The accuracy of sound duration representation in the human brain determines the accuracy of behavioural perception

Elena Amenedo and Carles Escera[1]
Department of Clinical Psychology and Psychobiology, University of Santiago de Compostela, Campus Sur S/N,
15706 Santiago de Compostela, Galicia, Spain
[1]Neurodynamics Laboratory, Department of Psychiatry and Clinical Psychobiology, University of Barcelona, Spain

Abstract

In recent years, the links between neural activity and perception have been an area of interest in cognitive neuroscience. Combined psychophysiological and psychophysical experiments provide a new powerful tool for establishing the relationship between neural activity and perceptual performance. In animals, intracellular recordings combined with psychophysical detection indices have revealed that a particular neuron or set of neurons can play a critical role in the generation of a perceptual event, showing detection functions (referred to as neurometric functions) which are remarkably similar to psychophysical detection functions, or psychometric functions (Parker & Newsome (1998) *Annu. Rev. Neurosci.*, **21**, 227–277). As noninvasive techniques for recording neural activity are now available, studies combining neuroelectric and psychophysical measures in humans are sparse. In the present study, the accuracy of the human brain in detecting differences in sound duration and the subject's ability to perceive the same differences were tested by means of mismatch negativity (MMN) and the distance between the distributions of false alarms and hits (sensitivity index *d'*), respectively. It was found that the accuracy of the human auditory system to represent sound duration information is related to the duration context in which the sounds are heard, and that these contextual representations determine the accuracy of perception at the behavioural level.

Introduction

Cognitive neuroscience is interested in specifying the ways in which environmental information is represented in the human brain to originate perception. The rationale is that if the neural activity of a particular brain region represents a stimulus dimension, then it is likely that this region has a direct role in mediating behavioural responses to that stimulus dimension.

The representation of acoustic information in the human brain has gained considerable ground since Näätänen *et al.* (1978) found that auditory stimulus deviance *per se* produces a measurable brain response, the mismatch negativity (MMN) event-related brain potential (ERP), to any change occurring in the acoustic environment (Näätänen, 1990; Näätänen & Alho, 1997).

Since its first observation, MMN has proved to be a powerful tool to reveal auditory-cortex neural traces of simple and complex auditory stimulus patterns (Ford & Hillyard, 1981; Sams *et al.*, 1985; Nordby *et al.*, 1988; Hari *et al.*, 1989; Kaukoranta *et al.*, 1989; Lounasmaa *et al.*, 1989; Näätänen *et al.*, 1989a, b; Lyytinen *et al.*, 1992; Paavilainen *et al.*, 1993; Schröger & Wolff, 1996; Joutsiniemi *et al.*, 1998), including phoneme representations (Dehaene-Lambertz, 1997; Näätänen *et al.*, 1997; Kraus *et al.*, 1998) and their development with language acquisition (Cheour *et al.*, 1998).

Establishing the precision with which sounds are represented at perceptual level has great relevance because in humans it is basic for speech perception (Phillips, 1993); a way to achieve this objective is to study the electrophysiological signals of perceptual accuracy. In this regard, MMN provides an objective measure of auditory discrimination ability, as MMN sensitivity to slight stimulus changes in the spectral content of an auditory stimulus approximates to the ability of detecting behaviourally these same changes (Lang *et al.*, 1990; Näätänen *et al.*, 1993; Aaltonen *et al.*, 1994; Kraus *et al.*, 1995; Tervaniemi *et al.*, 1997).

In spite of the above findings, no studies have combined MMN with psychophysical indexes of perceptual accuracy. The index of detection sensitivity more frequently employed in psychophysics is *d'*. This index constitutes an objective measure of the observer's confidence when deciding about the occurrence of a signal embedded in a streaming noise. The *d'* index has the advantage over the more frequently employed performance measures, such as the hit rate, of providing an estimate of the perceptual accuracy which is not affected by variables which shift the observer's response criterion during the execution of the task (Gescheider, 1997) as, for instance, responding fast at the expenses of committing more false alarms, or on the contrary, responding with confidence at the expense of slowing down the response speed.

