Separate Neural Processing of Timbre Dimensions in Auditory Sensory Memory

Anne Caclin[1,2,3,4], Elvira Brattico[5,6], Mari Tervaniemi[5,6], Risto Näätänen[5,6], Dominique Morlet[2,3,4], Marie-Hélène Giard[2,3,4], and Stephen McAdams[1,7]

Abstract

■ Timbre is a multidimensional perceptual attribute of complex tones that characterizes the identity of a sound source. Our study explores the representation in auditory sensory memory of three timbre dimensions (acoustically related to attack time, spectral centroid, and spectrum fine structure), using the mismatch negativity (MMN) component of the auditory event-related potential. MMN is elicited by a discriminable change in a sound sequence and reflects the detection of the discrepancy between the current stimulus and traces in auditory sensory memory. The stimuli used in the present study were carefully controlled synthetic tones. MMNs were recorded after changes along each of the three timbre dimensions and their combinations. Additivity of unidimensional MMNs and dipole modeling results suggest partially separate MMN generators for different timbre dimensions, reflecting their mainly separate processing in auditory sensory memory. The results expand to timbre dimensions a property of separation of the representation in sensory memory that has already been reported between basic perceptual attributes (pitch, loudness, duration, and location) of sound sources. ■

INTRODUCTION

Sounds are characterized perceptually by a number of attributes: their pitch, loudness, perceived duration, perceived position, and timbre. Timbre is defined as the set of perceptual attributes that allows one to distinguish among tones having the same pitch, loudness, and duration (American National Standards Institute, 1973). For example, each musical instrument has a particular timbre, and we can recognize that the same note played on a piano or a clarinet has been generated by different sources. This example highlights the relationship between timbre perception and sound-source recognition (Handel, 1995; McAdams, 1993). Despite this fundamental link, timbre perception remains less precisely understood than that of the other auditory attributes, both at the behavioral and neurophysiological levels (Menon et al., 2002; Hajda, Kendall, Carterette, & Harschberger, 1997). The difficulty in timbre studies arises from the variety of possible acoustic parameters that might underlie timbre perception.

Indeed, contrary to other auditory attributes that appear to rely principally on a single characteristic of the acoustic waveform (such as fundamental frequency for pitch height or intensity for loudness), timbre has been shown to rely on several acoustic dimensions and is therefore described as a multidimensional perceptual attribute. Psychophysical evidence for the multidimensional nature of timbre comes from studies of timbre identification, ratings on verbal scales, and dissimilarity ratings (see Hajda et al., 1997). The experiment reported here examines the multidimensional nature of timbre at the neurophysiological level, using the mismatch negativity (MMN) component of the auditory event-related potential (ERP).

Behavioral Studies of Timbre Perception

Early studies on timbre (Berger, 1964; Saldanha & Corso, 1964) assessed instrument identification by using recorded sounds and transformed versions of these sounds, which consisted in removing parts of the sounds or playing the sounds backward. Instrument recognition was altered under those conditions, suggesting that both spectral and temporal characteristics of the sounds were relevant for the task (see also Hajda, 1999). Verbal rating studies (Kendall & Carterette, 1993a, 1993b; von Bismarck, 1974) further showed that timbre relations could not be summarized along a single verbal axis, such as ''not brilliant'' to "brilliant" or ''not nasal'' to "nasal."

Timbre perception has been most often analyzed by applying multidimensional scaling models to dissimilarity

[1]Institut de Recherche et Coordination Acoustique/Musique (STMS-IRCAM-CNRS), Paris, France, [2]INSERM U280, Lyon, France, [3]Institut Fédératif des Neurosciences, Lyon, France, [4]University of Lyon 1, France, [5]University of Helsinki, Finland, [6]Helsinki Brain Research Center, Finland, [7]McGill University, Montréal, Canada

ratings (e.g., McAdams, Winsberg, Donnadieu, De Soete, & Krimphoff, 1995; Grey, 1977). In such an experiment, participants are requested to rate the dissimilarity for every pair of *n* sounds having different timbres. Multi-dimensional scaling techniques (e.g., Winsberg & De Soete, 1993; Carroll & Chang, 1970) then reveal, on the basis of these dissimilarities, a low-dimensional (usually Euclidean) space in which the distance between objects matches with the dissimilarity ratings as closely as possible. These multidimensional scaling studies of timbre dissimilarity ratings have shown that timbre relationships are well described in a multidimensional space (with two to four dimensions). Such a perceptual space, resulting only from participants' ratings, is known as a *timbre space*. A psychophysical interpretation for each dimension of this space is suggested by means of correlations between perceptual coordinates and parameters extracted from the acoustic waveforms.

Depending on the set of sounds being used and also on the group of listeners, the number of dimensions of the timbre space varies, as do their potential acoustic correlates. These acoustic parameters fall into three broad categories: temporal, spectral, and spectrotemporal descriptors. Consistently across studies, attack time (more precisely its logarithm) and spectral centroid (weighted mean of spectrum energy) have been reported as correlates of timbre space dimensions (see, e.g., Marozeau, de Cheveigné, McAdams, & Winsberg, 2003; Samson, Zatorre, & Ramsay, 1997; McAdams et al., 1995; Krumhansl, 1989; Grey, 1977). Attack time separates impulsive from sustained instruments, and spectral centroid reflects the global spectral distribution. Other parameters have been less systematically reported: spectrum fine structure (Krimphoff, McAdams, & Winsberg, 1994), spectral spread (Marozeau et al., 2003), or spectral flux (McAdams et al., 1995), the latter being a measure of the variation of the spectrum over time. Altogether, these studies confirm that timbre indeed comprises several perceptual dimensions, with each of these dimensions corresponding to a single acoustic parameter. The exact number of dimensions needed to account for timbre relationships, as well as the most appropriate acoustic descriptors for timbre dimensions, remain open questions. We recently undertook a confirmatory analysis of the relevance of several acoustic parameters for timbre perception (Caclin, McAdams, Smith, & Winsberg, 2005). Using carefully controlled synthetic tones, we confirmed the perceptual relevance of attack time, spectral centroid, and spectrum fine structure (see Samson et al., 1997, and Miller & Carterette, 1975, for a similar confirmatory analysis). Each of these parameters accounted for one dimension of a 3-D timbre space. Conversely, we have shown in the same study that a spectrotemporal parameter (spectral flux) was less salient than these three parameters and that spectral flux saliency was context dependent.

Given this multidimensional account of timbre, one might ask whether the various aspects of the sound that underlie timbre perception are processed separately or holistically in the auditory system. We have chosen to address this issue at the level of auditory sensory memory on theoretical and practical grounds that are discussed below.

Representation of Auditory Attributes in Sensory Memory

Auditory events leave a transient memory trace, lasting up to 10–20 sec, in the human auditory system known as auditory sensory (or echoic) memory (see Cowan, 1984). Behavioral research using in particular the interference paradigm introduced by Deutsch (1970) has shown that different auditory attributes are represented separately in auditory sensory memory as, for example, pitch and timbre (Krumhansl & Iverson, 1992; Semal & Demany, 1991) or pitch and loudness (Clément, Demany, & Semal, 1999). At the neurophysiological level, auditory sensory memory organization has mostly been studied using the MMN component of the ERP (Näätänen, Gaillard, & Mantysalo, 1978) and its magnetic counterpart denoted MMNm.

MMN as a Probe of Auditory Sensory Memory Organization

MMN is a brain response elicited by virtually any detectable violation of an acoustic regularity in a sound sequence. It is typically recorded after infrequent deviant tones inserted in sequences of sounds composed primarily of the same standard tone. A number of arguments converge on the hypothesis that MMN reflects the neuronal mismatch between the deviant stimulus and the trace in auditory sensory memory of the events that have just happened (see Näätänen & Winkler, 1999, for a discussion). It is therefore possible to use MMN as a probe to uncover the organization of the memory traces upon which it is based (as reviewed in Ritter, Deacon, Gomes, Javitt, & Vaughan, 1995).

MMN has been widely used to study normal and abnormal auditory processing, and this interest stems in particular from the fact that MMN can be elicited even without attentional focus. MMN can indeed be recorded while the participant is not paying attention to the sounds and is engaged in another activity such as reading or watching a silent movie. It can even be recorded during sleep (Loewy, Campbell, & Bastien, 1996). But the most important feature of the MMN in the present context is that the memory traces upon which it is based stand in close relationship with the actual percept (see Näätänen & Winkler, 1999). For example, an MMN is elicited by a subtle change in a sound sequence only when the participants have learned to discriminate

the deviant from the standard (Näätänen, Schröger, Karakas, Tervaniemi, & Paavilainen, 1993). Furthermore MMN amplitude and latency are correlated with the degree of perceived dissimilarity between the standard and the deviant (Toiviainen, Tervaniemi, Louhivuori, Saher, Huotilainen, & Näätänen, 1998; Tiitinen, May, Reinikainen, & Näätänen, 1994), as well as with hit rate and reaction time in a change detection task (Novitski, Tervaniemi, Huotilainen, & Näätänen, 2004).

