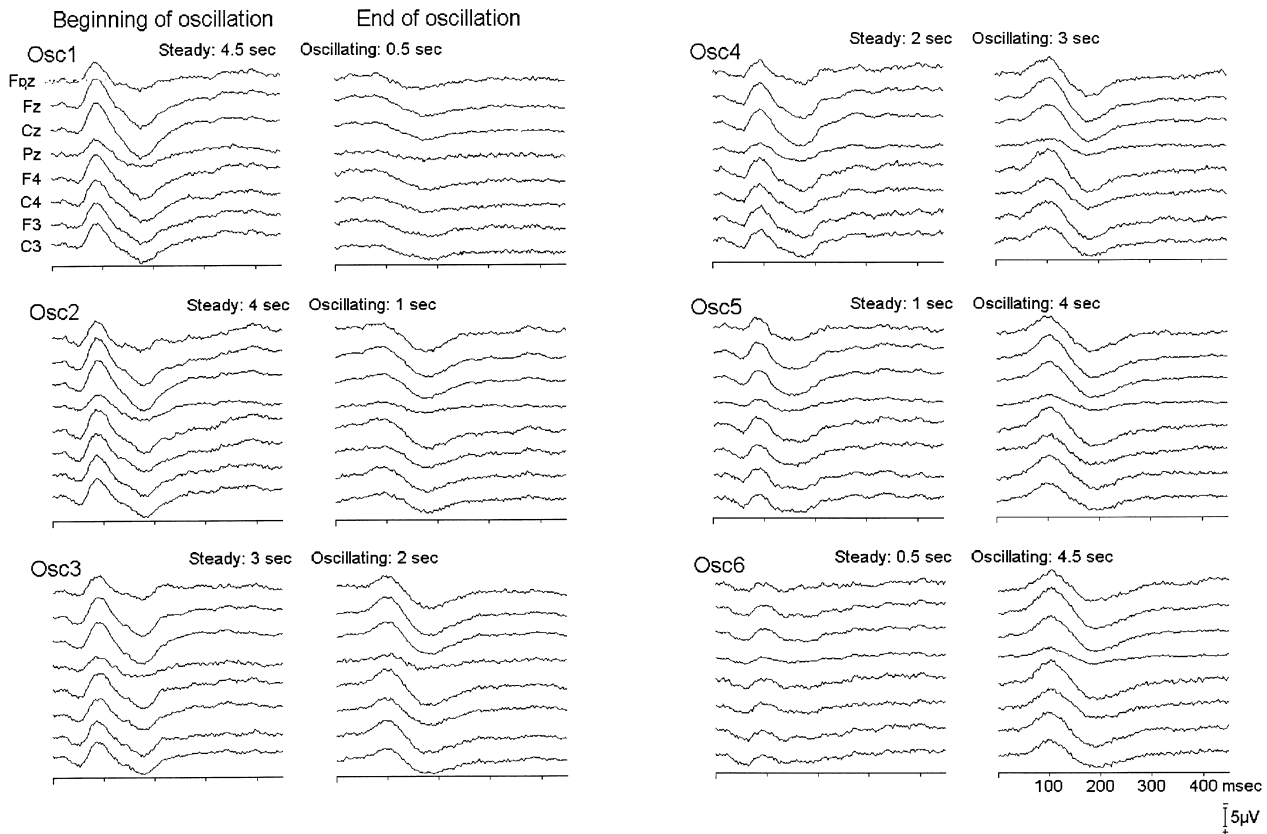
Fig. 4. Group mean waveforms of C-responses (left) and M-responses (right) at the beginning and end of pitch oscillation in each condition. It can be perceived that the CN1 is usually of similar amplitude at Fz and Cz, while the MN1 is slightly larger at Fz.

stimulus durations except the shortest (Fig. 5). When the amplitude data were collapsed across stimulus durations, the main effect of electrode location was highly significant for N1 ($F(7, 144) = 5.925$, $P < 0.0001$) and P2 ($F(7, 144) = 10.169$, $P < 0.0001$). The effect of response type was not significant, nor was there a significant interaction between response type and electrode location, although for P2 the interaction approached significance ($F(7, 144) = 1.616$, $P = 0.135$). It could not be proven from these data, therefore, that the two responses were differently distributed on the scalp, although the general tendency for M-potentials to be more anteriorly distributed was as previously noted.