The objective of the present study was to determine the accuracy with which sound duration is represented in the human brain at perceptual level, where it is accessible, for instance, for language perception and recognition. For this purpose, neuroelectric responses

and performance were measured to changes in sound duration in two separate sessions with the aim of investigating any relationship between MMN elicitation and behavioural detection.

Materials and methods

Subjects

Ten right-handed healthy volunteers (18–26 years, mean age = 21.7 ± 2.1, 3 males) with normal hearing participated in the experiment.

Stimuli

The stimuli were spectrally rich tones consisting of four harmonic pure tones of 500, 1000, 1500 and 2000 Hz, which were scaled in intensity by 1/2, 1/4, 1/6 and 1/8, respectively.

Stimulation procedure

Subjects were presented with 20 different blocks of 1000 spectrally rich tones. In separate blocks, sounds of 50 or 100 ms (including 5 ms of rise and fall times) of contextual duration were randomly interspersed with 'deviant' spectrally rich tones, these being 10, 20, 30, 40, and 60% shorter or longer than the contextual duration. Thus, within each block, only one deviant and one standard tone was presented (Fig. 1). In the first part of the experiment (the ERP recording sessions), the MMN was measured while subjects sat comfortably in a reclining chair and read a self-selected book ignoring the auditory stimulation. In the second part (the behavioural-detection sessions), the subject's ability to behaviourally detect the deviant stimuli was tested. Subjects attended to the same stimulus sequences presented in the ERP sessions, and responded to the deviant spectrally rich tones by pressing a response-button as quickly and accurately as possible.

Sounds were binaurally delivered at an intensity of 85 dB SPL (sound pressure level) and with a constant interstimulus interval (ISI) of 400 ms from onset-to-onset. The standard stimuli showed a probability of 0.8 and the deviant stimuli a probability of 0.2 in each stimulus block. The order of blocks was counterbalanced across subjects in the first and second sessions of the experiment.

Electrophysiological recordings

Measurements were performed in an electrically and acoustically shielded room. The electroencephalogram (EEG) was continuously recorded (bandpass 0.1–100 Hz, A/D rate = 500 Hz) from eight scalp locations of the 10–20 International System: Fp1, Fp2, F3, Fz, F4, C3, Cz, C4, and from two additional electrodes placed at the left mastoid (LM) and right mastoid (RM). The horizontal electrooculogram (EOG) was recorded with an electrode attached to the outer canthus of the right eye. The vertical EOG was assessed using the recordings from Fp1 and Fp2 electrodes. The common reference electrode was placed on the tip of the nose.

Data processing

The ERPs were averaged off-line for standard and deviant stimuli, separately in each block and subject, for a period of 600 ms including a prestimulus baseline of 100 ms. Periods in which the EEG or EOG recordings exceeded ±100 µV, as well as the first five periods of each block, were automatically excluded from the average. Standard-stimulus trials immediately following deviant-stimulus trials were also excluded from the averages. Individual ERPs were digitally band-pass filtered between 0.1 and 30 Hz. Difference waveforms were obtained by subtracting from each subject separately the ERPs elicited to standard stimuli from those elicited to the same stimuli when they were the deviant stimuli in the corresponding block. In
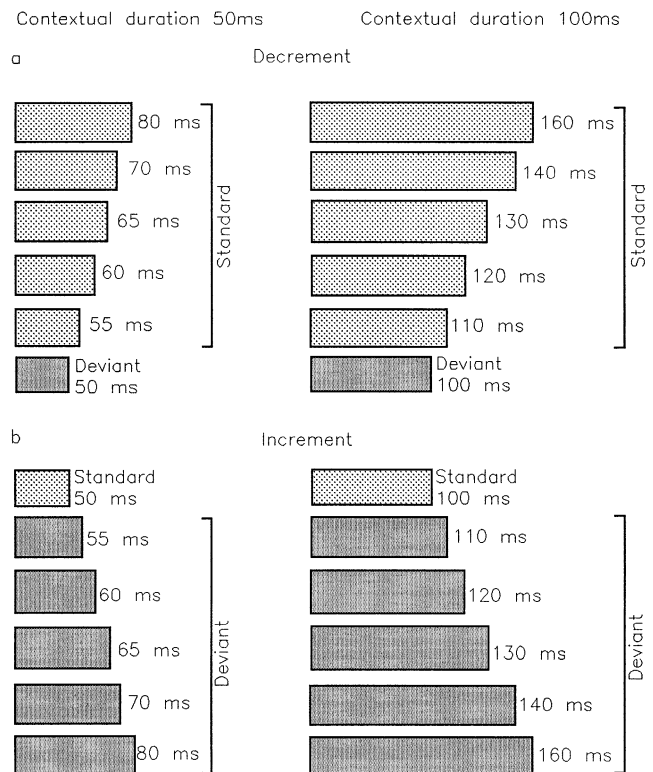

