Neural encoding of sound duration persists in older adults

Bernhard Ross a,b,⁎, Joel S. Snyder c, Meaghan Aalto a, Kelly L. McDonald a, Benjamin J. Dyson d, Bruce Schneider e, Claude Alain a,e

a Rotman Research Institute, Baycrest Centre, Toronto, Ontario, Canada M6A 2E1
b Department of Medical Biophysics, University of Toronto, Ontario, Canada M5G 2M9
c Department of Psychology, University of Nevada, Las Vegas, NV 89154, USA
d Department of Psychology, Ryerson University, Toronto, Ontario, Canada M5B 2K3
e Department of Psychology, University of Toronto, Ontario, Canada M5G 2M9

A R T I C L E   I N F O

Article history:
Received 4 February 2009
Revised 9 April 2009
Accepted 10 April 2009
Available online 22 April 2009

A B S T R A C T

Speech perception depends strongly on precise encoding of the temporal structure of sound. Although behavioural studies suggest that communication problems experienced by older adults may entail deficits in temporal acuity, much is unknown about the effects of age on the neural mechanisms underlying the encoding of sound duration. In this study, we measured neuromagnetic auditory evoked responses in young, middle-aged and older healthy participants listening to sounds of various durations. The time courses of cortical activity from bilateral sources in superior temporal planes showed specific differences related to the sound offsets indicating the neural representation of onset and offset markers as one dimension of the neural code for sound duration. Model free MEG source analysis identified brain areas specifically responding with an increase in activity to increases in sound duration in the left anterior insula, right inferior frontal, right middle temporal, and right post-central gyri in addition to bilateral supra-temporal gyri. Sound duration-related changes in cortical responses were comparable in all three age groups despite age-related changes in absolute response magnitudes. The results demonstrated that early cortical encoding of the temporal structure of sound presented in silence is little or not affected by normal aging.

© 2009 Elsevier Inc. All rights reserved.

Introduction

The ability to perceive the duration of sound is important for understanding speech. For instance, in many languages differences in the duration of a vowel on the order of tens of milliseconds may distinguish between words of different meaning. Although it is well recognized that speech perception depends on the processing of the fine temporal structure of sounds (e.g., envelope), little is known about how sound duration itself is encoded in human cortex. Studies measuring neuroelectric brain responses to changes in sound duration have been shown as mismatch negativity (MMN) in auditory evoked potentials (AEP) and most importantly the size of the response determined the accuracy of perceptual performance (Amenedo and Escera, 2000). Other studies have shown that the amplitude of the N1 wave, in EEG a negative deflection with maximum at vertex electrodes, peaking at about 100 ms after sound onset, increases as sound duration increases up to 40 ms (Alain et al., 1997; Forss et al., 1993; Gage and Roberts, 2000; Joutsiniemi et al., 1989; Onishi and Davis, 1968). These observations support the hypothesis that neurons can act as a linear integrator of the acoustical energy with their output being proportional to the sound duration. Furthermore, neurons tuned to a specific sound duration between 30 and 300 ms have been found in auditory cortex of the cat (He et al., 1997). In human, Alain et al. (1997) showed that the N1c subcomponent of the N1, recorded at mid temporal electrode sites with peak latency of about 130 ms, increased strongly with duration increase up to 24 ms, suggesting a short integration time and triggering by the sound onset. In contrast, the amplitude of the N1a subcomponent, recorded from same electrodes with peak latency around 90 ms, increased more for longer sound durations, indicating integration of ongoing sound. In addition to integration mechanisms complex networks have been proposed in which the onset and offset of sound is marked and a counting mechanism estimates how many time units elapsed between both markers (Creelman, 1962). The integrator model would appear to be suitable for short sound durations only and unsuitable across longer time intervals that become more and more vulnerable to noise and stimulus intensity. In these cases the latter stopwatch-like mechanism would be more effective. However, in principle, the marking of sound onsets and offsets could serve as an important encoding principle even for short tone durations. Related onset and offset markers have been identified in the auditory thalamus of the guinea pig for sound duration as short as 50 ms (He, 2001).

⁎ Corresponding author. Baycrest Centre, Rotman Research Institute, 3560 Bathurst Street, Toronto, Ontario, Canada M6A 2E1. Fax: +1 417 785 2662.
E-mail address: bross@rotman-baycrest.on.ca (B. Ross).

