Effects of an acute acoustic trauma on the representation of a voice onset time continuum in cat primary auditory cortex

Masahiko Tomita, Arnaud J. Noreña, Jos J. Eggermont *

Departments of Physiology and Biophysics, and Psychology, Neuroscience Research Group, University of Calgary, 2500 University Drive N.W., Calgary, Alta., Canada T2N 1N4

Received 1 December 2003; accepted 2 March 2004

Available online 24 March 2004

Abstract

Here we show that hearing loss associated with an impairment of speech recognition causes a decrease in neural temporal resolution. In order to assess central auditory system changes in temporal resolution, we investigated the effect of an acute hearing loss on the representation of a voice onset time (VOT) and gap-duration continuum in primary auditory cortex (AI) of the ketamine-anesthetized cat. Multiple single-unit activity related to the presentation of a /ba/-/pa/ continuum – in which VOT was varied in 5-ms steps from 0 to 70 ms – was recorded from the same sites before and after an acoustic trauma using two 8-electrode arrays. We also obtained data for gaps, of duration equal to the VOT, embedded in noise 5 ms after the onset. We specifically analyzed the maximum firing rate (FRmax), related to the presentation of the vowel or trailing noise burst, as a function of VOT and gap duration. The changes in FRmax for /ba/-/pa/ continuum as a function of VOT match the psychometric function for categorical perception of /ba/-/pa/ modeled by a sigmoid function. An acoustic trauma made the sigmoid fitting functions shallower, and shifted them toward higher values of VOT. The less steep fitting function may be a neural correlate of an impaired psychoacoustic temporal resolution, because the ambiguity between /ba/ and /pa/ should consequently be increased. The present study is the first one in showing an impairment of the temporal resolution of neurons in AI caused by an acute acoustic trauma.

© 2004 Elsevier B.V. All rights reserved.

Keywords: Noise trauma; Hearing loss; Voice onset time; Speech perception; Temporal resolution

1. Introduction

Animal vocalizations and speech in humans, are characterized by distinct amplitude fluctuations. It has been demonstrated that the auditory system makes effective use of these temporal cues for perception. In this context, Shannon et al. (1995) showed that noise-bands modulated with the temporal envelope of speech, producing an acoustical signal with strongly degraded spectral information but with preserved temporal envelope cues, was sufficient to allow a correct identification of consonants, vowels and words. Moreover, amplitude fluctuation can be used by the auditory system in auditory scene analysis. Indeed, different frequency components coherently modulated may be interpreted as being related to the same source or auditory object (Bregman et al., 1990).

Amplitude fluctuations can be defined as periodic (as in vowels) or aperiodic. The voice onset time (VOT), the interval with aspiration noise between the consonant release and voicing onset, is an aperiodic amplitude fluctuation. The difference between /ba/ and /pa/ phonemes, for instance, depends to a large extent on the difference in VOT (Kuhl and Miller, 1978). A VOT of 0 ms is generally associated with the perception of /ba/, whereas a VOT of 25 ms and higher results in the perception of /pa/. The perception changes abruptly from voiced to voiceless at this boundary (Abramson and Lisker, 1970). Interestingly, this boundary does not depend on a specific human
speech mechanism and was reported in chinchillas (Kuhl and Miller, 1978) and monkeys (Morse and Snowden, 1975; Sinnott and Adams, 1987) at approximately the same value. The fact that the perceptual boundary was reported in other mammals than humans suggests that evolution may have taken advantage of specific properties of the central nervous system to design speech-sounds (Kuhl and Miller, 1978). VOT boundaries also shift with place of articulation and the boundaries for /ta/–/da/ and /ka/–/ga/ are higher than for the /ba/–/pa/ distinction. 

Neural correlates of these perceptual boundaries have been investigated in the auditory nerve of chinchillas (Sinex and McDonald, 1988, 1989), inferior colliculus of chinchillas (Chen et al., 1996) and in the auditory cortex of monkey (Steinschneider et al., 1995, 2003) and cat (Eggermont, 1995a,b). Steinschneider et al. (1995) examined multi-unit activity elicited by the consonant vowel syllable /da/ and /ha/ that varied in VOT (0, 20, 40 and 60 ms). They found responses time-locked to stimulus onset and to the onset of voicing with 40 ms VOT. The response to voicing onset was markedly diminished with 20 ms and was similar to that evoked by /da/ with 0 ms VOT. These results were seen in 59% penetrations, which display the “double-on” response pattern. Eggermont (1995b) studied the representation of VOT for a /ba/–/pa/ continuum in which VOT was varied in 5-ms steps from 0 to 70 ms. He found that the mean value for the minimum detectable VOT, which was represented in onset responses to both the voiceless and voiced parts of the sound, was 46 ms in adult cat. Comparing the results obtained in awake monkeys (Steinschneider et al., 1995, 2003) and anesthetized cats (Eggermont, 1995a,b) suggests that the minimum VOT is little affected by anesthesia.