FIG. 1. Experimental conditions. Structure of the 20 blocks of stimuli for decrements and increments in contextual duration of 50 ms (left column) and 100 ms (right column). Light grey bars represent the standard stimuli delivered in each block, and dark grey bars represent the deviant stimuli in each block. Stimulus duration and category (standard or deviant) is specified at the right of each stimulus bar.

these difference waveforms, mean MMN amplitudes were calculated in the latency window between 150 and 250 ms, determined on the basis of the total average ERPs from each condition. The presence of the MMN in each condition was tested by two-tailed one-group t-tests of the difference mean amplitude calculated at Fz (where MMN was largest).

Reaction time (RT) of hits (responses to deviant stimuli) was computed for each subject and condition from the second session. Probabilities of hits and false alarms (responses to standard stimuli) were also recorded and transformed to z scores. The detection sensitivity measure d' was then calculated as the difference between the z scores for the probability of false alarms and hits (defined as $d' = z_{P(FA)} - z_{P(H)}$, where FA = false alarms and H = hits; Green & Swets, 1988; Gescheider, 1997). In this study, perfect detection (probability of hits 1.0, and probability of false alarms, 0.0) should give a d' value of 8.6, while chance responses (when the probabilities of hits and false alarms are equal, 0.5) should give no difference between the z scores, i.e. a value for d' of 0.0.

Results

Figure 2 illustrates the MMN responses recorded at the frontal Fz electrode. Duration deviance affected the mean amplitude of MMN at Fz (repeated measures ANOVA, $F_{4,36} = 4.8$, $P < 0.02$, $\varepsilon = 0.58$) irrespective of contextual duration ($F_{1,9} = 0.37$, $P = 0.6$) or decrement-increment in duration ($F_{1,9} = 0.3$, $P = 0.7$). MMN mean amplitude increased linearly as a function of the logarithm of the duration difference (Fig. 3a). MMN was reliably elicited by duration differences of 20% and above for both decrements and increments
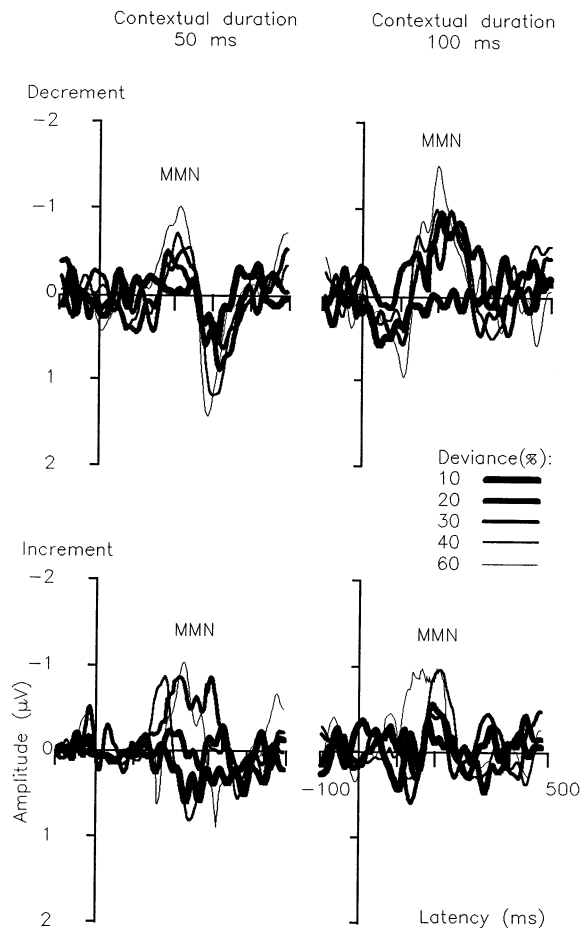
FIG. 2. Total mean ERP difference waveforms at Fz showing the MMN for decrements (upper panel) and increments (lower panel) in contextual duration of 50 ms (left column) and 100 ms (right column).