1053-8119/—see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2009.04.051
Temporal acuity in hearing has been mainly studied as performance in detecting a silent gap (Phillips, 1999) or periodic modulation in sound (Joris et al., 2004). Continuing this work, our present study aims to further understand the neural networks involved in processing of sound duration in the range of tens of milliseconds. With magnetoencephalography (MEG) we recorded brain activity to sounds that varied in duration and analyzed spatial distributions of underlying cortical source activity. The hypothesis was that neuro-magnetic source imaging would reveal distinct brain areas specifically responsive to sound of different duration and more importantly it would provide the time course of activation with high temporal resolution, which could inform us about the employed encoding strategy. Furthermore we addressed the question whether encoding of sound duration changes over the adult life span according to the hypothesis that aging impairs temporal acuity in sound processing. This question is important for understanding of the speech communication problem in the elderly, which may result from deterioration in encoding elementary sound features or from deficits in higher order function of inhibiting irrelevant and focusing attention on relevant sound. Indication for age-related decline in temporal acuity came from a study of Ostroff et al. (2003), who measured auditory AEPs to small increases in sound duration (i.e., 2–10 ms in 2-ms steps) across young, middle-aged and older adults and confirmed the increase in N1 amplitude with sound duration for all ages despite age-related differences in absolute N1 amplitudes. However, increase in P2 amplitude was observed in young and middle-aged but not in the older adults. Correspondingly, smaller mismatch negativity has been found in older adults for changes in the duration of a silent gap in sound and this effect remained even when young and older adults were matched in terms of their gap thresholds (Alain et al., 2004; Bertoli et al., 2002).

Methods

Participants

Seventeen young (mean age 24 years (range 20–36 years); 12 female), twelve middle-aged (mean age 47 years (38–54); 6 female), and thirteen older (mean age 71 years (61–81); 7 female) healthy adults, recruited from the local community and laboratory personnel, participated in the study. Although, all participants had pure-tone thresholds less than or equal to 30 dB hearing level (HL) between 250 and 2000 Hz in both ears, mean audiometric thresholds across this frequency range increased gradually between age groups \( F(2,40)=14.1, p<0.001 \) from 2.4 dB HL in the young adults, 4.5 dB HL in the middle-aged group, to 13.3 dB HL in the group of older adults, which was significantly higher compared to young or middle-aged adults \( p<0.001 \) in both cases). Participants provided informed consent in accordance with the guidelines established by the University of Toronto and Baycrest Centre for Geriatric Care.

Stimuli and procedure

During one block of 24 min recording time, participants listened passively to 1064 sound stimuli of which equal numbers had durations of 34, 46, 64, and 76 ms, respectively. Half of the stimuli contained a gap (Fig. 1), however, the auditory evoked fields (AEFs) elicited by those gap stimuli are reported elsewhere. Stimuli were 1-kHz pure tones with the envelope shaped by a Gaussian with standard deviation of 0.5 ms as described by Schneider and Hamstra (1999) and amplitudes adjusted for equal energy. Diotic stimuli were presented at 80 dB sound pressure level (SPL) via an OB 822 Clinical Audiometer through ER30 transducers (Etymotic Research, Elk Grove, USA) and connected with 1.5 m of length matched plastic tubing and foam earplugs to the participant’s ears. Sound transmission through plastic tubing of such length was required because the sound transducers had to be placed at a sufficient distance from the MEG sensor to avoid any interference between the stimulus signal and the recorded brain activity. Below 2000 Hz the frequency characteristic of the sound transmission system was relatively flat (±6 dB) and the phase characteristic was linear as tested using a 2-cm² coupler. All types of different stimulus sounds were presented in random order with stimulus onset asynchrony (SOA) uniformly randomized between 850 and 1850 ms under control of Neuroscan Stim Software. No specific task was required for the participants who watched a silent movie of their own choice.

Data acquisition and analysis

The MEG was recorded in a magnetically shielded room using a 151-channel whole head neuro-magnetometer (OMEGA, VSM Med-tech Inc., Vancouver, Canada). In order to minimize head movement, participants were in supine position with their head resting in the helmet shaped MEG sensor. Neuro-magnetic activity was sampled at rate of 625 Hz after 200 Hz low-pass filtering and was recorded continuously. We performed two complementary types of data analysis. First we modeled the averaged MEG data with single dipoles in left and right auditory cortices and calculated the waveforms of source strength, which allowed studying the sequence of positive and negative waves in the responses like the more conventional analysis of waveforms obtained from selected electrodes in EEG or sensors in MEG. Second, we performed a model free analysis of source activity in order to identify the brain areas involved in encoding of sound duration.