The role of aperiodic amplitude fluctuations in the understanding of speech has also been studied by using gap-detection paradigms (Schneider and Hamstra, 1999; Snell and Frisina, 2000; Snell et al., 2002). Eggermont (1995a, 1999, 2000) focused on the cortical representation of gaps embedded in noise. For short leading noise-burst duration (5 ms), he found mean values for the minimum neural gap threshold (MNGT) around 35 ms, and similar to those obtained with /ba/–/pa/ continuum.

Hearing loss and aging, conditions known as being associated with an impairment of speech understanding (CHABA, 1998; Humes, 1991; Abel et al., 1990), are both linked to a decrease in temporal resolution (Moore, 1985, 1993; Gordon-Salant and Fitzgibbons, 1993; Snell et al., 2002; Mazelova et al., 2003). Thus, impairment in temporal resolution, as measured in a gap-detection task (Walton et al., 1998) or evoked potentials by a VOT continuum (Tremblay et al., 2003), may be related to the decline in the understanding of speech. It has been demonstrated that gap-detection thresholds were impaired in older subjects compared to those in younger subjects (Snell, 1997). Consistent with this finding, Walton et al. (1998) reported an age-related alteration of temporal resolution of neurons in inferior colliculus. It is suggested that the decline in temporal resolution during aging may involve, in addition to a deficit in the function of the cochlea, a central component. Numerous studies have shown that a hearing loss is followed by changes in central inhibition (Rajan, 1998, 2001; Wang et al., 2002; Norrena et al., 2003). However, there is no electrophysiological study in the literature addressing the effect of hearing loss on the cortical representation of VOT, i.e., on the temporal resolution of the auditory system.

In the present study, we investigated the effect of an acute hearing loss on the representation of a voice onset time continuum (ba/–/pa/ phonemes) in primary auditory cortex. Multiple single-unit (MU) activity related to the presentation of the /ba/–/pa/ continuum — in which VOT was varied in 5-ms step from 0 to 70 ms — was recorded from the same recording sites before and after an acoustic trauma. Moreover, MU activity related to the presentation of gaps inserted 5 ms after the onset of the noise (“early” gap) in a 1-s noise-burst was also recorded. Gap duration was varied in 5-ms step from 0 to 70 ms. The goal of our study is to gain insight into the effects of an acute hearing loss on the cortical representation of VOT and gaps. More generally, our study is aimed at pointing out potential central changes after a hearing loss that are associated with a decline in speech understanding in humans.

2. Method

2.1. Animal preparation

The care and the use of animals reported on in this study were approved by the Life and Environmental Science Animal Care Committee of the University of Calgary (BI-2001-021). All cats were anesthetized with administration of 25 mg/kg of ketamine (100 mg/ml) injected intra-muscularly followed after approximately 10 min by 20 mg/kg of pentobarbital sodium (65 mg/ml) intra-muscularly and 0.25 ml/kg body weight of a mixture of 0.1 ml acepromazine (0.25 mg/ml) and 0.9 ml of atropine methyl nitrate (5 mg/ml) subcutaneously. Lidocain (20 mg/ml) was injected subcutaneously and a skin flap and muscle tissue overlying right temporal bone was removed and the skull cleared. A large screw was cemented upside down on the skull with dental acrylic. Two 8 mm diameter holes were trephined over the right temporal cortex so as to expose parts of primary auditory cortex. The holes were enlarged to expose both anterior and posterior ecysylvian sulci if needed. The dura was left intact, and the brain was covered with light mineral oil to prevent tissue drying. Then the cat was placed in a sound-treated room on a vibration iso-
2.2. Peripheral threshold estimation