4. Discussion

Long-latency AEPs to sudden or ramp-like changes in the pitch of a continuous pure tone received some attention in the earlier literature (Clynes, 1969; Arlinger et al., 1976; Kohn et al., 1978). A P1/N1/P2 complex was evoked at the beginning of ramp changes but not at the end, suggesting that this might represent the onset of auditory change or 'motion'. The findings are compatible with those obtained more recently by Makela et al. (1987), who recorded

magnetic fields P60m, N100m and P200m to amplitude and frequency modulation of continuous pure tones and concluded that the processes are independent of one another. Subsequently, Makela et al. (1988) compared the magnetic fields evoked by the onset of a noise burst with the fields of similar latency (cf. 100 ms) evoked when the noise changed abruptly to a square-wave tone. Their different recovery cycles and relatively low mutual interaction suggested that distinct processes are engaged by the onset of a sound and change in its spectral composition. In another magnetoencephalographic study, Lavikainen et al. (1995) concluded that the N1m produced by a brief change in the frequency of a pure tone was a combination of onset-N1m and MMNm fields. This cannot easily be reconciled with our own finding (Jones et al., 1998) that the N1 potential evoked by spectral change of a continuous complex tone was more posteriorly distributed than the N1 at tone onset. Addition of an MMN to the onset-N1 should conversely result in an anterior shift. Rather than a compound of onset-N1 and MMN potentials, we would argue that the CN1 represents a unitary process activated by change in the spectral pattern of a sound.

We have also previously argued (Vaz Pato and Jones, 1999) that the MN1 at the end of a period of rapid pitch oscillation must represent a different process from the CN1
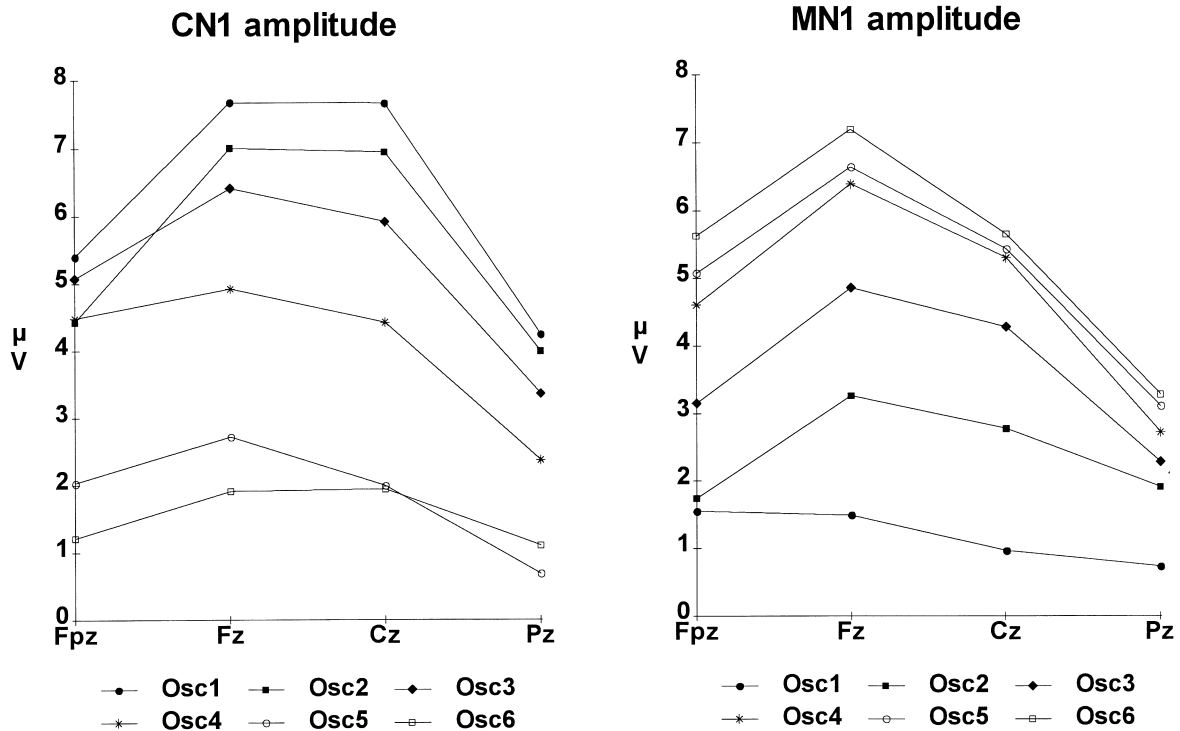
CN1 amplitude

MN1 amplitude

Fig. 5. Sagittal profile of mean CN1 and MN1 amplitudes at 4 midline scalp electrodes in each condition.