Dipole source analysis

Epochs of MEG data including 200 ms of pre- and 600 ms of post-stimulus activity were averaged separately for each stimulus type and digitally low-pass filtered at 20 Hz. Individual dipole source analysis was performed on the data grand averaged across all stimulus types. The source model consisted of single equivalent current dipoles in the left and right temporal lobes and was approximated to the 40 ms interval around the peak of the N1 wave. Source waveforms for each stimulus condition were calculated based on the fixed dipole locations and orientations. The

Fig. 1. Waveforms of the auditory stimuli. The sounds of 34, 46, 64, and 76 ms total duration, respectively, were either continuous tones or contained a temporal gap. The gap lengths were 4 or 16 ms, the length of the leading marker was 10 or 40 ms, and the trailing marker was 20 ms long. The stimulus amplitudes were adjusted for equal energy.
polarity of source waveforms was defined for a negative going wave at 100 ms latency. Peak amplitudes relative to the baseline amplitude and latencies relative to the sound onset were determined as the largest positivity (P) or negativity (N) in the individual source waveforms in the time intervals of 30–70 ms for P1m, 70–150 ms for N1m, and 140–240 ms for P2m. AEF amplitudes and latencies were analyzed using a mixed model repeated measures ANOVA with age as the between-groups factor and signal duration, and hemisphere as the within-group factors. When appropriate, the degrees of freedom were adjusted with the Greenhouse–Geisser epsilon (\(\varepsilon\)). Probability estimates were based on the reduced degrees of freedom although the original degrees of freedom are reported.

Event related beamforming analysis

The synthetic aperture magnetometry (SAM) minimum-variance beamformer algorithm (Van Veen et al., 1997) was used as a spatial filter to estimate the source activity on a lattice of 5 mm spacing across the whole brain volume. Waveforms of averaged source activity across all trial for each stimulus type were calculated following the event-

![Figure 2](image1)

**Fig. 2.** Grand averaged waveforms of cortical source activity. (a) Evoked responses to the four sounds of different duration in the left hemisphere for the three age groups, (b) averaged across groups, (c) for the right hemisphere and all age groups, and (d) averaged across groups. The inserts in b and d depict the N1 peak on enlarged scales.

![Figure 3](image2)

**Fig. 3.** Waveforms of responses to the 46, 64, and 76-ms tones after subtraction of the response to the shortest tone of 34 ms duration. (a) Difference waveforms for the three subject groups and left and right hemispheres. (b) Difference waveforms grand averaged across subject groups and hemispheres. The time axis was adjusted to the stimulus offset, which emphasizes the common peak latency for the negative peak at 75 ms after stimulus offset.
related SAM (ER-SAM) approach (Cheyne et al., 2006). The beamformer analysis using the algorithm as implemented in the VSM software package was based on individual multi sphere models, for which single spheres were locally approximated for each of the 151 MEG sensors to the three-dimensionally digitized head shape. The ER-SAM procedure results in a \( z \)-score of the source activity, which often is normalized to an estimate of the sensor noise (Robinson and Vrba, 1999). However, normalization with respect to the spontaneous brain activity seems more meaningful. Therefore, we calculated the mean variance across all time points within an epoch (McCubbin et al., 2008) and normalized the estimates of source activity by this variance measure, which resulted in time series of signal-to-noise ratios (SNR) for each voxel. Volumetric maps of group mean SNR values at selected time points were overlaid to the anatomical image of a template brain (colin27, Montreal Neurological Institute) (Holmes et al., 1998) and were visualized with AFNI software (National Institute of Mental Health, Bethesda, MD, USA) (Cox, 1996). Linear regression of the experimental parameter stimulus duration (34, 46, 64, and 76 ms) to the corresponding four data sets of averaged source activity identified those brain regions, which showed maximum modulation of activity as effect of the stimulus duration. A one-way ANOVA was performed for each voxel to test the effect of stimulus duration. Source waveforms of brain activity at local maxima of the linear regression were calculated and compared between age groups.

**Results**

**Dipole source activity in auditory cortices**

Auditory evoked magnetic fields were successfully recorded in all subjects and grand averaged source waveforms of AEFs from left and right hemispheres for the stimuli with four different durations and the three age groups are shown in Figs. 2a, c. In all three age groups, responses to each stimulus showed clearly pronounced P1m, N1ms, and P2m waves with peak latencies around 70, 105, and 170 ms, respectively. Individual P1m–N1m–P2m peaks could be identified in all subjects for all stimuli in left and right hemispheres. Increasing P1m amplitudes with advancing age became obvious as a consistent between-group effect in left and right hemispheres, whereas N1m were almost replicated between age groups except smaller N1m amplitudes in the young group in the left hemisphere only. Larger variability between age groups can be seen in the latency range of the P2m wave. In order to visualize the effect of stimulus duration, the AEF waveforms were grand averaged across the three age groups (Figs. 2b, d). The P1m amplitudes were almost exactly reproduced for different stimulus durations. The N1m amplitudes increased with increasing stimulus duration; however the changes were small and the amplitude increase did not follow the stimulus duration monotonically. Larger changes in the AEF waveforms were observed in the latency range of the P2m wave. The P2m amplitude seemed to decrease for longer stimulus duration, although, this may be caused by an actual decrease of the P2m response itself or due to a superimposed wave with negative polarity. Fig. 3 was composed to investigate the stimulus duration-related changes in the latency range between 100 and 200 ms in more detail. The assumption for this analysis was that the response to the shortest tone of 34 ms duration reflected mainly the effect to the stimulus onset and could be used as a model for the onset response. The response to the 34 ms stimulus was subtracted from all responses to sounds of longer duration. The resulting