Prior to acute recordings and in one-third of the cases after the end of experiment, approximately 6 h after the exposure, peripheral hearing sensitivity was determined from auditory brainstem response (ABR) threshold. For this purpose, tone pips with frequencies of 3, 4, 6, 12, 16, 24 and 32 kHz were presented at 10/s in an anechoic room from a speaker placed 45° from the midline and at 55 cm distance from the cat’s head. The sound-treated room was made anechoic for frequencies above 625 Hz by covering walls and ceiling with acoustic wedges (Sonex 3°) and by covering exposed parts of the vibration isolation frame, equipment and floor with wedge material as well. Calibration and monitoring of the sound field was done using a Brüel and Kjær (type 4134) microphone placed above animal’s head and facing the loud speaker. ABRs were recorded, with needle electrodes in ipsi- and contralateral muscles covering both mastoids, in response to γ-function shaped tone pips with an effective (50% from peak) duration of 15 ms. In this recording montage, the first negative–positive component representing the compound action potential of the auditory nerve is frequently the largest component. The signals were amplified between 300 and 3000 Hz using a DAM 500 (World Precision Instruments) differential amplifier and averaged with a Brüel and Kjær (type 2034) dual signal analyzer in the signal enhancement mode. Artifact rejection and local lidocaine infusion were used to avoid contamination of the ABR by muscle action potentials. At high intensity levels 20–50 averages sufficed but at near threshold values 200 averages were obtained and repeated once. Step size was 10 dB, except around threshold where it was 5 dB. Threshold was defined as the lowest intensity that yielded a reproducible response.

2.3. Acoustic stimulus presentation

Stimuli were generated using MATLAB® and transferred to the DSP boards of a TDT (Tucker Davis Technologies) sound delivery system. Acoustic stimuli were presented in an anechoic room from a speaker placed 45° from the midline 55 cm distance from the cat’s left ear. Characteristic frequency (CF) and tuning properties of individual neuron were determined with γ-shape envelope tone pips. These tone pips with a half-peak-amplitude duration of 15 ms and a γ-function-shaped envelope, were presented at a rate of 1/s in a pseudo-random frequency order at fixed intensity level. The 81 different frequencies used were equally spaced logarithmically between 625 Hz and 20 kHz (or between 1.25 and 40 kHz) so that 16 frequencies were present per octave. Each recording session comprised of up to nine different intensity series in 10 dB steps of these tone pip stimuli.

After the frequency–tuning properties of the cells at each electrode were determined, the /bal–/pal/ continuum with VOTs ranging from 0 to 70 ms in steps of 5 ms was presented in random order as described previously (Eggermont, 1995b). These phonemes were generated with a parallel/cascade Klatt-synthesizer KLSYN88a, using a 20-kHz sampling frequency. The total duration of the stimulus was 250 ms, regardless the duration of the VOT, and the onset noise burst was 5 ms in duration (leading noise burst). The dominant frequency ranges for the vowel part were \( F_0 = 120 \) Hz, \( F_1 = 700 \) Hz, and \( F_2 = 1200 \) Hz. The fundamental frequency started at 120 Hz, remained at that value for 100 ms, and then dropped from there to 100 Hz at the end of the vowel. \( F_1 \) started at 512 Hz and increased from 25 ms to 700 Hz, \( F_2 \) started at 1019 Hz and increased from 25 ms to 1200 Hz, and \( F_3 \) changed in the same time span from 2153 to 2600 Hz. In order to match the CF of neuons to the spectral content of the vowel, its spectrum was transformed so that the upper edge of the \( F_2 \) band was above the CF of neurons. The phonemes were presented once every 2 s. The only parameter that changed in the continuum was the VOT; formant transition durations were not altered. Twenty repetitions were presented at each VOT value so that the entire stimulus sequence lasted 600 s.

Gaps ranging from 5 to 70 ms in duration were placed in two positions in wideband noise bursts of 1 s in duration that were presented once per 3 s in random order. The first position, the “early gap,” started 5 ms after the noise burst onset, and the second position, the “late gap,” was positioned 500 ms after noise burst onset. Only results for the early gap will be presented as they relate to the VOT results (Eggermont, 1999). Each gap stimulus was presented 15 times. The noise burst used consisted of “frozen” noise, i.e., the pseudo-random noise sequence was the same for all conditions. These stimuli (a continuum of /bal–/pal/ phonemes and a stimulus detecting early and late gap) were presented at peak intensities of 65 or 75 dB SPL and the intensity level was same before and after the acoustic trauma.
2.4. Acoustic trauma