in 50 ms and 100 ms contextual duration (pooled $t(9) = 3.2$, $P < 0.02$; Fig. 2)

RT was affected by the deviance in duration ($F_{4,36} = 27.6$, $P < 0.0001$, $\varepsilon = 0.38$). However, contextual duration ($F_{1,9} = 3.22$, $P = 0.1$) and direction of deviance ($F_{1,9} = 1.68$, $P = 0.23$) failed to show any significant effect on this variable. RT decreased linearly as a function of deviance in the contextual duration of 50 ms (49 ms/ decrement $F_{1,48} = 12.5$, $P < 0.001$, and 59 ms/increment $F_{1,48} = 25.3$, $P < 0.0001$), and in the contextual duration of 100 ms (35 ms/ decrement $F_{1,48} = 10.5$, $P < 0.01$, and 42 ms/increment $F_{1,48} = 15.5$, $P < 0.001$).

Deviance affected the behavioural detection of subjects (repeated measures ANOVA, $F_{4,36} = 41.2$, $P < 0.0001$, $\varepsilon = 0.6$), which was unaffected by contextual duration ($F_{1,9} = 1.9$, $P = 0.5$) or direction of deviance ($F_{1,9} = 0.09$, $P = 0.8$). The behavioural detection measure d' increased linearly as a function of the logarithm of the duration difference (Fig. 3b), showing that the distance between the subject's gaussian distributions of false alarms and hit responses increased linearly, as the MMN amplitude, as a logarithmic function of the magnitude of deviance in sound duration (Fig. 3).

MMN mean amplitude and RT were highly correlated for decrements (Pearson's R = 0.98, two-tailed $P < 0.01$) and increments (Pearson's R = 0.89, two-tailed $P < 0.05$) in contextual duration of 50 ms, and for decrements (Pearson's R = 0.85, two-tailed $P < 0.05$) and increments (Pearson's R = 0.93, two-tailed $P < 0.03$) in the contextual duration of 100 ms.
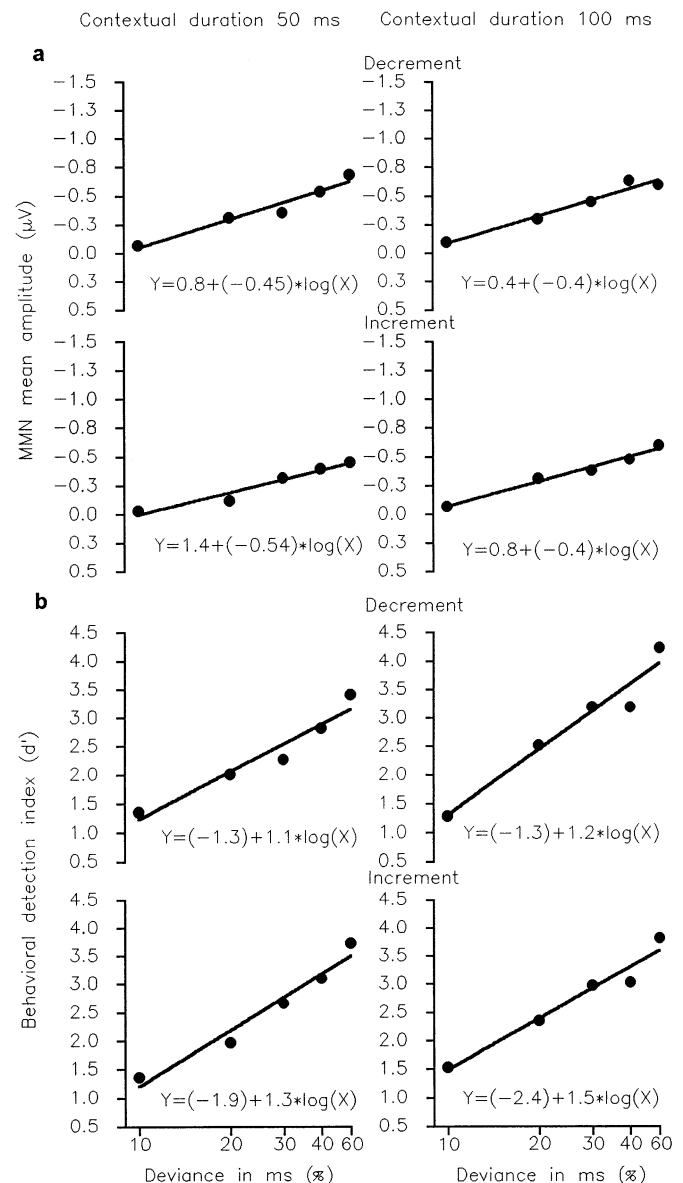