MMN and Auditory Attributes

The processing of basic auditory attributes and, in particular, pitch, loudness, perceived duration, and location, has been analyzed by using the MMN. The central question is to determine whether these attributes are processed separately or holistically (i.e., as a gestalt) in the auditory system. A number of studies using the MMN support the view that in auditory sensory memory, auditory features are represented separately. A first line of evidence arises from studies of the location of intracerebral MMN generators. In general, it is known that MMN relies on the activation of bilateral sources in the supratemporal auditory cortex (Alho, 1995). Additional generators have been described particularly in the frontal areas (e.g., Giard, Perrin, Pernier, & Bouchet, 1990). The exact location of the temporal sources of the MMN depends on the particular acoustic feature that distinguishes the standard from the deviant, at least for frequency, intensity, and duration (Rosburg, 2003; Frodl-Bauch, Kathmann, Moller, & Hegerl, 1997; Giard et al. 1995). Separate processing of different features is further supported by an additivity of the responses to unidimensional changes (Wolff & Schröger, 2001; Schröger, 1995b). This means that the MMN elicited by bidimensional deviants (e.g., frequency and location, see Schröger, 1995b) is well predicted by the sum of the MMNs following the corresponding unidimensional deviants, except at frontal sites for frequency and intensity (Wolff & Schröger, 2001). This additivity even applies to deviations along two acoustic features that both contribute to perceived auditory location: interaural time and level differences (Schröger, 1995a; see also Winkler, Tervaniemi, Schröger, Wolff, & Näätänen, 1998). However, deviants with three deviating features elicit a smaller MMN than predicted by the sum of the unidimensional deviants (Paavilainen, Valppu, & Näätänen, 2001), revealing complex attention-dependent interactions in the processing of simultaneously deviating features.

A variety of experimental manipulations have been used to assess whether basic auditory features were processed independently, as suggested by studies of MMN generator location and MMN additivity, or as gestalts. For example, when multiple standards are used that differ from one another along acoustic features other than the feature of interest, an MMN is observed, suggesting that the feature of interest is processed separately from the irrelevant features (Gomes, Ritter, & Vaughan, 1995). Furthermore, when presenting several deviants in a row, MMN amplitude is not reduced for the second or third deviant if they differ from the standards along different dimensions (Nousak, Deacon, Ritter, & Vaughan, 1996), whereas it is reduced when the successive deviants are identical (Sams, Alho, & Näätänen, 1984). Finally, MMN amplitude reflects the probability of a particular feature rather than the probability of a combination of features (Deacon, Nousak, Pilotti, Ritter, & Yang, 1998). All these arguments converge on the idea that the strength of the representation of each acoustic feature in auditory sensory memory evolves independently from that of others, suggesting independent storage of these features.

Nevertheless, gestalt storage is also possible in sensory memory, as revealed by the fact that an MMN is recorded after deviants differing from the standards only by the conjunction of two features (Sussman, Gomes, Nousak, Ritter, & Vaughan, 1998; Gomes, Bernstein, Ritter, Vaughan, & Miller, 1997). There were three standards intermixed in these studies, and individual features of the deviant were already present in one of the three standards, but not their conjunction. This result applies to timbre (taken as a whole) as well, as conjunction deviants involving timbre and pitch elicit an MMN (Takegata et al., 2005). Although it might seem at first sight that these results are hardly reconcilable with separate processing of the very same acoustic dimensions, it is possible to propose mechanisms that account for both types of results (Sussman et al., 1998; Ritter et al., 1995). The idea is that different levels of information (e.g., basic auditory attributes or higher level grouping properties) are retrieved from sensory analysis and stored in auditory sensory memory so that the deviance detection system underlying MMN generation can operate on these different types of information.

In the experiments reported here, we investigated whether timbre dimensions exhibit properties of processing separability in auditory sensory memory, as observed for intensity, frequency, duration, and location, using the MMN as a probe.

Neural Basis of Timbre Perception

Neuropsychological and brain imaging studies of timbre perception have focused on delineating the brain areas involved in timbre processing, and in particular have tested whether there is a hemispheric dominance for this processing. Timbre judgments have been investigated in patients with unilateral lesions of the temporal lobe. An impairment in timbre discrimination tasks has been reported after right temporal lobe lesions (Samson & Zatorre, 1994; Milner, 1962), hence the longstanding idea of a right-hemispheric dominance in timbre processing. But this dominance is likely to be

mostly the consequence of the task being used because abnormal timbre spaces are observed in dissimilarity rating studies after temporal lesions in both hemispheres (Samson, Zatorre, & Ramsay, 2002). Furthermore, in healthy participants, passive listening to different timbres leads to bilateral activations in temporal auditory regions (Caclin, Menon, Krasnow, Mackenzie, Smith, & McAdams, 2003; Menon et al., 2002).

A few electroencephalography (EEG) or magneto-electroencephalography (MEG) studies have focused on the effect of musical training on timbre processing (Pantev, Roberts, Schulz, Engelien, & Ross, 2001; Pantev, Oostenveld, Engelien, Ross, Roberts, & Hoke, 1998; Crummer, Walton, Wayman, Hantz, & Frisina, 1994). In a timbre discrimination task, the rapidity of timbre processing as measured with the P3 latency was faster in musicians, especially in those with absolute pitch (Crummer et al., 1994). Using MEG recordings, Pantev et al. (1998) have shown that N1m generators were more strongly activated for piano tones than pure tones in musicians but not in nonmusicians. Furthermore, this N1m enhancement was larger for the timbre of the instrument practiced by the listener (Pantev et al., 2001). These studies indicate that timbre processing can be modulated by musical experience, but shed little light on timbre processing itself.

The processing of single timbre dimensions in auditory sensory memory has been analyzed using the MMN (e.g., rise time in Lyytinen, Blomberg, & Näätänen, 1992; or spectral centroid in Toiviainen et al., 1998). Tervaniemi, Winkler, and Näätänen (1997) have found that timbres are categorized preattentively, as an MMN was elicited by pure tones presented among harmonic, complex tones with different missing fundamentals. This result implies that at the level of auditory sensory memory, all the missing fundamental tones were grouped into a "rich" timbre category that was distinguished from the "pure" category. Toiviainen et al. (1998) have explored MMN after changes in spectral centroid. The MMN amplitude and latency were correlated with the degree of perceived dissimilarity between the standard and different types of deviants. Altogether, these studies reveal that information concerning the categorical and continuous nature of timbre is encoded in sensory memory, but the relationships between the processing of different timbre dimensions remain so far unexplored.

Rationale of the Study

The present study aims to determine whether timbre dimensions are processed independently in auditory sensory memory, using the MMN as a probe. Given the complexity of the acoustic underpinnings of timbre perception, a major issue was to construct an appropriate and fully controlled stimulus set. It involves two main choices: first, selecting acoustic parameters, and second,

choosing the values to be taken by each parameter. Synthetic tones were used to ensure complete control of the stimuli. We selected three acoustic parameters on the basis of a previous psychophysical analysis of acoustic correlates of timbre dimensions (Caclin et al., 2005): attack time, spectral centroid, and spectrum fine structure, the latter being modeled by a selective attenuation of the even harmonics. Our psychophysical work has indeed revealed that these three parameters accounted for three orthogonal dimensions of a perceptual timbre space (Caclin et al., 2005). The values along each acoustic dimension were chosen according to a dissimilarity criterion, as MMN amplitude and latency are correlated with perceived dissimilarity (e.g., Tiitinen et al., 1994). Uni-, bi-, and tridimensional timbre changes were tested. We hypothesized that separate processing of these three timbral dimensions in auditory sensory memory would be revealed by anatomically distinct MMN generators for each dimension as well as by an additivity of unidimensional MMNs.

METHODS

Participants

Nineteen paid listeners (aged 19–54 years, all right-handed, 15 women) were recruited to participate in this experiment. Six of them had practiced music regularly. Participants were naive as to the purpose of the experiment.