**Table 1**

| Brain region with largest modulation of activity as function of stimulus duration. |
|----------------------------------------|-----------------|------------------|-----------------|-----------------|
| Brain region                             | Talairach        | Latency (ms)     | Regression       | \( F(3,162) \) | \( p \)         |
| Left hemisphere                          |                 |                  |                  |                 |
| Insula                                  | \(-37, 5, -2\ mm| 234              | 0.0282/ms        | 5.18            | 0.0019         |
| Supra-temporal gyrus                    | \(-44, -29, 9\ mm| 138              | 0.0281/ms        | 2.73            | 0.0354         |
| Right hemisphere                        |                 |                  |                  |                 |
| Inferior frontal gyrus                  | \(54, 12, 5\ mm | 234              | 0.0236/ms        | 4.27            | 0.0062         |
| Middle-temporal gyrus                   | \(59, -1, -8\ mm| 256              | 0.0379/ms        | 4.62            | 0.0040         |
| Post-central gyrus                      | \(55, -17, 16\ mm| 234              | 0.0289/ms        | 4.15            | 0.0072         |
| Supra-temporal gyrus                    | \(55, -10, 4\ mm| 138              | 0.0279/ms        | 2.89            | 0.0371         |

The regions were identified as local maximum of the coefficient of linear regression of stimulus duration to the activity in each voxel.

Fig. 4. RMS values of magnetic fields recorded with frontal (F), central (C), temporal (T), parietal (P), and occipital (O) groups of MEG sensors above the left (L) and right (R) hemispheres. For clarity the responses to the shortest and longest tone durations of 34 and 76 ms are shown.
difference waveforms, shown in Fig. 3, exhibit a negative wave in the 100–200 ms interval after stimulus onset and increasing amplitude with longer stimulus duration. Visual inspection revealed consistent waveforms across age groups and hemispheres and therefore the waveforms were grand averaged as shown in Fig. 3b. One striking observation was that the negative peak in the difference waveforms always occurred approximately 75 ms after stimulus offset regardless of the stimulus duration. The constant latency with respect to the stimulus offset rather than stimulus onset suggests that the falling slope of the stimulus elicited the response. In order to emphasize the relationship between sound offset and the negative peak in the difference waves the time axis of the waveforms was aligned to the

Fig. 5. Brain areas with largest modulation of activity by stimulus duration in the latency range of 150 to 250 ms after stimulus onset. (a) Color coded is the coefficient b of linear regression of the stimulus duration to the source activity at each voxel (e.g. b = 0.03 corresponds to an increase of activity by the factor of 1.2 when the stimulus duration increases by 40 ms) (b) Time courses of activity at the location of strongest duration effect in the left hemisphere (note that the peak latency increases accordingly to the stimulus duration), (c) Three areas with local maximum of the duration effect were identified in the right hemisphere. (d) Time courses of activation in the right post-central gyrus with a first around 114 ms for all stimuli and a second peak with progressively increasing latency according to the stimulus duration. (e) Activation in right inferior frontal gyrus shows a single peak at 150 ms latency, (f) whereas the activation in right middle temporal gyrus shows double peaks with latencies according to stimulus onset and offset.
stimulus offset. Fig. 3b reveals also a nonlinear relation between response magnitude and duration. The upper trace exhibits a barely recognizable peak as the effect of 12 ms increase in duration from 34 to 46 ms. The more pronounced difference between the two lower waveforms indicate that the same difference of 12 ms between 64 and 76 ms duration has a much stronger effect on the AEF.