FIG. 3. Regression functions of the psychophysiological and psychophysical detection measures. (a) Mismatch negativity (MMN) mean amplitude as a function of the magnitude of deviance in sound duration for decrements (upper functions) and increments (lower functions). (b) d' values as a function of the magnitude of deviance in sound duration for decrements (upper functions) and increments (lower functions). Left column, contextual duration of 50 ms. Right column, contextual duration of 100 ms. The regression equations are shown below each regression line. Values on the abscissa are on a logarithmic scale.

MMN mean amplitude determined subject's behavioural discrimination as indicated by d' for decrements ($R^2 = 0.9$, one-way ANOVA $F_{1,48} = 6.8$, $P < 0.02$) and increments ($R^2 = 0.9$, $F_{1,48} = 13.2$, $P < 0.001$) in contextual duration of 50 ms, and for decrements ($R^2 = 0.8$, $F_{1,48} = 6.07$, $P < 0.02$) and increments ($R^2 = 0.9$, $F_{1,48} = 5.9$, $P < 0.02$) in contextual duration of 100 ms.

Discussion

The main finding of this study was that brain detection, as indicated by MMN amplitude, of decreasing or increasing changes in the

duration of a spectrally rich tone of either 50 or 100 ms predicted the accuracy of behavioural detection of these same changes.

In time perception studies, it remains to be determined whether the perceived duration of a sound depends on its perceived energy (Carbotte & Kristofferson, 1973; Allan, 1979), and the present results indicate that the cerebral detection of changes in sound duration of approximately a few tens of milliseconds is independent of sound energy, i.e. perceptual loudness, which was of course higher in the present experiment when the stimuli had a contextual duration of 100 ms. As brain detection underlies perception, these findings constitute a direct support to the behavioural results of Rammsayer & Lima (1991)**,** who found that duration discrimination of sounds less than 100 ms was independent of signal energy.

In the present experiment, the minimum change in stimulus duration necessary for eliciting a significant MMN was ± 10 ms with a contextual duration of 50 ms**,** and ± 20 ms with a contextual duration of 100 ms. That is, a deviation of 20% in stimulus duration was required to elicit an MMN in the two different contextual duration conditions**,** and for decrements and increments in contextual duration. This indicates that brain detection of duration deviance is independent of the direction of change, and that it depends on contextual information about duration, represented in this study by the duration of standard stimuli, rather than on physical length of the sounds, and hence that the behavioural perception of sound duration is related to the duration context in which these sounds are listened. The contextual nature of this detection system is in accordance with the contextual nature of speech perception (Celsis *et al.*, 1999), where, for instance, the duration of vowels is relative depending on variables such as the rate of speech, and listeners often identify vowels as a function of their relative durations (Lieberman & Blumstein, 1988).