Stimuli

Eight stimuli were chosen so as to form a cube with edges of equal perceived dissimilarity within the calibrated timbre space (see Figure 1 and Table 1). The choice was made on the basis of a reanalysis of the data of the last experiment reported in Caclin et al. (2005), which was a dissimilarity rating experiment with synthetic tones similar to those used in the present study. A standard tone was chosen arbitrarily with a sharp attack (ATT), a low spectral center of gravity (SCG), and a regular spectrum without attenuation of the even harmonics (EHA). Starting from this standard tone, three unidimensional (1-D) deviants were constructed by modifying one of the three acoustic parameters of interest in such a way that these three unidimensional deviants were equally dissimilar relative to the standard. Finally three bidimensional (2-D) deviants and one tridimensional (3-D) deviant were obtained by combining the values chosen for the standard and the unidimensional deviants along the three acoustic dimensions.

All tones had 311-Hz fundamental frequency (E♭4) and were constructed by additive synthesis (20 harmonics). The temporal envelope of the sounds was composed of a linear attack, followed by a plateau and an exponential decay (see Table 1). The attack time
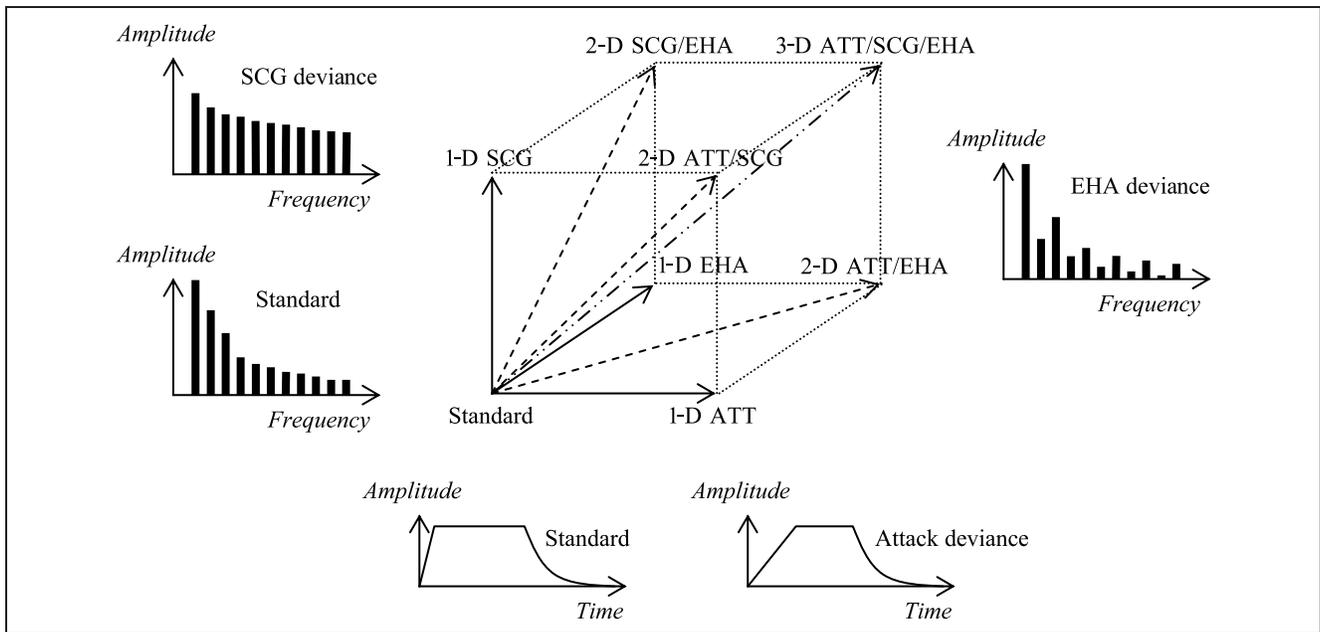
Figure 1. Stimulus construction. All stimuli have the same pitch, loudness, and perceived duration but vary along three timbral dimensions. ATT = attack time; SCG = spectral center of gravity; EHA = even-harmonic attenuation. The edges of the stimulus cube represent equal perceived dissimilarity. Arrows represent the seven timbre changes from the standard, with solid lines for unidimensional (1-D) deviants and dashed lines for multidimensional (2-D and 3-D) deviants.

could be 15 or 200 msec, and the duration of the plateau was adjusted to keep perceived sound duration constant (see Caclin et al., 2005). The global shape of the spectrum was fixed by a power relationship between amplitude and harmonic rank, which determined the value of the SCG (see Caclin et al., 2005, for details). SCG could be 3 or 4.3 in harmonic rank units. Finally, the local shape of the spectrum was adjusted with a selective attenuation of even harmonics (EHA). Again, this parameter could take two values, 0 or −6.4 dB. The levels were corrected to keep loudness constant across all sounds (see Table 1).

These eight stimuli were presented in two types of sequences with a constant 1-sec stimulus onset asynchrony: test sequences and equiprobable sequences. In test sequences, the standard was presented with a probability of 79% and each of the seven deviants with a probability of 3%, with at least two intervening standards between two consecutive deviants. In equiprobable sequences, each of the eight sounds was presented with an equal probability of 12.5%. Test sequences comprised a total of 6000 sounds, leading to 180 presentations of each type of deviant. Equiprobable sequences comprised 1440 sounds (180 of each of the eight stimuli)

Table 1. Parameters Describing the Stimuli

Timbre	F0 (Hz)	ATT (msec)	Plateau (msec)	Decay (msec)	SCG (Harmonic Rank)	EHA (dB)	Amplitude Correction (dB)
Standard	311	15	400	200	3	0	0
1-D ATT	311	<u>200</u>	247	200	3	0	0
1-D SCG	311	15	400	200	<u>4.3</u>	0	−2.5
1-D EHA	311	15	400	200	3	−<u>6.4</u>	0
2-D ATT/SCG	311	<u>200</u>	247	200	<u>4.3</u>	0	−2.5
2-D ATT/EHA	311	<u>200</u>	247	200	3	−<u>6.4</u>	0
2-D SCG/EHA	311	15	400	200	<u>4.3</u>	−<u>6.4</u>	−2.5
3-D ATT/SCG/EHA	311	<u>200</u>	247	200	<u>4.3</u>	−<u>6.4</u>	−2.5

EHA is given relative to the odd harmonics. Amplitude correction was applied after equalization of RMS amplitude. Deviating features are <u>underlined</u>. F0 = fundamental frequency; ATT = attack time; SCG = spectral center of gravity; EHA = even-harmonic attenuation.

and were always presented in the middle of the experiment. New sequences were created for each participant. Participants watched a silent subtitled movie during the recordings and were instructed not to pay attention to the sounds. The entire experiment, including electrode positioning, lasted around 3 hr.

EEG Recordings

EEG was acquired continuously using 63 Ag/AgCl electrodes inserted in a cap. The reference electrode was placed at the tip of the nose. Signals were amplified (Neuroscan Inc., El Paso, TX), sampled at 250 Hz and filtered online (DC to 50 Hz).

For each participant, each of the eight sounds, and each of the two types of sequences, ERPs were computed at each electrode over a 1000-msec temporal window from 200 msec prestimulus to 800 msec poststimulus. The 100-msec period preceding the onset of the sound was taken as baseline. Artifact rejection consisted of removing the trials with variation of more than 150 μV or less than 2 μV in a sliding 100-msec window at any of the 63 electrodes, so that epochs with rapid ocular artifacts or signal saturation were excluded. The data from four participants (including one musician) were removed, as less than 100 single trials were averaged for at least one type of sound. ERPs were digitally filtered using a 2- to 25-Hz band-pass filter (slopes, 12 dB/octave), and only the −100- to +700-msec time window was retained for analysis. Finally, grand-average ERPs (mean ERPs across the 15 remaining participants) were computed.

Data Analysis

For each type of deviant, MMN was sought in the subtraction between the ERPs to this sound when presented in the test versus in equiprobable sequences (i.e., deviant − corresponding equiprobable). As opposed to the traditional deviant minus standard subtraction, this procedure avoids contamination by differences in the sensory processing of different types of sounds because the only difference between the two ERPs compared is the status of the sound in the sequence (see Sinkkonen & Tervaniemi, 2000, for a discussion). It also allows minimization of the contribution of refractory neurons to the MMN (see Jacobsen, Schröger, Horenkamp, & Winkler, 2003).

Amplitudes and Latencies of Unidimensional MMNs

MMN was defined as the most negative wave at frontal sites in the 100- to 270-msec time window in the deviant minus equiprobable ERPs. For each type of deviant and each participant, MMN latency and amplitude were measured as the mean latency and amplitude of the most negative wave at nine electrodes (Fz, Cz, an electrode at the midline between Fz and Cz, and the three adjacent electrodes on the right and left hemispheres). Amplitudes and latencies of the three unidimensional MMNs were compared by using one-way analyses of variance, with Timbre Dimension as a three-level within-subject factor. Where appropriate, the results were corrected by using the Greenhouse–Geisser procedure, and epsilon and corrected p are reported. Significant effects were explored by using Tukey HSD post hoc tests.