Effects of age and sound duration on the AEF waveforms

Quantitative evaluation of effects of age and hemisphere on the peak latencies and amplitudes of the P1m, N1m, and P2m was performed with a mixed-design ANOVA. The main effect of age on the P1m amplitude was significant ($F(2,38) = 8.15, p < 0.005$) with larger responses in older adults compared to young and middle-aged adults (see also Figs. 2a, c). Moreover, the P1m responses were larger ($F(1,38) = 8.57, p < 0.01$) in the left hemisphere than the right hemisphere, however, a group×hemisphere interaction was not significant. The ANOVA on the N1m amplitude yielded no main effect of age on the N1m amplitude ($F(2,38) < 1.0$). However, an interaction between hemisphere and age ($F(2,38) = 4.35, p < 0.05$) was caused by larger responses in the right hemisphere for young adults but larger responses in the left hemisphere for middle-aged and older adults. The N1m amplitude increased progressively with signal duration as revealed by a main effect of sound duration (linear trend, $F(1,38) = 12.70, p < 0.001$). This duration-related increase in N1m amplitude was comparable in both hemispheres (hemisphere×sound duration, linear trend, $F(1,38) < 1.0$). More importantly, an interaction between group and sound duration was not significant (linear trend, $F(1,38) = 1.1$), indicating that the duration-related increase in N1m amplitude was not affected by normal aging. The P2m wave was larger over the left than the right hemisphere, ($F(1,38) = 17.65, p < 0.001$). There was no main effect or interactions involving age on the P2m amplitude.

The mean amplitudes of the difference waves in the latency range around 150 ms were analyzed with a mixed-design ANOVA with duration (46, 64, 76 ms), hemisphere, and age as factors (see Fig. 3). This revealed a significant main effect of duration ($F(2,78) = 61.52, p < 0.001$) with linear ($F(1,39) = 71.31, p < 0.001$) and quadratic ($F(1,39) = 16.41, p < 0.001$) trends, reflecting an increase in duration-related negativity with longer stimulus durations. There were no main effects of hemisphere or age, and no interactions between any of the variables ($F<1$ in all cases).

![Fig. 6. Overlay of maps of brain activity in the latency range of 120 to 140 ms for the four different stimulus durations and time courses of activation in left and right supra-temporal gyri. Note the progressive dissociation of left hemispheric activity into two separated sources with increasing stimulus duration.](image-url)
For latencies, age had a main effect on P1m ($F(2,39)=25.5$, $p<0.001$) with shorter latency in middle-aged (64 ms) and older (65 ms) compared to young adults (80 ms), ($p<0.005$ in both cases). P1m responses were later ($F(1,39)=11.63$, $p<0.005$) in the left hemisphere than the right hemisphere, however, a group×hemisphere interaction was not significant. The participant’s ages had no main effects on the N1m and P2m latencies nor were interactions between age, duration, and/or hemisphere significant. The N1m latency slightly increased with sound duration, which was best modeled as a quadratic trend ($F(1.38) = 6.08$, $p<0.02$).

**Source analysis based on beamforming**

Although the recorded magnetic field distribution in the latency range of the N1m maximum could be explained to large degree with a pair of single equivalent current dipoles (ECD) it is commonly assumed that more complex configurations of multiple sources underlay the auditory evoked response. Furthermore, changes in stimulus duration may affect brain activity outside the auditory cortex sources as estimated with the N1 model. Support for such an assumption came from inspection of the magnetic field distribution of the evoked response as shown in Fig. 4. Whereas the effect of increasing N1m amplitudes with longer stimulus duration was mostly expressed in left and right temporal sensors, the stimulus duration-related changes in later components showed a different spatial distribution, indicating a different source configuration. Therefore, a model free approach to decomposition of source activity was used to identify brain areas related to variation of stimulus duration.

**Duration dependent source activity**

Volumetric distributions of source activity were estimated for each participant and each stimulus condition. The linear regression of the stimulus duration to the grand averaged data identified those brain regions, which showed the strongest change in activity in relation to the stimulus duration. A list of brain areas is given in Table 1 together with results of the one-way ANOVA confirming significance for the effect of stimulus duration. The brain areas with strongest effect in the 200 ms latency range are visualized in Fig. 5. Maps of regression coefficients, which are measures of the effect size of stimulus duration on changes in brain activity are color-coded overlaid to three orthogonal slices of an MR image of a template brain in Fig. 5a. The cursor is centered at the maximum effect in the anterior part of the insula in the left hemisphere. Time courses of grand averaged activation at this voxel showed that the peak activation and the latency increased with increasing stimulus duration (Fig. 5b). The peak latency was 135 ms for the shortest and 180 ms for the longest sound duration. The latency increase of 45 ms corresponds to stimulus duration increase of 42 ms, which means that the peak latency was approximately 102 ms with respect to the sound offset.