The acoustic trauma was induced by a 1-h exposure to a pure-tone at the maximum output level. The trauma tone frequency (TF) was set at 5 or 6 kHz. The sound was presented by a high-power amplifier (Samson Servo 240) and loudspeaker (Yorkville 120 W) placed 75 cm away from the cat’s head. Measured at cat’s head, the trauma tone fundamental frequency had a level of 115–120 dB SPL; whereas the second and third harmonics had a level of 25 and 38 dB lower than that of fundamental frequency, respectively. The pure tone exposure was administered about 8 h after the start of the preparation, following typically 4 h of recording at the selected recording sites. The post-exposure threshold at CF was generally obtained with γ-shape envelope tone pips between 1 and 1.5 h after the end of the exposure, and just prior to recording responses to the VOT and gap stimuli.

2.5. Recording and spike separation procedure

Two arrays of eight electrodes (Frederic Haer Corp.) each with impedances between 1 and 2 MΩ were used. The electrodes were arranged in a 4×2 configuration with inter-electrode distance within rows and columns equal to 0.5 mm. Each electrode array was oriented with the long axis parallel to the midline in primary auditory cortex located in majority between anterior and posterior ectosylvian sulci. The arrays were manually and independently advanced using Narishige M101 hydraulic microdrives. The signals were amplified 10,000 times using a Frederic Haer Corp. HiZx8 set of amplifiers with filter cut-off frequencies set at 300 Hz and 5 kHz. The amplified signals were processed by a DataWave multi-channel data acquisition system. Spike sorting was done off line using a semi-automated procedure based on principal component analysis (Eggermont, 1990) implemented in MATLAB. The spike times and waveforms were stored. The data presented in this paper represent only well separated single units that, because of their regular spike wave form, likely are dominantly from pyramidal cells, thereby minimizing potential contributions from thalamocortical afferents or interneurons (Eggermont, 1996). For statistical purposes, the separated single-unit spike trains were added again to form a multi-unit spike train. We have previously shown that the CFs of neurons recorded from the same electrode are similar (Eggermont, 1996).

2.6. Data analysis

To assess frequency–tuning properties, the peak number of action potentials in the post-stimulus time histogram (PSTH, 5 ms bins) calculated over the first 100 ms after γ-tone presentation was estimated. The peak counts for three adjacent frequencies were then combined, in order to reduce variability, and divided by number of stimuli and presented as a firing rate per stimulus. This resulted in 27 frequency bins covering 5 octaves so that the final frequency resolution for determining the CF was approximately 0.2 octaves. The results were calculated per stimulus intensity, and were combined into an intensity–frequency–rate profile from which tuning curves, rate-intensity functions and iso-intensity rate-frequency contours could be derived (Eggermont, 1996) using routines implemented in MATLAB. The frequency–tuning curve was defined for a firing rate at 25% of the maximum peak-firing rate. This was about 10–20% above the background firing-rate, but as the latter was dependent on the level of stimulus-induced suppression, the tuning curve criterion based on a percentage of peak firing rate was preferred over that based on increase over background activity. In order to calculate threshold shifts with the acoustic trauma for each cat, all tuning curves for each cat were drawn as the contour lines at 25% from maximum response as function of a frequency before and after the acoustic trauma (Fig. 1(a) and (b)). The lower envelope of the set of frequency-tuning curves was traced (Fig. 1(c)) and the difference of the values at the lowest threshold between before and after the noise exposure was estimated as the cortical threshold shift as a function of frequency (Fig. 1(d)).