Additivity of Unidimensional MMNs

To assess the additivity hypothesis, we compared MMNs for bi- or tridimensional deviants with the sum of the corresponding unidimensional MMNs in the 80- to 270-msec time window. For each of the 2-D cases and the 3-D case, we computed for each electrode and each time point a paired-value t test. To correct for multiple comparisons, a difference was considered significant only when observed for at least six consecutive time points corresponding to 20 msec (see Guthrie & Buchwald, 1991, for a discussion).

Dipole Modeling

Possible anatomical differences between the intracerebral generators of MMNs for changes along different timbre dimensions were sought by using equivalent current dipoles (ECDs) to model the scalp potential distributions in the grand-average ERPs. A three-layer spherical head model was used, with a nonlinear iterative procedure (Marquardt algorithm) to estimate dipole parameters (see also Giard et al., 1995; Giard, Perrin, Echallier, Thévenet, Froment, & Pernier, 1994). For each of the three unidimensional deviants, we modeled MMN with two symmetrical dipoles. Distinct orientation of the dipoles in the two hemispheres was allowed. Location and orientation of the dipoles were constrained to be constant across time, whereas their amplitudes were time variant. For each of these three MMNs, modeling was performed in the time window where the deviant and equiprobable ERPs were significantly different from each other at Fz or at either of the mastoids.

Once the model was selected, using 400 Monte Carlo simulations, we estimated a confidence interval for the position and orientation of the dipoles (see also Giard et al., 1995). The 99% confidence interval was defined as the ellipsoid in which 99% of the Monte Carlo samples fell. Direct comparison of these confidence intervals across timbre dimensions allows us to determine whether the dipoles fitting the different MMNs are located and/or oriented similarly or not. More specifically, the locations of two dipoles were considered different when their
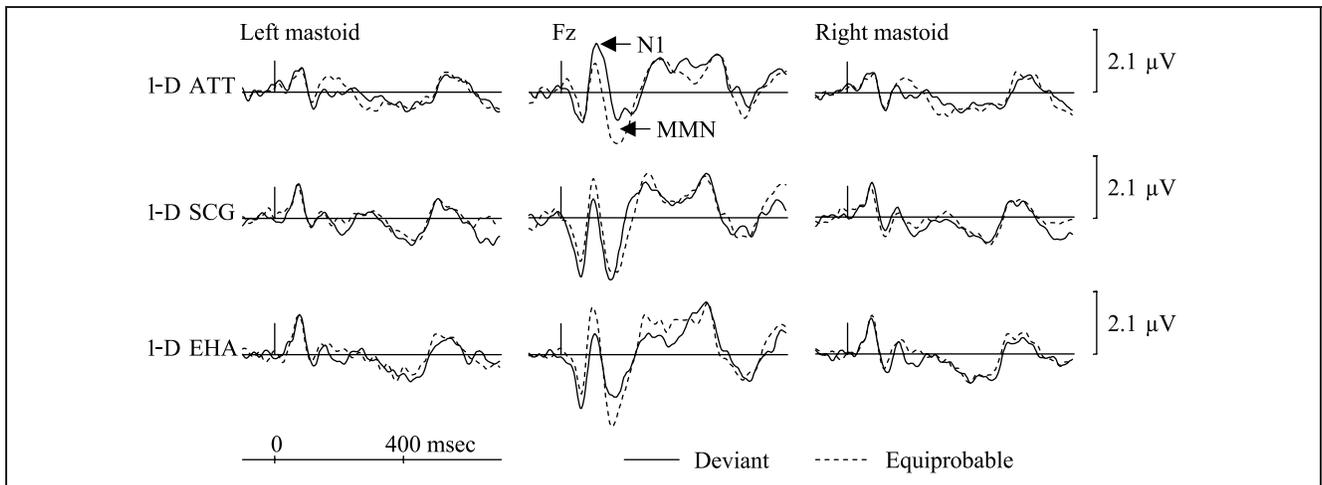
Figure 2. Grand-average ERPs for the three unidimensional deviants (solid lines) and their equiprobable counterparts (dashed lines) at Fz and the mastoids. Time and amplitude scales are the same for all curves. ERP negativity is up (as in Figures 3 and 4).

confidence intervals did not overlap. The same criterion was applied for dipole orientation comparisons.

RESULTS

Grand-average ERPs obtained for equiprobable sounds and unidimensional deviants are presented in Figure 2. Their subtraction (deviant − equiprobable) can be seen in Figure 3 (unidimensional deviants) and Figure 4 (multidimensional deviants). For each deviant, we ob-

served in the subtraction ERP a negative wave at frontal sites, peaking around 150–200 msec (see Table 2), accompanied by a polarity reversal at the mastoids. This topography and latency range is the one expected for the MMN. We can thus conclude that a change along any of the three timbre dimensions (and their combinations) investigated here elicits an MMN. A closer look at the results reveals that for most of the seven MMNs recorded in this study, multiple peaks are observed, even for unidimensional deviants (see Figure 3), suggesting that multiple processes are actually in operation.
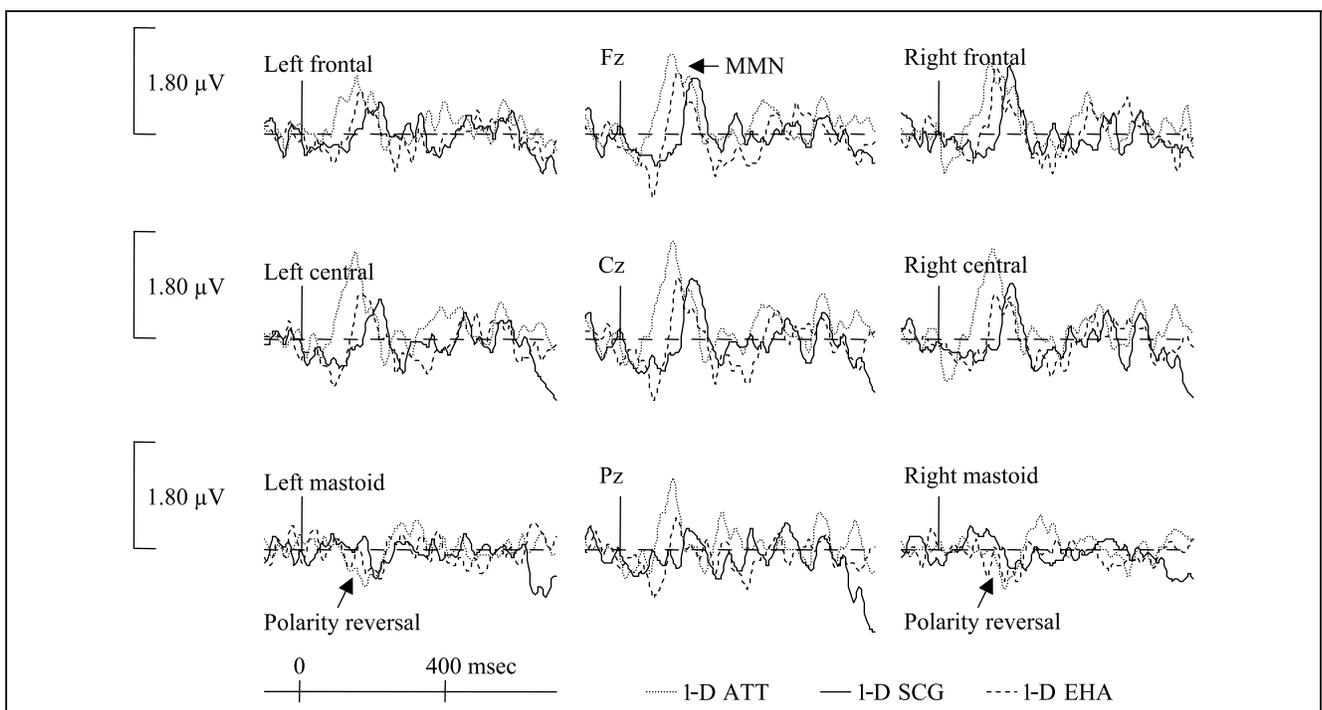
Figure 3. Subtraction waves (deviant minus equiprobable ERPs) showing the MMNs for the three unidimensional timbre changes at nine electrodes. Time and amplitude scales are the same for all curves.
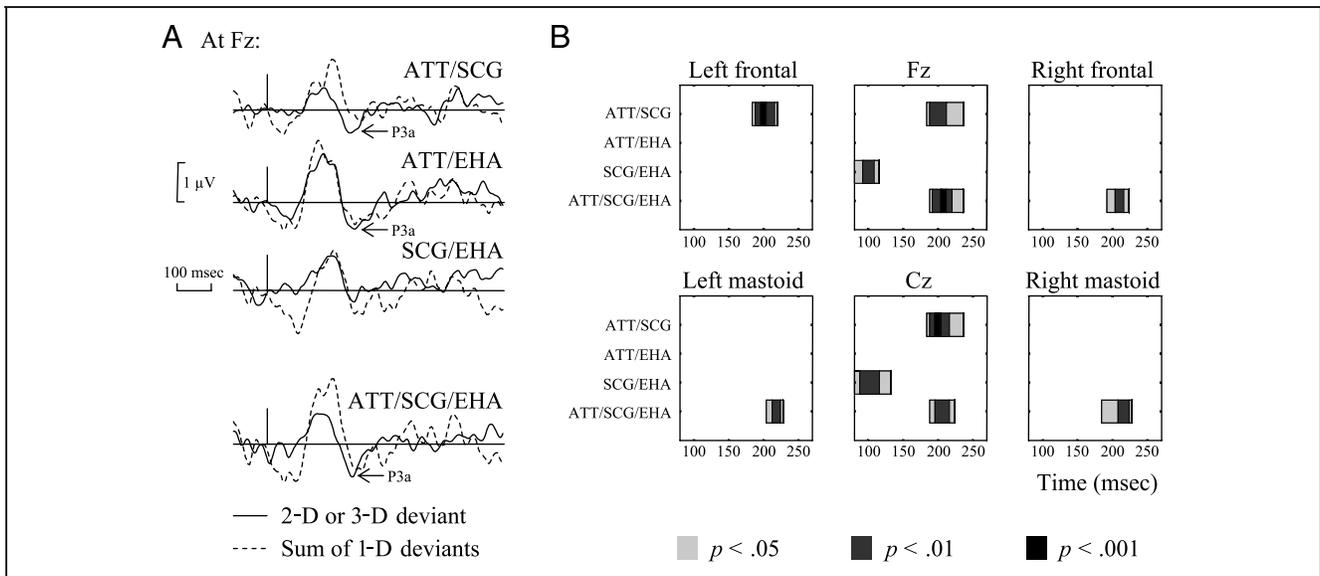
Figure 4. Additivity of unidimensional MMNs: (A) 2-D and 3-D MMNs (solid lines) and sums of the corresponding unidimensional MMNs (dashed lines) at Fz. (B) Detection of differences between multidimensional MMNs and the sum of the corresponding unidimensional MMNs in the 80- to 270-msec time window at a subset of the electrodes of Figure 3 (black, $p < .001$; dark gray, $p < .01$; light gray, $p < .05$).