The effect of stimulus duration on the activity in the right hemisphere is characterized by three distinct peaks (Fig. 5c, Table 1). The corresponding time courses of activation are shown in Figs. 5d–f. The activity in the right inferior frontal gyrus showed a single peak at about 150 ms latency (Fig. 5e). The two more posterior sources exhibited pronounced bimodal time courses of activity. The more superior source in the right post-central gyrus (Fig. 5d) showed a peak at about 114 ms independently of stimulus duration and a second peak with latency of 186 ms for the sound of 34 ms duration and 225 ms for the 76-ms duration stimulus. Thus, the latency of the second peak was about 151 ms with respect to the stimulus offset. Similarly, the latency of the first peak in the source activity of right middle temporal gyrus (Fig. 5f) was about 103 ms for all stimulus durations, while the latency of the second peak increased from 234 ms for the 34-ms stimulus to 263 ms for 76-ms stimulus.

Color coded maps of activation in left and right supra-temporal gyri, the locations of auditory cortices, are overlaid to the template MRI in Fig. 6 for the four different stimulus durations. The time courses
of activation showed a dominant peak in the latency range of the N1m response with a mean of 110 ms in the right and 133 ms in the left hemisphere. One remarkable observation in the activation maps is that the left hemispheric activation, which appeared as a single source in STG for the 34 ms stimulus, dissociated into two sources when the stimulus duration increased.

Effects of age on the distributed source activity

The ANOVA for the amounts of activity in the locations of maximum effect of stimulus duration did not show significant effects of age nor interactions between the factors age and stimulus duration. Examples for the source activity, which was observed for the three age groups, are given in Fig. 7. The time series of activation in the anterior part of left insula (Fig. 7a) showed larger variability than the corresponding grand average in Fig. 2b. Nevertheless, the peak amplitudes increased when the stimulus duration was prolonged from 34 to 76 ms for all age groups (see Fig. 7b). The peak latencies were more variable than in the grand average because of larger residual noise. However, a systematic shift toward longer latencies with increasing stimulus duration can be observed from the slopes of activation, with the rising slopes well aligned and the falling slopes progressively delayed for all age groups. The mean peak latencies seem to increase with age from 144 ms in the young group to 209 ms in the middle-aged and older group. Also the modulation of activity by the factor stimulus duration seems to be smaller in the group of older participants than in the other groups. Indeed the effect of stimulus duration on the left insula activation was smallest in the older group (Fig. 7b). However, the error bars indicating the 95% confidence intervals of the mean demonstrate that the difference was not significant. Whereas the effect of stimulus duration was significant for the five discussed areas in all age groups, no differences between age groups were significant.

Discussion

Using the beamformer approach for neuromagnetic source imaging a network of cortical areas was identified which responded systematically to changes in stimulus duration. Specific changes in activity were observed in the latency range of the N1m and P2m response in the anterior part of the left insula, right inferior frontal gyrus, right middle temporal gyrus, right post-central gyrus, and bilateral superior temporal gyri. Although, magnitude and latencies of AEFs changed with age, the effects of sound duration were consistently observed in all age groups.

Hemispheric asymmetry of the AEF

Although the sound stimuli were presented simultaneously to both ears with equal intensity, we observed distinct hemispheric asymmetries in the brain responses. The P1m wave of the AEF, generated in medial part of Heschl’s gyrus (Godey et al., 2001), was larger in the left than the right hemisphere and this laterality was consistent across age groups. Our results corroborated previous reports of left lateralized P1m in response to diotic sounds (Chait et al., 2004) and supported the hypothesis that asymmetrical sound processing is already established in primary auditory cortices within the first 100 ms after sound onset. Such hemispheric asymmetry in response magnitude reflects structural asymmetry in humans with typically larger primary auditory cortices in the left hemisphere (Galaburda et al., 1978; Penhune et al., 1996; Rademacher et al., 1993). Functional relevance of such asymmetry has been demonstrated in fMRI studies (Devlin et al., 2003). In contrast to left lateralization of P1m, the N1m response was largest in the right hemisphere suggesting multiple auditory pathways involving both hemispheres. Our result of distributed source configurations for sound duration-sensitive areas in left and right hemispheres is in line with the concept of specific asymmetries in processing of elementary auditory features involving both hemispheres (Jamison et al., 2006).