---

Fig. 1. Changes in frequency tuning of MU responses recorded in one cat. The frequency–tuning curves of each recording site obtained before and after the acoustic trauma are shown in panels (a) and (b), respectively. Panel (c) shows the lower envelopes of all frequency–tuning curves in panels (a) and (b). The difference of the value estimated from contour lines in panel (c) are plotted at 3, 4, 6, 8, 12, 16, 24 and 32 kHz in panel (d). ABR threshold shifts obtained 5 h after the trauma are shown as dot line in panel (d).
Post-stimulus time histograms with a 5-ms binwidth were used to represent the recordings for the /ba/-/pa/ continuum and early stimuli. From these PSTHs both the peak latency and the spike count were calculated. The minimum neural VOT (MNVOT) was defined as the minimum VOT at which neural responses are time-locked to both the burst onset and the onset of voicing. Similarly, the minimum neural gap threshold (MNGT) for “early” gap was defined as the minimum gap duration at which neural responses are time-locked to both the leading and the trailing noise burst. The responses related to the vowel and the trailing noise burst were analyzed only when an onset response to the leading noise burst was clearly visible. The MNVOT and MNGT were obtained from inspection of the dot displays and depended on the number of stimulus presentation and the firing rate of the neurons. We presented 20 repetition for each VOT duration of a /ba/-/pa/ continuum and required at least four time-locked spikes per 20 presentations to the onset of voicing for assigning the MNVOT. Similarly, in gap stimuli 15 repetition for each gap duration were represented and three time-locked spikes per 15 presentations to the trailing noise burst were required to assign the MNGT.

Moreover, the maximum firing rate (FRmax) related to the responses for the voicing or the trailing noise burst for each VOT or gap duration were derived from PSTHs. In order to avoid the increase in firing rate related to a late rebound response, only spikes that occurred within a time window of 30 ms starting at a time equal to the peak latency of the leading noise burst plus VOT or gap duration were included in the analysis.

All statistical analyses were performed using Statview 5®. Illustrations and figures were prepared with MATLAB® and SigmaPlot®.

3. Results

Recordings were made from the right auditory cortex in 13 cats. For a total of 46 recording sites out of 103 that showed reliable frequency–tuning curves before and after the noise trauma, we recorded clear MU responses to the VOT or gap stimuli approximately 1–1.5 h before and approximately 1.5–2 h after the acoustic trauma. The consonant–vowel stimuli were frequency transformed so that the upper edge of the F2 band was above the neuron’s pre-trauma CF. The study presented here includes data collected in the same animals as those in two other studies. Namely, the effects of an acoustic trauma on spontaneous activity and neural synchrony were described in Noreña and Eggermont (2003). Moreover, the effects of an acoustic trauma on frequency response properties were reported in Noreña et al. (2003). The cortical threshold shift in five cats as a function of the difference between the frequency at which ABR thresholds have been measured and the TF. One notes that ABR threshold shift is around 40 dB in average above the TF. The cortical threshold shift, for the recording sites that yielded responses to the /ba/-/pa/ and gap stimuli, as a function of the difference between the pre-trauma CF and the TF. The cortical threshold shift is around 20 dB above the TF and is larger for the frequency band above TF than
below TF. The average threshold elevation is 13.1 dB for MUs with CF below and above TF and significantly different from 0 ($t(43) = 5.287, p < 0.001$). As noted above, these threshold shifts were calculated at the CFs, which typically changed dramatically after the noise trauma and thus do not represent the peripheral hearing loss at particular frequencies. The increase in the thresholds at CF in the subset of 46 recordings analyzed here was nearly identical to that in the overall set of 124 neurons analyzed previously (Noreña et al., 2003).

3.2. Individual example

Fig. 4 shows neural responses obtained at the same recording site for presentation of the /ba/-/pa/ continuum ((a)–(c)) and the “early” gap stimuli ((d)–(f)) (see Section 2). Left-hand panels ((a) and (d)) show dot displays: each dot corresponds to a spike. The dot displays are organized vertically according to gap duration or VOT and horizontally for time since leading noise burst onset. Middle panels ((b) and (e)) represents the PSTHs (the bin size is 5 ms) of the data shown in panels (a) and (d), respectively. The oblique dotted lines indicate the time windows (30-ms duration) in which the FRmax, for each gap duration or VOT, is derived. Right-hand panels ((c) and (f)) show the FRmax, for each gap duration or VOT, related to the presentation of the trailing noise or the vowel.

One notes first that the responses related to the presentation of the /ba/-/pa/ continuum are similar to those related to the presentation of the “early” gap stimuli. In both cases, neural responses are evoked at the onset of the stimuli. Moreover, neurons are activated by the trailing stimulus only if it is presented at more than
40 ms after the leading stimulus. Namely, the MNVOT and MNGT have similar values, around 40 ms. In this example, the background firing rate is very low; neurons do not respond to the second stimulus (vowel or trailing noise) for VOT below 40 ms.