Amplitudes and Latencies of Unidimensional MMNs

MMN amplitudes did not differ as a function of the deviating timbre dimension, $F(2,28) = 0.851$, $\varepsilon = 0.945$, $p = .43$, but MMN latencies depended on the deviating timbre dimension, $F(2,28) = 6.134$, $\varepsilon = 0.769$, $p = .01$. A Tukey HSD post hoc test revealed that this effect was mostly due to a difference between the ATT and SCG deviants ($p = .006$). The EHA deviant fell between the other two deviants and was marginally different from the SCG deviant ($p = .054$), but not significantly different from the ATT deviant ($p = .61$). This latency difference can be easily observed in Figure 3.

Table 2. MMN Amplitudes and Latencies Averaged across Nine Frontocentral Electrodes (as Described in Methods) for Each Deviant Type (See Table 1)

Deviant Type	Latency (msec)	Amplitude (μV)
1-D ATT	149 (±31)	−2.5 (±0.87)
1-D SCG	201 (±46)	−2.2 (±0.76)
1-D EHA	164 (±32)	−2.3 (±0.86)
2-D ATT/SCG	169 (±49)	−2.3 (±0.83)
2-D ATT/EHA	161 (±27)	−2.9 (±1.32)
2-D SCG/EHA	174 (±36)	−2.7 (±0.92)
3-D ATT/SCG/EHA	168 (±43)	−2.5 (±1.04)

The data reported are the mean (±SD) across the 15 participants. All amplitudes were significantly different from zero ($p < .0001$).

Additivity of Unidimensional MMNs

To assess unidimensional MMN additivity, MMNs for 2-D and 3-D changes were analyzed by comparing them with the sum of the corresponding unidimensional deviants (see Figure 4).

2-D Attack and Spectral Center of Gravity (ATT/SCG)

For this pair of dimensions, significant differences were observed at frontocentral electrodes around 200 msec. For example, this difference was significant at Fz from 184 to 232 msec, corresponding roughly to the latency range of the 1-D SCG MMN. The absolute value of the amplitude was larger for the sum of the unidimensional MMNs than for the bidimensional MMN, indicating that the additive model overestimates the MMN response in this time window. No such differences were observed over the mastoids.

2-D Attack and Even-harmonic Attenuation (ATT/EHA)

No differences were observed at any sites between the sum of unidimensional MMNs and the bidimensional MMN, suggesting an additivity of these unidimensional MMNs.

2-D Spectral Center of Gravity and Even-harmonic Attenuation (SCG/EHA)

For this last pair of dimensions, a significant difference was observed at frontocentral sites around 100 msec, for example between 72 and 132 msec at Fz. Nevertheless,

this appears to be due to a selective attenuation of N1 for the 1-D EHA deviant relative to its corresponding equiprobable ERP (see Figures 2 and 3). This unexpected selective attenuation of N1 did not occur for the 2-D deviant, hence the difference observed here. More importantly, in the latency range of the MMNs for these two dimensions (i.e., from about 120 to 230 msec), there was no difference between the sum of the unidimensional MMNs and the bidimensional MMN. The results therefore suggest that the MMNs for these two dimensions are additive.

3-D Deviant (ATT/SCG/EHA)

Finally, for the tridimensional MMN, a subadditivity (i.e., an overestimation of the 3-D MMN by the additive model) was observed at around 200 msec, both at frontocentral sites and over the mastoids. This difference was significant between 192 and 228 msec at Fz, between 204 and 224 msec at the left mastoid, and between 184 and 224 msec at the right mastoid.

Dipole Modeling

As can be seen in Figure 5A, the topographies of the three unidimensional MMNs were rather similar to each other, with a large frontocentral negativity in the early part, which persists and is later associated with a polarity reversal at the mastoids. This suggests that the general mechanisms involved are rather similar across timbre dimensions. Nevertheless, subtle topographical differences remain possible, and were assessed by using dipole modeling. The modeling procedure was conducted in the latency ranges where the equiprobable and deviant ERPs for each timbre were found to be significantly different: 112–208 msec for ATT, 180–228 msec for SCG and 124–224 msec for EHA. In a study similar to the present one but dealing with tone frequency, intensity, and duration, Giard et al. (1995) have shown that topographical differences between unidimensional MMNs were not much affected by the relative latencies of these MMNs (which can be controlled by modifying the degree of dissimilarity between the standard and the deviants). We therefore expected any topographical difference between the unidimensional MMNs recorded here to reflect genuine differences between timbre dimensions and not to result from the latency disparity between unidimensional MMNs.

The results of ECD modeling are shown in Figure 5B. The residual variance in the data not accounted for by the model was 9.1%, 4.8%, and 9.7% for ATT, SCG, and EHA, respectively. As can be seen in Figure 5B, the confidence intervals for the ECD locations for ATT and EHA did not overlap, with the ATT dipoles located more centrally, whereas the confidence interval for the location of the SCG dipoles overlapped with those of the ATT and EHA

dipoles. Therefore, it was possible to distinguish only between the location of the ATT and EHA dipoles, and the SCG dipoles were found in between them.

Additionally, dipole-orientation differences were observed (see Figure 5B). In the right hemisphere, because there was no overlap of the confidence intervals for the orientation of the three dipoles, we can conclude that dipole orientation was significantly different within all three pairs of dimensions. In the left hemisphere, the orientation of the ATT dipole was distinct from that of the two spectral deviant dipoles, but the orientation of the two spectral deviants' dipoles cannot be distinguished, as indicated by a large overlap of their confidence intervals. In general, ECD orientation differences without significant location differences are considered to reveal subtle anatomical separation of the generators that is not captured by location results, probably as a consequence of the rather low spatial resolution of EEG (e.g., Giard et al., 1995). This was the case here between SCG and ATT in both hemispheres and between SCG and EHA in the right hemisphere. Taken together, the present ECD location and orientation results suggest at least partially separate MMN generators for the three timbre dimensions tested in the right hemisphere. In the left hemisphere, only the ATT dipole can be separated from the other two.