In psychophysical detection, the index to measure the observer's sensitivity to discriminate between signal and noise is represented by the distance between the normal (z) distribution of false alarms (responses to noise) and hits (responses to signal) (Gescheider, 1997; Green & Swets, 1988). As the signal strength increases, the sensitivity raises, showing a logarithmic, referred to as psychometric function**,** which describes the behavioural detection curve (Parker & Newsome, 1998), as was observed in our behavioural data. The new finding of this study was that MMN, a neuroelectric brain response recorded noninvasively to any change occurring in the acoustic environment (Näätänen *et al.*, 1978; Näätänen, 1992), was elicited to changes in sound duration with a 'neurometric function' which predicted the psychometric function seen at behavioural level. As an example, the R^2 value of 0.9 found for decrements in the contextual duration of 50 ms indicates that an average increment of 0.45 μV in MMN mean amplitude predicted on average an increment of 1.1 units of d'.

It has been previously shown that the latency of MMN predicts behavioural speed (RT) in humans (Tiitinen *et al.*, 1994). In the present study RT decreased linearly with the increase in duration deviance and was highly correlated to MMN, supporting the fact that the neuroelectric activity involved in MMN generation is related to behavioural speed. RT, however, depends on a decision making process, so that speed may be preferred at the expenses of making more false alarms, or, on the contrary, precision may be preferred at the expense of slowing down the responding process. Therefore, RT measures do not indicate the subject's confidence in responding to target stimulation. As d' provides an estimate of the confidence of behavioural discrimination independent of response criterion variability, the present results show how the accuracy of conscious perception of sound duration is determined by the accuracy of the brain to represent such duration.

Abbreviations

A/D rate, analogue-to-digital conversion rate; ANOVA, analysis of variance; ERP, event-related potential; MMN, mismatch negativity; RT, behavioural speed.

References

Allan, L.G. (1979) The perception of time. *Percept. Psychophys.*, **26**, 340–354.

Aaltonen, O., Eerola, O., Lang, A.H., Uusipaikka, E. & Tuomanien, J. (1994) Automatic discrimination of phonetically relevant and irrelevant vowel parameters as reflected by mismatch negativity. *J. Acoust. Soc. Am.*, **96**, 1489–1493.

Carbotte, R.M. & Kristofferson, A.B. (1973) On energy-dependent cues in duration discrimination. *Percept. Psychophys.*, **14**, 501–505.

Celsis, P., Doyon, B., Boulanouar, K., Pastor, J., Démonet, J.-F. & Nespoulous, J.-L. (1999) ERP correlates of phoneme perception in speech and sound contexts. *Neuroreport*, **10**, 1523–1527.

Cheour, M., Ceponiene, R., Lehtokoski, A., Luuk, A., Allik, J., Alho, K. & Näätänen, R. (1998) Development of language-specific phoneme representations in the infant brain. *Nature Neurosci.*, **1**, 351–353.

Dehaene-Lambertz, G. (1997) Electrophysiological correlates of categorial phoneme perception in adults. *Neuroreport*, **8**, 919–924.

Ford, J.M. & Hillyard, S.A. (1981) Event-related potentials, ERPs, to interruptions of steady rhythm. *Psychophysiology*, **18**, 322–330.

Gescheider, G.A. (1997) *Psychophysics: The fundamentals.* 3rd edn. Lawrence Erlbaum, Mahwah, NJ.

Green, D.M. & Swets, J.A. (1988). *Signal Detection Theory and Psychophysics*. Peninsula Publishing, Los Altos, CA, USA.

Hari, R., Joutsiniemi, S.-L., Hämäläinen, M. & Vilkman, V. (1989) Neuromagnetic responses of human auditory cortex to interruptions in a steady rhythm. *Neurosci. Lett.*, **99**, 164–168.

Joutsiniemi, S.-L., Ilvonen, T., Sinkkonen, J., Huotilainen, M., Tervaniemi, M., Lehtokoski, A., Rinne, T. & Näätänen, R. (1998) The mismatch negativity for duration decrement of auditory stimuli in healthy subjects. *Electroencephalogr. Clin. Neurophysiol.*, **108**, 154–159.

Kaukoranta, E., Sams, M., Hari, R., Hämäläinen, M. & Näätänen, R. (1989) Reactions of human auditory cortex to changes in tone duration. *Hear. Res.*, **41**, 15–22.