The data were fitted by a sigmoid function
\[ Y = y_0 + a / \left(1 + e^{k \frac{(X-x_{50})}{C_0}}\right), \]
where \(Y\) is the FRmax, \(X\) is VOT or gap duration in ms, \(x_{50}\) is the VOT or gap duration corresponding to the 50% point of the function, \(y_0\) is the background activity, \(a\) is the difference between maximum and minimum of the sigmoid function, and \(k\) is a factor defining the slope of the function. The middle (50%) points between minimum value and maximum value of FRmax from the fitting curves to obtained data (panels (c) and (f)) are very close to the value of the minimum VOT or gap duration compared to that derived from visual inspection in the left (dot displays) and middle panels (PSTH). Namely, MNVOT as estimated from \(x_{50}\) is 42 ms for /ba/-/pa/ stimulus continuum and the MNGT is estimated at 37 ms.

3.3. Group data for the /ba/-/pa/ continuum and the “early” gap condition

As shown in Fig. 3, the average CF-threshold shift is larger for the frequency band above TF than that below TF. Consequently, we divided the data of the MU recordings into two groups based on the CF and TF. However, we found no significantly difference (with respect to onset response, FRmax, MNVOT and MNGT) between the groups with CF above and below TF, and thus pooled the data into a single group. All the following results then concern the entire group. Fig. 5 shows the distribution of MNVOT and MNGT values before and after the acoustic trauma. One notes that the distribution in MNVOT and MNGT values presents a peak around 40 ms before the acoustic trauma. After the acoustic trauma, the distribution is more uniform and the number of MUs with MNVOT and MNGT above 40 ms is increased. The averaged MNVOT is 31 ms (SD = 13 ms) and 35 ms (SD = 17 ms) before and after the acoustic trauma, respectively. The averaged MNGT is 36 ms (SD = 17 ms) and 38 ms (SD = 18 ms) before and after the acoustic trauma, respectively. The right-hand panels show the distribution in changes of MNVOT and MNGT with the acoustic trauma. The number of MUs with difference of MNVOT and MNGT above 0 ms is larger than that below 0 ms, although the increase of MNVOT and MNGT with the acoustic trauma was not significantly different using one sample t-test hypothesized mean of zero (/ba/-/pa/; \(p = 0.11\); early gap; \(p = 0.17\)).

Fig. 6 shows the changes in MNVOT and MNGT as a function of the threshold shift at CF for the /ba/-/pa/ continuum (a) and the “early” gap condition (c). And panels (b) and (d) show the changes in MNVOT and
MNGT as a function of the difference between the pre-trauma CF and the TF for the /ba/-/pa/ continuum (b) and the “early” gap condition (d). One observes a tendency for the MNVOT or the MNGT to be dependent on threshold shift, but this was only significant for MNVOT ($R^2 = 0.206, p = 0.013$). Changes in MNVOT or MNGT were not significantly correlated with the difference between the pre-trauma CF and TF.

Fig. 7 shows the averaged PSTHs in response to the leading noise burst. We averaged the PSTHs obtained from MU recordings of stimulus representation with 25–45 ms of VOT or gap duration. For each bin, the changes in PSTH (before vs. after the trauma) were statistically tested (paired $t$-test). There was no significant difference between firing rates before and after the acoustic trauma in both the /ba/-/pa/ continuum and “early” gap conditions.

Fig. 8 shows the average of the normalized FRmax as a function of VOT or (“early”) gap duration. The normalized FRmax were obtained by dividing the FRmax obtained at each VOT- or gap-duration by the highest FRmax to the trailing stimuli over all VOTs or gap durations. As shown for individual recordings in Fig. 4, the largest FRmax were usually obtained for highest values of VOT or gap duration, namely around 70 ms. An ANOVA showed that the data were very well fitted by a sigmoid function (Eq. (1)); pre-trauma /ba/-/pa/: $R^2 = 0.96, p < 0.001$, post-trauma /ba/-/pa/: $R^2 = 0.98$, $p < 0.001$, pre-trauma early gap: $R^2 = 0.97, p < 0.001$, and post-trauma early gap: $R^2 = 0.97, p < 0.001$. 