at the beginning. Since the CN1 is almost completely abolished when pitch changes occur at a rate of 16/s, it would not be logically plausible for a large CN1 to be generated by the last change. By varying the rate of changes between 8, 12 and 16/s it was shown that the MN1 latency was relatively constant, not with respect to the last pitch change which actually occurred but with respect to the next expected change which failed to occur. It is conceivable that the sudden non-occurrence of a pitch change, expected on the basis of extrapolation of the preceding modulation pattern, might cause the CN1 generator to become dishabituated. However, in this as in the previous study there was a tendency for the N1 at the end of the period of pitch modulation to have a scalp distribution anterior to that of the CN1, suggesting a different underlying neuronal population. In neither study could unequivocal statistical proof of the different scalp distributions be demonstrated, but this might be addressed with detailed mapping studies and more powerful statistical techniques. Further evidence for distinct processes comes from the fact that the CN1, but not the MN1, was consistently preceded by a P1 potential. In the earlier study the MN1 distribution was found to be similar to that of the mismatch negativity evoked by the same tones presented discontinuously with occasionally deviant pitch. Our hypothesis, therefore, is that the MN1 represents the mismatch negativity recorded in a relatively pure form, without distortion by overlapping onset-N1 or change-N1 components.

If the amplitudes of CN1 and MN1 were only dependent on the refractory period of the generating process, and not

on memory accumulation of the preceding sound pattern, no difference would be expected between the 6 conditions since in every case the stimulus onset asynchrony was 5 s. Therefore, each change should find the relevant process in a similar refractory state. The fact that both responses were strongly influenced by the duration of the preceding stimulus can therefore plausibly be attributed to the accumulating strength of a memory trace. Further studies using a longer duty cycle would be necessary to ascertain the time course of memory accumulation more precisely, but present evidence suggests an exponential process with a time constant in the order of 3–5 s. This agrees with many estimates of the duration of the 'long auditory store' (Cowan, 1984).

We were somewhat surprised to find the dependency of CN1 amplitude on the duration of the preceding stimulus to be very similar to that of MN1, since the findings of a previous study (Jones et al., 1998) had suggested a different time constant involved in the C-process. The effect of introducing silent gaps of variable duration between two tones of different spectral composition (timbre) suggested that the processes responsible for CN1 and CP2 only become active with gaps less than about 200 ms (this experiment has yet to be performed with tones of the same timbre but different pitch, but in other respects the responses to pitch and timbre change were very similar to one another). There is considerable psychophysical evidence (reviewed by Cowan, 1984) for a 'short auditory store' of up to 200 ms in duration that is responsible for integration of sounds over such periods. In particular, the studies of Efron (1970a,b,c) indicate that the

perceived durations of very brief sounds tend to be consistently extended up to 130–180 ms, this representing the duration of auditory 'persistence'. Electro- and magnetoencephalic studies of the MMN (Tervaniemi et al., 1997; Winkler et al., 1998; Yabe et al., 1998) have also found evidence for sound integration occurring over such intervals, so it appears that both C- and M-processes involve the participation of both the short and the long auditory store.

Also unexpectedly, no overall effect of stimulus duration on the amplitude of CP2 and MP2 was found, while CP1 showed the converse tendency to CN1, becoming larger after shorter periods of stimulus constancy. It is possible that the latter effect was illusory, caused by progressive 'unmasking' of the CP1 as the overlapping CN1 became smaller. We cannot at present account for the fact that CP2 and MP2 amplitudes were apparently independent of stimulus duration, although their latencies did show a significant tendency to increase.

Since the CN1 and MN1 were not absolutely distinguishable in respect of their scalp distributions, it might be argued that the increasing amplitude of CN1 with longer preceding periods of pitch constancy might have been due to an increasing contribution from the population of neurones responsible for MN1. In fact, however, the topographic distinction between the two potentials became more rather than less evident after longer stimulus durations. Also, whereas MN1 showed a significant tendency for its latency to increase with increasing stimulus duration, the (just nonsignificant) trend for CN1 was in the opposite direction and the differential effect when CN1 latency was compared with MN1 latency was highly significant. A possible reconciliation of these findings is that the N1 latency was in both cases longer when the total number of pitch changes in the preceding 5 s was greater. Perhaps, therefore, the time the C- and M-processes take to occur is determined by the degree of complexity, not just of the immediately preceding sound but of the whole image contained in the long auditory store.