Kraus, N., McGee, T.J., Carrell, T. & Sharma, A. (1995) Neurophysiologic bases of speech discrimination. *Ear Hear.*, **16**, 19–37.

Kraus, N., McGee, T.J. & Koch, D.B. (1998) Speech sound representation, perception, and plasticity: a neurophysiologic perspective. *Audiol. Neurootol.*, **3**, 168–182.

Lang, A.H., Nyrke, T., Ek, M., Aaltonen, O., Raimo, Y. & Näätänen, R. (1990) Pitch discrimination performance and auditory event-related potentials. In Brunia, C.H.M., Gaillard, A.W.K. & Kok, A., (eds), *Psychophysiological Brain Research*, Vol. I. Tilburg University Press, Tilburg, pp. 294–298.

Lieberman, P. & Blumstein, S.E. (1988). *Speech Physiology, Speech Perception, and Acoustic Phonetics*. Cambridge University Press, Cambridge, NY.

Lounasmaa, O.V., Hari, R., Joutsiniemi, S.-L. & Näätänen, R. (1989) Multi-SQUID recordings of human cerebral magnetic fields may give information about memory processes. *Europhys. Lett.*, **9**, 603–608.

Lyytinen, H., Blomberg, A.-P. & Näätänen, R. (1992) Event-related potentials and autonomic responses to a change in unattended auditory stimuli. *Psychophysiology*, **29**, 523–534.

Näätänen, R., Gaillard, A.W.K. & Mäntysalo, S. (1978) Early selective-attention effect on evoked potential reinterpreted. *Acta Psychologica*, **42**, 313–329.

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

Näätänen, R., Paavilainen, P. & Reinikainen, K. (1989b) Do event-related

potentials to infrequent decrements in duration of auditory stimuli demonstrate a memory trace in man? *Neurosci. Lett.*, **107**, 347–352.

Näätänen, R. (1990) The role of attention in auditory information processing as revealed by event-related brain potentials. *Behav. Brain Sci.*, **13**, 199–290.

Näätänen, R. (1992) *Attention and brain function.* Lawrence Erlbaum, Hillsdale, NJ.

Näätänen, R., Schröger, E., Karakas, S., Tervaniemi, M. & Paavilainen, P. (1993) Development of a memory trace for a complex sound in the human brain. *Neuroreport*, **4**, 503–506.

Näätänen, R. & Alho, K. (1997) Mismatch Negativity – The measure for central sound representation accuracy. *Audiol. Neurootol.*, **2**, 341–353.

Näätänen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huotilainen, M., Livonen, A., Vainio, M., Alku, P., Ilmoniemi, R.J., Luuk, A., Allik, J., Sinkkonen, J. & Alho, K. (1997) Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature*, **385**, 432–434.

Nordby, H., Hammerborg, D., Roth, W.T. & Hugdahl, K. (1988) Event-related potentials to time-deviant and pitch-deviant tones. *Psychophysiology*, **25**, 249–261.

Paavilainen, P., Jiang, D., Lavikainen, J. & Näätänen, R. (1993) Stimulus duration and the sensory memory trace: an event-related potential study. *Biol. Psychol.*, **35**, 139–152.

Parker, A.J. & Newsome, W.T. (1998) Sense and the single neuron: Probing the physiology of perception. *Annu. Rev. Neurosci.*, **21**, 227–277.

Phillips, D.P. (1993) Representation of acoustic events in the primary auditory cortex. *J. Exp. Psychol. Hum. Percept. Perform.*, **19**, 203–216.

Rammsayer, T.H. & Lima, S.D. (1991) Duration discrimination of filled and empty auditory intervals: Cognitive and perceptual factors. *Percept. Psychophys.*, **50**, 565–574.

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

Schröger, E. & Wolff, C. (1996) Mismatch response of the human brain to changes in sound location. *Neuroreport*, **7**, 3005–3008.

Tervaniemi, M., Ilvonen, T., Karma, K., Alho, K. & Näätänen, R. (1997) The musical brain: Brain waves reveal the neurophysiological basis of musicality. *Neurosci. Lett.*, **226**, 1–4.

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