Fig. 6. Differences of MNVOT ((a) and (b)) and MNGT ((c) and (d)) before and after the acoustic trauma as a function of threshold shift at CF (left column) and as a function of the difference between CF and TF (right column).
One notices that the shape of the fitting curves before the trauma is similar for /ba/–/pa/ and “early” gap stimuli. Below a VOT or gap duration of 20 ms, the FRmax is low and constant and corresponds to the background activity (the trailing stimulus does not evoke stimulus-locked responses). Between 30 and 50 ms, the FRmax increases abruptly as a function of VOT and gap duration. Finally, the FRmax reaches a plateau from values of VOT or gap duration around 50 ms. The middle point ($x = 50$: the 50% detectable point) from minimum to maximum value of FRmax gives similar values for MNVOT and MNGT, namely 35 and 38 ms, respectively.

After the trauma, the shape of the fitting curves for both /ba/–/pa/ continuum and “early” gap condition is changed. The most striking difference compared to the pre-trauma condition is that the fitting curves are shifted toward longer values of VOT or gap duration for both conditions. And the slope of the fitting function is also less steep for /ba/–/pa/ continuum. Indeed, slope values derived from the fitted data were 0.017 and 0.014 for /ba/–/pa/ continuum in the pre and post-trauma condition, respectively. For the gap condition the slope stayed the same at 0.019. As a consequence, the estimated MNVOT and MNGT from the fit functions are increased after the trauma compared to the pre-trauma condition: MNVOT and MNGT derived from the fitted data for post-trauma condition were 46 and 51 ms, respectively. Moreover, FRmax does not reach a plateau at high values of VOT or gap duration after the trauma. Again, after the trauma, the shape of the fitting curves was similar between /ba/–/pa/ and “early” gap stimuli.

As we described above, the shape of the fitting curve to data obtained before the trauma consists of three parts divided by VOT or gap duration, a short VOT or gap duration part with mostly background activity, a moderate duration part with abrupt changes of FRmax and a long duration plateau. In order to estimate the change of these FRmax values with the acoustic trauma, we divided the data into three groups by VOT or gap duration, namely short (0–20 ms) duration group, border (25–45 ms) duration group around MNVOT and MNGT, and long (50–70 ms) duration group. The change of FRmax was statistically tested using a repeated measures ANOVA with repeated factor of time of recording (before and after the acoustic trauma). For /ba/–/pa/ continuum, there was a significant decrease of FRmax with the acoustic trauma for the border group ($F = 9.829$, $p = 0.002$) and no significant change in the other two groups. For the early gap condition, the FRmax in the short duration group and border group...
were also significantly decreased with the acoustic trauma \((F = 12.48, \ p = 0.0005, \ F = 27.18, \ p < 0.0001\), respectively).

4. Discussion

The results of the present study can be summarized as follows. There were no significant differences in the mean MNVOT and MNGT to the /bal-/pa/ and gap-in-noise stimuli. The mean MNVOT and MNGT values as estimated from the dot-displays were not significantly affected by the trauma. However, the acoustic trauma significantly shifted the sigmoid functions used to fit the average data toward higher values of VOT or gap duration, and the VOT fit curve became less steep. Before the trauma, the mean 50% points of the fit curves were 35 and 38 ms for /bal-/pa/ and gap stimuli, respectively. After the trauma, the mean 50% points were increased to 46 and 51 ms for /bal-/pa/ and gap stimuli, respectively.

For the pre-trauma condition, our results are the same as those from the previous studies of Eggermont (1995a,b, 1999): MNGT and MNVOT presented similar mean values around 35–40 ms. Moreover, these results have been corroborated by another study where authors have recorded auditory evoked potentials elicited by stop consonant–vowel syllables directly from Heschl’s gyrus and temporal gyrus in awake humans (Steinschneider et al., 1999). This once more underlines the minor effect that anesthesia has on our findings. Furthermore, the recordings of the gap and /bal-/pa/ stimuli were at most 4 h apart, the earlier recording typically occurring at least 5 h after the pentobarbital injection. At this time one does not expect any change resulting from lingering barbiturate effects. We have previously reported about the lack of changes in temporal response properties as a function of time after anesthesia onset (Eggermont, 1991), and we have extensive recordings of VOT studies in animals (Tomita et al., unpublished) with a permanent noise trauma where we did not observe any changes that could be attributed to time after the single pentobarbital injection at the start of the experiment. Thus we conclude that the changes that we observe are due to the intervening noise trauma.