The C-potentials were previously found to be attenuated by the presence of a background tone, but the effect was less marked when the latter was of different timbre to the test tone (Jones et al., 1998). This suggested that the C-potentials are generated subsequent to the perceptual 'grouping' of the frequency components of the test tone (Bregman, 1990), which can more effectively be distinguished as a single entity when the background tone is of dissimilar timbre (it is well-known that simultaneously present musical tones of the same or different pitch tend to contrast more effectively when played by instruments of different type). Since the auditory cortex is organized tonotopically, low frequencies generally represented more laterally on the supratemporal plane and high frequencies more medially (recently demonstrated in man by Howard et al., 1996), spectral grouping is likely to involve the co-ordinated firing or transcortical 'binding' of neurones whose 'best frequency' characteristics correspond to the spectral components of the sound. The C-process may therefore be considered as reflecting a change in the spectral binding pattern when one complex tone changes to another.

If the CN1 and MN1 represent processes respectively concerned with analysis of change in the spectral and temporal dimensions of the sound, it remains to be considered what processes are represented by the onset-N1 to a brief tone burst or click. Depending on the circumstances, several processes are known to contribute to the onset-N1, some of them exogenous and some endogenous. However, in their comprehensive review, Näätänen and Picton (1987) identified only a single 'true' N1 component arising in the supratemporal cortex of each hemisphere. It is possible that the 'true' onset-N1 may represent a primary cortical response while the CN1 and MN1 reflect secondary responses, but if so it is surprising that the refractory period of the onset-N1 is also in the order of several seconds, suggesting again the interposition of the long auditory store. Alternatively, it is arguable that even the 'true' onset-N1 may be a hybrid of more than one component. At the onset of a sound new information is conveyed, not only in the spectro-temporal domain but also in the spatial dimensions. N1-like potentials are evoked by change in the spatial characteristics of a sound, chiefly represented by the interaural time and intensity differences (Ungan et al., 1989; McEvoy et al., 1990; Jones et al., 1991; Ungan and Özmen, 1996), and these same processes would probably be activated by the onset of sound after silence.

The main conclusion from the study, therefore, is that both the C-process of spectral pattern analysis represented by CN1 and the M-process concerned with analysis of the temporal sound pattern represented by MN1 appear to depend on accumulation of a sound image in the long auditory store. Further study of these and the processes concerned with change in the spatial parameters of sounds may help illuminate how the brain perceives auditory 'objects', distinguishing them from the background and extracting information from their pattern of spectral changes over time.

Acknowledgements

References

Alho K. Cerebral generators of mismatch negativity (MMN) and its magnetic counterpart (MMNm) elicited by sound changes. Ear Hear 1995;16:38–50.

Alho K, Tervaniemi M, Huotilainen M, Lavikainen J, Tiitinen H, Ilmoniemi RJ, Knuutila J, Näätänen R. Processing of complex sounds in the human auditory cortex as revealed by magnetic brain responses. Psychophysiology 1996;33:369–375.

Arlinger SD, Jerlvall LB, Ahren T, Holmgren EC. Slow evoked cortical

responses to linear frequency ramps of a continuous pure tone. Acta Physiol Scand 1976;98:412–424.

Bregman AS. Auditory scene analysis, Cambridge, MA: MIT Press, 1990.

Clynes M. Dynamics of vertex potentials, the R-M brain function. In: Donchin E, Lindsley LB, editors. Averaged evoked potentials: methods, results, evaluation, Washington, DC: National Aeronautics and Space Administration (NASA SP-19), 1969. pp. 363–374.

Cowan N. On short and long auditory stores. Psychol Bull 1984;96:341–370.

Efron R. Effects of stimulus duration on perceptual onset and offset latencies. Percep Psychophys 1970a;8:231–234.

Efron R. The relationship between the duration of a stimulus and the duration of a perception. Neuropsychologia 1970b;8:37–55.

Efron R. The minimum duration of a perception. Neuropsychologia 1970c;8:57–63.

Eggermont JJ. Temporal modulation transfer functions for AM and FM stimuli in cat auditory cortex. Effects of carrier type, modulating waveform and intensity. Hear Res 1994;74:51–66.

Gaese BH, Ostwald J. Temporal coding of amplitude and frequency modulation in the rat auditory cortex. Eur J Neurosci 1995;7:438–450.

Heil P, Rajan R, Irvine DR. 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 1992;63:135–156.

Howard 3rd MA, Volkov IO, Abbas PJ, Damasio H, Ollendieck MC, Granner MA. A chronic microelectrode investigation of the tonotopic organization of human auditory cortex. Brain Res 1996;724:260–264.

Jones SJ, Pitman JR, Halliday AM. Scalp responses to coherence and discoherence of binaural noise and change in the inter-aural time difference: a specific binaural evoked potential or a "mismatch" response? Electroenceph clin Neurophysiol 1991;80:147–154.