DISCUSSION

For the three dimensions of timbre investigated in the present study, an MMN is elicited by a change along any single dimension and along any combination of two or three dimensions. The latency of the unidimensional MMNs depended on the dimension despite equal rated dissimilarity between the standard and the unidimensional deviants. These unidimensional MMNs are at least partially additive, and their generators appear separable to some extent on anatomical grounds. These results are discussed in light of the current knowledge about timbre processing and auditory sensory memory organization.

Dimension Dependence of the Relationship between MMN and Perceived Dissimilarity

The first striking result of the current study was that despite careful control of our stimuli to ensure equal perceived dissimilarity between the standard and the three unidimensional deviants, unidimensional MMN latencies were very variable: the latency difference between the MMNs for ATT and SCG amounted to 60 msec in the grand-average ERP, and to 50 msec for EHA and SCG. Although the finding of the shortest MMN latency for the ATT dimension could be easily explained because attack is among the initially available auditory information, the latency difference observed between the two spectral dimension MMNs was rather unexpected.
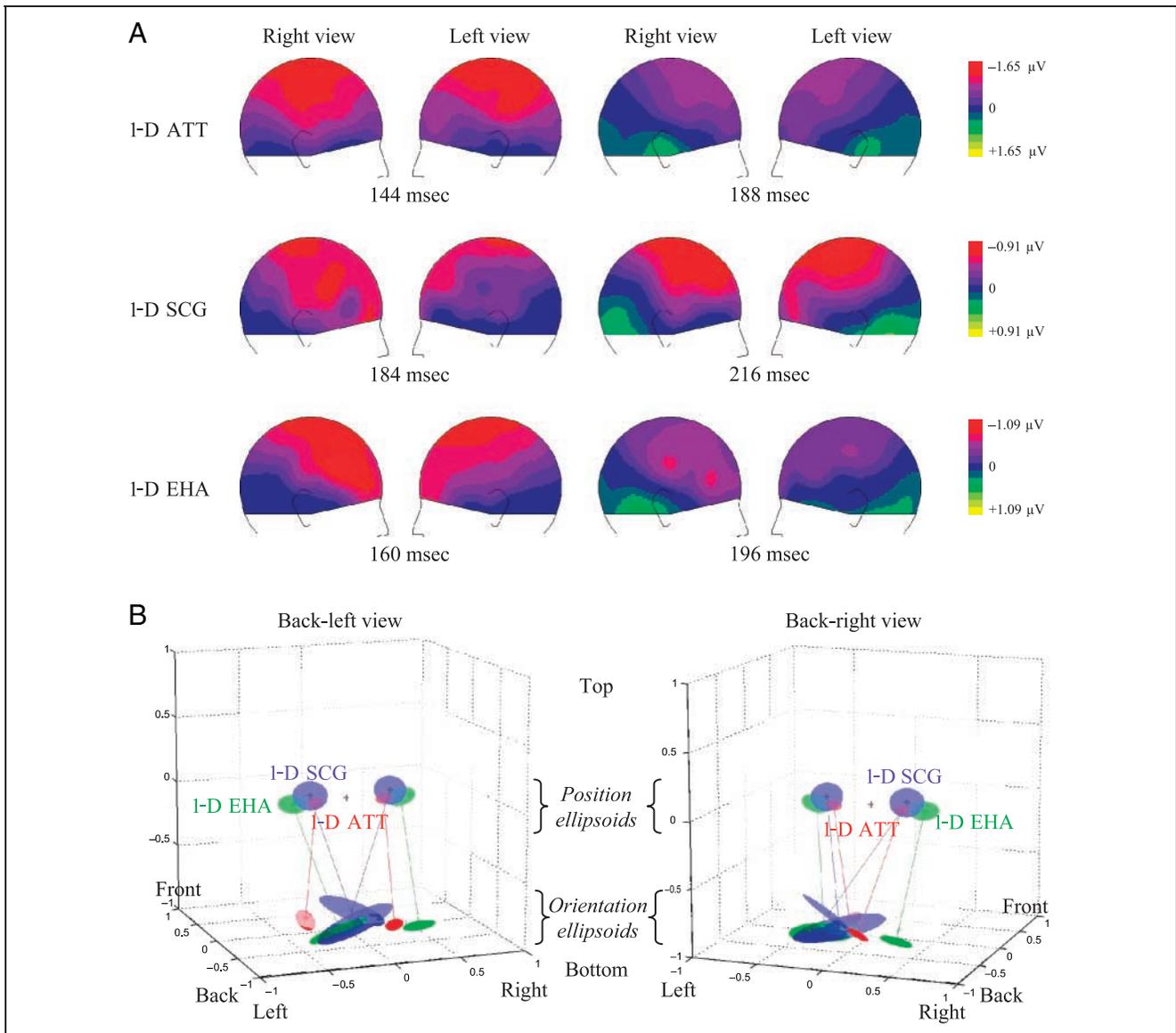
Figure 5. (A) Topographies and (B) equivalent current dipoles (ECDs) for unidimensional MMNs. (A) Topographies are presented at two different latencies for each MMN, corresponding to the early and late parts of the responses. The amplitude scale is adjusted for each timbre dimension. (B) 3-D representations of dipole-modeling results (left and right dipole positions for each timbre dimension were constrained to be symmetrical). The left and right panels are viewed from back–left and back–right, respectively. The sphere of radius one modeling the head is depicted in gray. ATT, SCG, and EHA dipoles are represented in red, blue, and green, respectively. The line segments are vectors of length 1, starting from dipole locations (at their top end) and indicating dipole orientation. Dipole position and orientation are represented with their 99% confidence ellipsoids. For dipole orientation, the confidence interval is the uncertainty of the location of the end of the orientation vector.

Within a single dimension, it has been shown that there is a correlation between MMN latency (and amplitude) and perceived dissimilarity (see Tiitinen et al., 1994, for frequency; Toiviainen et al., 1998, for spectral centroid). Our results therefore support the view that this relationship depends on the dimension tested, as equal perceived dissimilarities across dimensions do not lead to MMNs at the same latencies.

More generally, this result could imply that there is a characteristic time needed before information pertaining to one dimension becomes available for comparison with the traces in echoic memory, and that this charac-

teristic time depends on the auditory dimension. This interpretation warrants further testing.

Separate Processing of Timbre Dimensions in Auditory Sensory Memory

As highlighted in the Introduction, separate processing of auditory dimensions in sensory memory can be revealed by an additivity of unidimensional MMNs and by anatomically distinct generators for these MMNs. In our timbre data, additivity was observed for ATT and EHA and for SCG and EHA (at least in the time window of the MMN

response). For the ATT/SCG pair and for the triple deviant, multidimensional MMNs were subadditive (relative to the sum of unidimensional MMNs). Two alternative interpretations can be considered here: Either the ATT/SCG pair is different from the other two pairs and does not exhibit additivity, or other variables are confounding a genuine MMN additivity for this pair of dimensions. The first explanation seems unlikely because there is no experimental evidence that the processing of temporal aspects of the sounds would be interacting more strongly with that of global spectral characteristics such as SCG than with that of local spectral characteristics such as EHA. Put differently, if any pair of dimensions among the three tested here could have been thought beforehand to exhibit strong interactive behavior, it would have been the pair with the two spectral dimensions (SCG and EHA) rather than the ATT/SCG pair.

With these considerations in mind, we reviewed the existing MMN literature to assess whether there exist experimental variables that are likely to mask MMN additivity. Our starting point was that MMNs for ATT and SCG were the two unidimensional MMNs separated by the largest latency difference in our study. A first possibility is that the double deviance elicits a small P3a, which would cancel out the later part of the mismatch response corresponding to the end of SCG mismatch processing. Inspection of the ERP morphologies suggests that this could account for the subadditivity observed here (Figure 4). It is also possible that when the latency difference between the two unidimensional MMNs is large enough, the processing of the first mismatch prevents complete processing of the second. The available data are compatible with this critical time window view. For example, unidimensional MMNs for frequency and duration are additive when they occur at similar latencies (Levänen, Hari, McEvoy, & Sams, 1993), whereas they are subadditive when they are separated by ~70 msec (Czigler & Winkler, 1996), which is very similar to the present latency difference between the ATT and SCG MMNs. When the latency difference is larger than 200 msec, then two individualized responses are recorded for bidimensional deviants (e.g., Winkler, Czigler, Jaramillo, Paavilainen, & Näätänen, 1998, for intensity and frequency glide direction). It is also possible that in the present data, there is a combination of the two effects (P3a superposition and MMN subadditivity for critical latency differences). Altogether it appears most likely that MMNs are additive for the three pairs of timbre dimensions investigated here, and that this phenomenon would be better observed in conditions where the unidimensional MMNs occur at similar latencies. This would require adjusting the perceived dissimilarities at different values along the three dimensions.