Age-related changes in the AEF

Normal aging was associated changes in the morphology of response waveforms. Most prominent age-related modification of the AEF was an increase in P1m amplitude, which is consistent with findings from prior studies using different material and presentation rate (Kovacevic et al., 2005; Pekkonen et al., 1995; Ross and Tremblay, 2009). The decrease in P1m latency with age is likely concomitant with the P1m amplitude increase. Because of partial overlap of P1m and N1m in the AEF waveform the P1m peak shifts to earlier latency when P1m increases but N1m stays unchanged. Age-related increases in middle- (e.g. Na, Chao and Knight, 1997; Woods and Clayworth, 1986) and long-latency (e.g. N1, Alain and Woods, 1999) evoked responses have also been observed using neuroelectric recordings, and were associated with impairment in inhibitory control mediated by the prefrontal cortex (Alain and Woods, 1999; Chao and Knight, 1997). Such a proposal is supported by lesion studies in humans, which showed greater middle-latency auditory evoked responses following lesions to the dorsolateral prefrontal cortex (Alho et al., 1994; Knight et al., 1989). Moreover, age-related changes in the afferent auditory pathway could account for enhanced P1m amplitudes. Indeed, animal studies have shown that aging impairs inhibitory processes in the dorsal cochlear nucleus (Caspar et al., 2005), which could yield larger cortical evoked responses. In rodents, induced sensorineural hearing loss increased thalamocortical excitation and excitability in auditory cortex (Kotak et al., 2005). In humans, aging decreased contralateral suppression, as evidenced by reduced otoacoustic emissions in the presence of contralateral white noise (Kim et al., 2002). This lack of suppression is consistent with reduced inhibitory controls from the medial olivocochlear system on the outer hair cell, which could also affect amplitudes and latencies of cortical evoked responses. Therefore, the enhanced P1m amplitude may arise from age-related decline in inhibitory control that occurs at various levels in the ascending auditory pathways as well as in the higher cortical areas such as in the prefrontal cortex.

Another effect of aging on the AEF was the increase in N1m amplitude in the left hemisphere, which was already evident in the middle-aged group. The increase in N1m amplitude with age diminished the right hemispheric N1m dominance in the young toward symmetrical responses in the older groups. For speech stimuli N1 amplitudes were most pronounced in the left hemisphere, however, the responses were balanced in older listeners (Bellis et al., 2000). Age-related reduction of functional asymmetry has been shown for prefrontal activity during cognitive control of semantic and working memory, and perception (Cabeza, 2002) and is discussed in relation to compensatory mechanism or alternatively different trajectories of aging for both hemispheres. Our finding of change in laterality in basic sensory function may indicate more global reduction of asymmetry.

Effects of sound duration on the AEFs

Sound duration-related changes became most obvious when we subtracted a model of the onset response from the waveforms of cortical source activity. The resulting difference waveforms showed a negative wave that overlapped with the P2m deflection and apparently reduced the response amplitude at about 200 ms with increasing sound duration. This does not contradict previous findings of increasing P2 with stimulus duration (Alain et al., 1997; Ostroff et al., 2003), because the very short stimuli in these studies likely have not elicited a noticeable off response. In the current study, the latency of the negative deflection elicited by sound offset overlapped with the
positive P2m wave elicited by sound onset. The effect of the sound offset was noticeable nonlinear with respect to the change in sound duration. The increase of 12 ms in duration between the 34 and 46 ms stimuli elicited a just noticeable response whereas the same absolute increase between 64 and 76 ms elicited the strongest response. This indicates to a lower limit of about 40 ms for using the sound offset marker for duration encoding and very effective representation of the offset marker for longer sound duration. The lower limit of 40 ms coincides with the upper limit of 40 ms for which sound duration-related increase in the N1 amplitude had been shown. Thus the data demonstrate a tradeoff between a linear integrator mechanism for short sound duration and importance for the onset and offset markers for sound durations longer than 40 ms.

The duration-related difference waves were comparable in all three age groups and from this perspective processing of sound duration based on detecting the sound offset is little affected by normal aging. An effect of age had been shown for shorter sound of 10 to 18 ms in duration in an EEG study (Ostroff et al., 2003). These findings from EEG and MEG recordings corroborate results of a behavioural study, which showed that older adults need a larger proportion of duration increase (Weber fraction) before the difference becomes noticeable for stimuli shorter than 20 ms than do younger adults, but that the age difference disappears at longer durations (Bergeson et al., 2001). Behavioural findings for even longer sound durations (<250 ms) showed increasing duration difference limens in older adults (Fitzgibbons and Gordon-Salant, 1994). This cannot be directly transferred to behavioural performance in discrimination of sound duration on the order of tens of milliseconds, which is comparably poor in all ages (Fitzgibbons et al., 2007). One possibility for differences between behavioural performance and early evoked responses could be that sound duration is adequately encoded in primary auditory cortex, but that older adults have problems or need more time to access and use that information in order to make their response. Using complex sounds comprised of either all tuned or one mistuned harmonic, Alain and McDonald (2007) found an early mistuned-related change in neuroimagnetic brain activity, indicating the segregation of sound into multiple objects, that was comparable in young and older adults, despite an age-related decline in listeners’ likelihood of hearing the mistuned harmonic as a separate object. Coupled with our results, these findings suggest dissociation between neural activity indexing physical stimulus-related differences and the actual perceptual difference. Those studies, like our current one, presented stimuli in silence at comfortable loudness. Performances in noisy environment and with multiple speakers may be substantially degraded.