Jones SJ, Longe O, Vaz Pato M. Auditory evoked potentials to abrupt pitch and timbre change of complex tones. Electrophysiological evidence of "streaming"? Electroenceph clin Neurophysiol 1998;108:131–142.

Kohn M, Lifshitz F, Litchfield D. Averaged evoked potentials and frequency modulation. Electroenceph clin Neurophysiol 1978;45:236–243.

Kraus N, McGee T, Carrell TD, Sharma A. Neurophysiological bases of speech discrimination. Ear Hear 1995;16:19–37.

Lavikainen J, Huotilainen M, Ilmoniemi RJ, Simola JT, Näätänen R. Pitch change of a continuous tone activates two distinct processes in human auditory cortex: a study with whole-head magnetometer. Electroenceph clin Neurophysiol 1995;96:93–96.

Makela JP, Hari R, Linnankivi A. Different analysis of frequency and amplitude modulations of a continuous tone in the human auditory cortex: a neuromagnetic study. Hear Res 1987;27:257–264.

Makela JP, Hari R, Leinonen L. Magnetic responses of the human auditory cortex to noise/square wave transitions. Electroenceph clin Neurophysiol 1988;69:423–430.

McEvoy LK, Picton TW, Champagne SC, Kellett AJ, Kelly JB. Human evoked potentials to shifts in the lateralization of a noise. Audiology 1990;29:163–180.

McKenna TM, Weinberger NM, Diamond DM. Response of single auditory cortical neurons to tone sequences. Brain Res 1989;481:142–153.

Mendelson JP, Cynader MS. Sensitivity of cat primary auditory cortex (AI) neurons to the direction and rate of frequency modulation. Brain Res 1985;327:331–335.

Näätänen R. The mismatch negativity: a powerful tool for cognitive neuroscience. Ear Hear 1995;16:6–18.

Näätänen R, 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.

Phillips DP, Semple MN, Calford MB, Kitzes LM. Level-dependent representation of stimulus frequency in cat primary auditory cortex. Exp Brain Res 1994;102:210–226.

Ritter W, Deacon D, Gomes H, Javitt DC, Vaughan Jr HG. The mismatch negativity of event-related potentials as a probe of transient auditory memory: a review. Ear Hear 1995;16:51–66.

Scharf B, Houtsma AJ. Audition II: loudness, pitch, localization, aural distortion, pathology. In: Boff KR, Kaufmann L, Thomas JP, editors. Sensory processes and perception, Handbook of perception and human performance, 1. New York: Wiley, 1986. pp. 151–160.

Schröger E, Paavilainen P, Näätänen R. Mismatch negativity to changes in a continuous tone with regularly varying frequencies. Electroenceph clin Neurophysiol 1994;92:140–147.

Schröger E, Tervaniemi M, Wolff C, Näätänen R. Preattentive periodicity detection in auditory patterns as governed by time and intensity information. Cogn Brain Res 1996;4:145–148.

Tervaniemi M, Schröger E, Näätänen R. Pre-attentive processing of spectrally complex sounds with asynchronous onsets: an event-related potentials study with human subjects. Neurosci Lett 1997;227:197–200.

Ungan P, Özmen B. Human long-latency responses to brief interaural disparities of intensity. Electroenceph clin Neurophysiol 1996;99:479–490.

Ungan P, Sahinoglu B, Utkucal R. Human laterality reversal auditory evoked potentials: stimulation by reversing the interaural delay of dichotically presented continuous click trains. Electroenceph clin Neurophysiol 1989;73:306–321.

Vaz Pato M, Jones SJ. Cortical processing of complex tone stimuli: mismatch negativity at the end of a period of rapid pitch modulation. Cogn Brain Res 1999;7:295–306.

Wang X, Merzenich MM, Beitel R, Schreiner CE. Representation of a species-specific vocalization in the primary auditory cortex of the common marmoset: temporal and spectral characteristics. J Neurophysiol 1995;74:2685–2706.

Whitfield IC, Evans EF. Responses of auditory cortical neurons to changing frequency. J Neurophysiol 1965;28:655–672.

Winkler I, Czigler I, Jaramillo M, Paavilainen P, Näätänen R. Temporal constraints of auditory event synthesis: evidence from ERPs. NeuroReport 1998;9:495–499.

Yabe H, Tervaniemi M, Sinkkonen J, Huotilainen M, Ilmoniemi RJ, Näätänen R. Temporal window of integration of auditory information in the human brain. Psychophysiology 1998;35:615–619.