We have also investigated the hypothesis of separate representation of timbre dimensions in sensory memory with dipole modeling of the three unidimensional MMNs. The scalp topographies were well explained in all three cases by two symmetrical dipoles that are likely to be located in the supratemporal auditory cortex. The location of the ATT and EHA dipoles could be distinguished, with the ATT dipoles located more centrally than those for EHA, but the location of the SCG dipole could not be distinguished from the other two. Furthermore, the orientations of the three dipoles were different from each other in both hemispheres, except for SCG and EHA in the left hemisphere. This suggests that the temporal generators of these three MMNs are located close to one another in the cortex, but in cortical areas tilted differently with respect to the scalp. On these grounds, we may conclude that the anatomical substrates of MMN generation are at least partially separated for the three timbre dimensions considered. The lack of difference between the location and orientation of the SCG and EHA dipoles in the left hemisphere might be related to the predominance of the right hemisphere for certain types of spectral processing (e.g., Johnsrude, Penhune, & Zatorre, 2000).

If the mismatch process reflected by the MMN operates in the cortical areas where the underlying memory traces are encoded, then our dipole modeling results would imply anatomically separate representations of timbre dimensions in sensory memory (see Giard et al., 1995; Ritter et al., 1995). This separability should result in unidimensional MMN additivity (see Schröger, 1995b). The current study is actually the first one to test unidimensional MMN additivity and separation of generator localization within the same data set. The results add support to the commonly accepted idea that additivity and generator separation should occur together as the consequences of the very same neural organization.

Our findings supporting separate representation of timbre dimensions in auditory sensory memory may at first sight appear incompatible with a recent study showing that within-instrument timbre changes (a manipulation used to create different emotional valences) and between-instrument timbre changes lead to MMNs with indistinguishable topographies (Goydke, Altenmüller, Möller, & Münte, 2004). Yet it is very likely that both of these timbre manipulations actually lead to modifications of several of the acoustic dimensions investigated in the present study, and, therefore, the observed topographies would result from a combination of change detection along all these different dimensions.

Processing of Auditory Attributes

The results of the current study have broader implications for our understanding of auditory processing. A first point is that MMNs for timbre dimensions behave in the current study in a way similar to that of MMNs for acoustic dimensions such as frequency, duration, and intensity in the sense that they exhibit some degree of additivity and anatomical separation. This suggests that for at least some of the processes indexed by the MMN,

the relevant entities are timbre dimensions and not timbre as a whole. Furthermore, as it is likely that there exist more dimensions to timbre than the three investigated here, the question arises of how many such perceptual parameters are represented separately in sensory memory. It is an intriguing enterprise to try to reconcile this parcelled appearance of auditory sensory memory with unified percepts. As mentioned in the Introduction, anatomical separation of the MMN generators for acoustic dimensions such as frequency and intensity or frequency and location does not prevent gestalt storage of those features (Sussman et al., 1998; Gomes et al., 1997). Therefore, despite a certain degree of anatomical separation between the neural substrate of the processing of various acoustic dimensions in sensory memory, there probably exist neural mechanisms integrating these different processing streams. Understanding the interplay between attribute-based processing and grouping mechanisms remains a challenge for auditory neuroscience.

Finally, because our results support the hypothesis of separate representations of timbre dimensions in auditory sensory memory, one might ask how this separation is rooted in the functional architecture of the central auditory system. Separate representations of sound attributes in auditory sensory memory might be the consequence of the existence of parallel processing streams for these different attributes, but it may also be that different attributes are extracted sequentially, for example, along a pathway devoted to spectral processing for SCG and EHA. A corollary question is whether this separation exists at other stages of processing than sensory memory. For example, intensity and frequency are represented separately in sensory memory (Clément et al., 1999; Giard et al., 1995), but it is known that there exist interactions between the processing of these two dimensions, because they exhibit Garner interference (Melara & Marks, 1990; Grau & Kemler-Nelson, 1988), in which speeded classification along one dimension is affected by concomitant variation along the other dimension. Exploring other levels of timbre processing than that of sensory memory will be necessary to fully understand the relationships among the different dimensions of this multidimensional attribute.

REFERENCES

Alho, K. (1995). Cerebral generators of mismatch negativity (MMN) and its magnetic counterpart MMNm elicited by sound changes. *Ear & Hearing, 16,* 38–51.

American National Standards Institute. (1973). *American National Standard psychoacoustical terminology.* New York: American National Standards Institute.

Berger, K. W. (1964). Some factors in the recognition of timbre. *Journal of the Acoustical Society of America, 36,* 1888–1891.

Caclin, A., McAdams, S., Smith, B. K., & Winsberg, S. (2005). Acoustic correlates of timbre space dimensions: A confirmatory study using synthetic tones. *Journal of the Acoustical Society of America, 118,* 471–482.

Caclin, A., Menon, V., Krasnow, B., Mackenzie, K., Smith, B., & McAdams, S. (2003). fMRI correlates of timbre perception [Abstract]. *Neuroimage, 19,* 61.

Carroll, J. D., & Chang, J. J. (1970). Analysis of individual differences in multidimensional scaling via an n-way generalization of Eckart–Young decomposition. *Psychometrika, 35,* 283–319.

Clément, S., Demany, L., & Semal, C. (1999). Memory for pitch versus memory for loudness. *Journal of the Acoustical Society of America, 106,* 2805–2811.

Cowan, N. (1984). On short and long auditory stores. *Psychological Bulletin, 96,* 341–370.

Crummer, G. C., Walton, J. P., Wayman, J. W., Hantz, E. C., & Frisina, R. D. (1994). Neural processing of musical timbre by musicians, nonmusicians and musicians possessing absolute pitch. *Journal of the Acoustical Society of America, 95,* 2720–2727.

Czigler, I., & Winkler, I. (1996). Preattentive auditory change detection relies on unitary sensory memory representations. *NeuroReport, 7,* 2413–2417.

Deacon, D., Nousak, J. M., Pilotti, M., Ritter, W., & Yang, C. M. (1998). Automatic change detection: Does the auditory system use representations of individual stimulus features or gestalts? *Psychophysiology, 35,* 413–419.

Deutsch, D. (1970). Tones and numbers: Specificity of interference in immediate memory. *Science, 168,* 1604–1605.

Frodl-Bauch, T., Kathmann, N., Moller, H. J., & Hegerl, U. (1997). Dipole localization and test–retest reliability of frequency and duration mismatch negativity generator processes. *Brain Topography, 10,* 3–8.

Giard, M. H., Lavikainen, J., Reinikainen, K., Perrin, F., Bertrand, O., Pernier, J., et al. (1995). Separate representation of stimulus frequency, intensity, and duration in auditory sensory memory: An event-related potential and dipole-model analysis. *Journal of Cognitive Neuroscience, 7,* 133–143.

Giard, M. H., Perrin, F., Echallier, J. F., Thévenet, M., Froment, J. C., & Pernier, J. (1994). Dissociation of frontal and temporal components in the human auditory N1 wave: A scalp current density and dipole model analysis. *Electroencephalography and Clinical Neurophysiology, 92,* 238–252.

Giard, M. H., Perrin, F., Pernier, J., & Bouchet, P. (1990). Brain generators implicated in the processing of auditory stimulus deviance: A topographic event-related potential study. *Psychophysiology, 27,* 627–640.

Gomes, H., Bernstein, R., Ritter, W., Vaughan, H. G., Jr., & Miller, J. (1997). Storage of feature conjunctions in transient auditory memory. *Psychophysiology, 34,* 712–716.

Gomes, H., Ritter, W., & Vaughan, H. G., Jr. (1995). The nature of preattentive storage in the auditory system. *Journal of Cognitive Neuroscience, 7,* 81–94.

Goydke, K. N., Altenmüller, E., Möller, J., & Münte, T. F. (2004). Changes in emotional tones and instrumental timbre are reflected by the mismatch negativity. *Cognitive Brain Research, 21,* 351–359.

Grau, J. W., & Kemler-Nelson, D. G. (1988). The distinction between integral and separable dimensions: Evidence for integrality of pitch and loudness. *Journal of Experimental Psychology: General, 117,* 347–370.

Grey, J. M. (1977). Multidimensional perceptual scaling of musical timbres. *Journal of the Acoustical Society of America, 61,* 1270–1277.

Guthrie, D., & Buchwald, J. S. (1991). Significance testing of difference potentials. *Psychophysiology, 28,* 240–244.

Hajda, J. M. (1999). *The effect of time-variant acoustical properties on orchestral instrument timbres.* Unpublished PhD thesis, University of California, Los Angeles.