Brain areas encoding time duration information

Our findings of distinct brain areas involved in auditory temporal processing and specifically in duration processing on the scale of tens of milliseconds are consistent with several reports from previous studies. The strongest effect of stimulus duration in the left hemisphere was found in the anterior part of the left insula. Multiple anatomical connections between the auditory pathway and the insula cortex are known, and the importance of the insula cortex for basic auditory processing, specifically temporal processing, has been discussed (Bamiou et al., 2003). Moreover, activation of the left insula related to musical rhythm has been found in a PET study (Platel et al., 1997) in which brain activity in response to tones of irregular duration and temporal intervals was compared to regular tones in same pitch range. While patients with insular stroke had significant deficits in auditory temporal processing, sound duration discrimination was stronger affected in patients with left than right insular lesions (Bamiou et al., 2006). Right inferior frontal cortex activation has been shown in a PET study with pure-tone stimuli of 50 ms duration, which infrequently changed to 100 ms duration tones (Dittmann-Balcar et al., 2001). Broad right hemispheric activation, including inferior frontal and middle temporal gyr, has been shown for sound duration discrimination in fMRI (Belin et al., 2002). Moreover, activation of the right middle temporal gyrus was correlated with behavioural performance in the latter study. In contrast to PET and fMRI measures of changes in the BOLD signal, which integrate brain activity over a long time interval including perceptual and cognitive responses, the MEG study identified those cortical areas, which were differentially active as early as 100–200 ms after stimulus onset and offset.

Relation to components of the AEF

Eight different stimuli had been presented in randomized order in this study. On the first view, such stimulus paradigm does not constitute a classical mismatch design for which a standard stimulus is infrequently changed into a deviant version. In the present study, stimulus change occurred with respect to multiple stimulus features like gap duration or duration of the leading marker. Could the observed activity in the 200-ms latency range have been a part of the MMN generating network, given that multiple deviations of different stimulus features within the stimulus sequence can elicit multiple mismatch negativity responses (Naatanen et al., 2004)? We think, this is unlikely because the total amount of change was similar for all stimuli and equal size of change response would be expected for all stimulus durations, which was not the case. Moreover, the observed activity was strongly related to the parameter stimulus duration. The latencies of the duration-sensitive subcomponents of the N1, 110 ms in right and 133 ms in left auditory cortex, were longer than the 100 ms peak latency of the N1m observed in the AEF based on dipole modeling, and the peak latency of 160 ms for right inferior frontal activity was even longer. This is consistent with the report that the later components of the N1 with 100–130 ms latency were most responsive to sound duration (Alain et al., 1997). However, comparison between source activity and components of the AEF is not always feasible because of lacking polarity information.

Neural encoding of sound duration

Increases in magnitude with increasing stimulus duration had been found for the response components with peak latencies in the 100–130 ms range in left and right superior temporal gyr, in the right middle temporal gyrus and in right post-central gyrus. Whereas, the latencies of these components were constant despite changing stimulus duration, the later response components in the left insula, right middle temporal gyrus and in right post-central gyrus showed progressively increasing latencies in addition to increases in magnitude related to stimulus duration. The early responses likely reflect the proposed mechanism of temporal integration. It seems that sub-components of the N1 response show temporal integration at least up to 76 ms, the longest stimulus duration used in this study. This is a wider range of integration than the 20 to 40 ms suggested from previous ERP recordings (Alain et al., 1997; Forss et al., 1993; Gage and Roberts, 2000; Joutsiniemi et al., 1989; Onishi and Davis, 1968). The later components with latencies locked to the stimulus offset potentially encode the end point of the stimulus. Thus in the duration range of 34 to 76 ms, as tested in this study, potentially a combination of linear integration and detection of onset and offset events could be used as an effective mechanism of encoding stimulus duration.

Conclusions

We demonstrated the importance of detecting onset and offset markers in sound for neural encoding of sounds longer than about 40 ms. Shorter sounds are likely encoded entirely through an integration mechanism. Distributed cortical networks in left and right hemispheres are involved in sound duration encoding, which
takes place within the first 250 ms after sound onset. These observations were consistent across a wide range in the age of participants contributing to the notion that the neural encoding of elementary sound information is little affected by normal aging.

Acknowledgments

The study was supported by grants from the Canadian Institutes for Health Research (Grant No. 81135), the Canadian Foundation for Innovations and The Hearing Foundation of Canada.

References