Hajda, J. M., Kendall, R. A., Carterette, E. C., & Harschberger, M. L. (1997). Methodological issues in timbre research. In I. Deliège & J. Sloboda (Eds.), *Perception and cognition of music* (pp. 253–307). Hove, UK: Psychology Press.

Handel, S. (1995). Timbre perception and auditory object identification. In B. C. J. Moore (Ed.), *Hearing* (pp. 425–461). San Diego, CA: Academic Press.

Jacobsen, T., Schröger, E., Horenkamp, T., & Winkler, I. (2003). Mismatch negativity to pitch change: Varied stimulus proportions in controlling effects of neural refractoriness on human auditory event-related brain potentials. *Neuroscience Letters, 344,* 79–82.

Johnsrude, I. S., Penhune, V. B., & Zatorre, R. J. (2000). Functional specificity in the right human auditory cortex for perceiving pitch direction. *Brain, 123,* 155–163.

Kendall, R. A., & Carterette, E. C. (1993a). Verbal attributes of simultaneous wind instruments timbres: I. von Bismarck's adjectives. *Music Perception, 10,* 445–468.

Kendall, R. A., & Carterette, E. C. (1993b). Verbal attributes of simultaneous wind instruments timbres: II. Adjectives induced from Piston's orchestration. *Music Perception, 10,* 469–502.

Krimphoff, J., McAdams, S., & Winsberg, S. (1994). Caractérisation du timbre des sons complexes. II. Analyses acoustiques et quantification psychophysique [Characterization of the timbre of complex sounds. II. Acoustical analyses and psychophysical quantification]. *Journal de Physique, 4,* 625–628.

Krumhansl, C. L. (1989). Why is musical timbre so hard to understand? In S. Nielsen & O. Olsson (Eds.), *Structure and perception of electroacoustic sound and music* (pp. 43–53). Amsterdam: Elsevier.

Krumhansl, C. L., & Iverson, P. (1992). Perceptual interactions between musical pitch and timbre. *Journal of Experimental Psychology, Human Perception and Performance, 18,* 739–751.

Levänen, S., Hari, R., McEvoy, L., & Sams, M. (1993). Responses of the human auditory cortex to changes in one versus two stimulus features. *Experimental Brain Research, 97,* 177–183.

Loewy, D. H., Campbell, K. B., & Bastien, C. (1996). The mismatch negativity to frequency deviant during natural sleep. *Electroencephalography and Clinical Neurophysiology, 98,* 493–501.

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.

Marozeau, J., de Cheveigné, A., McAdams, S., & Winsberg, S. (2003). The dependency of timbre on fundamental frequency. *Journal of the Acoustical Society of America, 114,* 2946–2957.

McAdams, S. (1993). Recognition of sound sources and events. In S. McAdams & E. Bigand (Eds.), *Thinking in sound: The cognitive psychology of human audition* (pp. 146–198). Oxford: Oxford University Press.

McAdams, S., Winsberg, S., Donnadieu, S., De Soete, G., & Krimphoff, J. (1995). Perceptual scaling of synthesized musical timbres: Common dimensions, specificities and latent subject classes. *Psychological Research, 58,* 177–192.

Melara, R. D., & Marks, L. E. (1990). Perceptual primacy of dimensions: Support for a model of dimensional interaction. *Journal of Experimental Psychology: Human Perception & Performance, 16,* 398–414.

Menon, V. Levitin, D. J., Smith, B. K., Lembke, A., Krasnow, B. D., Glazer, D., et al. (2002). Neural correlates of timbre change in harmonic sounds. *Neuroimage, 17,* 1742–1754.

Miller, J. R., & Carterette, E. C. (1975). Perceptual space for musical structures. *Journal of the Acoustical Society of America, 58,* 711–720.

Milner, B. (1962). Laterality effects in audition. In V. B. Mountcastle (Ed.), *Interhemispheric relations and cerebral dominance* (pp. 177–195). Baltimore: Johns Hopkins Press.

Nousak, J. M., Deacon, D., Ritter, W., & Vaughan, H. G., Jr. (1996). Storage of information in transient auditory memory. *Cognitive Brain Research, 4,* 305–307.

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

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

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., & Winkler, I. (1999). The concept of auditory stimulus representation in cognitive neuroscience. *Psychological Bulletin, 125,* 826–859.

Paavilainen, P., Valppu, S., & Näätänen, R. (2001). The additivity of auditory feature analysis in the human brain as indexed by the mismatch negativity: $1+1\approx2$ but $1+1+1<3$. *Neuroscience Letters, 301,* 179–182.

Pantev, C., Oostenveld, R., Engelien, A., Ross, B., Roberts, L. E., & Hoke, M. (1998). Increased auditory cortical representation in musicians. *Nature, 392,* 811–814.

Pantev, C., Roberts, L. E., Schulz, M., Engelien, A., & Ross, B. (2001). Timbre-specific enhancement of auditory cortical representations in musicians. *NeuroReport, 12,* 169–174.

Ritter, W., Deacon, D., Gomes, H., Javitt, D. C., & Vaughan, H. G., Jr. (1995). The mismatch negativity of event-related potentials as a probe of transient auditory memory: A review. *Ear & Hearing, 16,* 52–67.

Rosburg, T. (2003). Left hemispheric dipole locations of the neuromagnetic mismatch negativity to frequency, intensity and duration deviants. *Cognitive Brain Research, 16,* 83–90.

Saldanha, E. L., & Corso, J. F. (1964). Timbre cues and the identification of musical instruments. *Journal of the Acoustical Society of America, 36,* 2021–2026.

Sams, M., Alho, K., & Näätänen, R. (1984). Short-term habituation and dishabituation of the mismatch negativity of the ERP. *Psychophysiology, 21,* 434–441.

Samson, S., & Zatorre, R. J. (1994). Contribution of the right temporal lobe to musical timbre discrimination. *Neuropsychologia, 32,* 231–240.

Samson, S., Zatorre, R. J., & Ramsay, J. O. (1997). Multidimensional scaling of synthetic musical timbre: Perception of spectral and temporal characteristics. *Canadian Journal of Experimental Psychology, 51,* 307–315.

Samson, S., Zatorre, R. J., & Ramsay, J. O. (2002). Deficits of musical timbre perception after unilateral temporal-lobe lesion revealed with multidimensional scaling. *Brain, 125,* 511–523.

Schröger, E. (1995a). Interaural time and level differences: Integrated or separated processing? *Hearing Research, 96,* 191–198.

Schröger, E. (1995b). Processing of auditory deviants with changes in one versus two stimulus dimensions. *Psychophysiology, 32,* 55–65.

Semal, C., & Demany, L. (1991). Dissociation of pitch from timbre in auditory short-term memory. *Journal of the Acoustical Society of America, 89,* 2404–2410.

Sinkkonen, J., & Tervaniemi, M. (2000). Towards optimal recording and analysis of the mismatch negativity. *Audiology & Neuro-Otology, 5,* 235–246.

Sussman, E., Gomes, H., Nousak, J. M., Ritter, W., & Vaughan, H. G., Jr. (1998). Feature conjunction and auditory sensory memory. *Brain Research, 793,* 95–102.

Takegata, R., Brattico, E., Tervaniemi, M., Varyagina, O., Näätänen, R., & Winkler, I. (2005). Preattentive representation of feature conjunctions for concurrent spatially distributed auditory objects. *Cognitive Brain Research, 25,* 169–179.

Tervaniemi, M., Winkler, I., & Näätänen, R. (1997). Pre-attentive categorization of sounds by timbre as revealed with event-related potentials. *NeuroReport, 8,* 2571–2574.

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.

Toiviainen, P., Tervaniemi, M., Louhivuori, J., Saher, M., Huotilainen, M., & Näätänen, R. (1998). Timbre similarity: Convergence of neural, behavioral and computational approaches. *Music Perception, 16,* 223–241.

von Bismarck, G. (1974). Timbre of steady sounds: A factorial investigation of its verbal attributes. *Acustica, 30,* 159–192.

Winkler, I., Czigler, I., Jaramillo, M., Paavilainen, P., & Näätänen, R. (1998). Temporal constraints of auditory events synthesis: Evidence from ERPs. *NeuroReport, 9,* 495–499.

Winkler, I., Tervaniemi, M., Schröger, E., Wolff, C., & Näätänen, R. (1998). Preattentive processing of auditory spatial information in humans. *Neuroscience Letters, 242,* 49–52.

Winsberg, S., & De Soete, G. (1993). A latent class approach to fitting the weighted euclidean model, CLASCAL. *Psychometrika, 58,* 315–330.

Wolff, C., & Schröger, E. (2001). Human pre-attentive auditory change-detection with single, double, and triple deviations as revealed by mismatch negativity additivity. *Neuroscience Letters, 311,* 37–40.