Syllables with a relatively long VOT – above 30 ms, were associated with a clear “double-onset” response, namely a neural response time-locked to the consonant release and one to voicing onset. On the other hand, syllables with a VOT below 30 ms generally elicited a reduced response – if any – to voicing onset. The latter results suggest “that the perceptual category boundary could reflect an average of the temporal activity patterns in auditory cortex (AI) that showed a double-on response” (Eggermont and Ponton, 2002). In addition, the fact that noise stimuli and different syllables are associated with similar MNGT or MNVOT around 35 ms in cortex suggests that these values reflect intrinsic cortical neural properties and are not directly related to perceptual boundaries. These points are developed below.

4.1. Neurophysiological mechanisms

Eggermont (1999, 2000) argued that cortical post-activation suppression might be the main neurophysiological mechanism accounting for the MNVOT and MNGT. Post-activation suppression may result from after hyperpolarization as well as from feed-back or feed-forward inhibition (Eggermont, 2000). As mentioned by Eggermont (1999), inhibition resulting from the involvement of GABA A receptor activation is difficult to distinguish from after hyperpolarization. They have approximately the same time constants and both start a few ms after the excitatory on-response of the neurons. Regardless the mechanism, post-activation suppression limits the temporal resolution of the auditory system in preventing the occurrence of locked responses to the trailing or voicing stimulus when gap duration or VOT are short. As modeled by Eggermont (2000), the post-activation suppression is expected to be prominent immediately after the onset of the leading stimulus or consonant release. This hypothesis is consistent with electrophysiological data showing that MNGT decreases as a function of leading stimulus duration (Eggermont, 2000).

Post-activation suppression, in constraining temporal neural responses, may then be responsible for the behavioral performances, namely for the categorical perceptual boundary of a speech continuum, for instance of /bal-/pa/ phonemes. In this context, it is noted that the perceptual boundary – around 25 ms (Kuhl and Miller, 1978) – is about 10 ms shorter the middle point of our normal data fitted with a sigmoid function (Fig. 8). This difference could be due to the anesthesia used as it increases inhibition and after hyperpolarization effects. Moreover, gap thresholds measured by Phillips et al. (1997) corroborate this. Namely, in a condition where the silent gap was marked by a wide-band noise (leading stimulus, intended to simulate a consonantal burst) and a low frequency narrow-band noise (trailing stimulus, intended to simulate the vowel), gap thresholds were found to match MNGTs found in Eggermont’s (1999, 2000) and in the present study. In addition, perceptual gap thresholds and MNGT present similar values as a function of leading burst duration (Eggermont, 2000).

However, the broad distribution of minimum gap and VOT thresholds also suggests that the perceptual decision is likely made on basis of the population firing activity as reflected in Fig. 8 and not on the distribution of the threshold values. In this respect, the activity in AI likely only provides a basis on which to set perceptual boundaries. If the activity produced by the consonant–
vowel phonemes in AI is affected so will be ability to discriminate between them.

4.2. Neural changes induced by the acoustic trauma

We have shown that the acoustic trauma impairs temporal resolution in AI: the sigmoidal fitting function shifted toward higher values of gap duration or VOT. Specifically, the middle points derived from these functions increased by about 10 ms (Fig. 8). Because the acoustic trauma induced a hearing loss (Figs. 1 and 3), it is important to separate the potential effect of the effective intensity level of stimulation on the neural responses from that of changes in temporal acuity. A decrease in the effective intensity level may have induced the impairment in temporal resolution we observed. For instance, Fig. 6 shows that the increase in MNVOT and MNGT is slightly dependent on the threshold shift. However, the unaffected FRmax values for large VOTs suggest that the changes in temporal resolution of cortical neurons after the acoustic trauma are not related to the effective intensity level. Fig. 8 shows that the FRmax (for gap durations or VOTs above 55 ms) is not decreased after the trauma; to the contrary, the largest FRmax values are even increased. This result suggests that, on average, neurons have roughly conserved the same responsiveness to the trailing stimulus. Eggermont (1995b, 1999) showed that MNGT or MNVOT could depend on the strength of the onset response related to the presentation of the leading stimulus or the consonant release. A decrease in the onset response related to the onset of the leading stimulus was associated with a decrease in MNGT and MNVOT. A decrease in effective level due to hearing loss may have decreased the neural responsiveness to the leading stimulus and consequently decreased the MNGT and MNVOT. This explanation is not corroborated by our results: the averaged onset responses related to the leading stimulus are not changed after the acoustic trauma compared to before (Fig. 7). This is likely due to the elevated sound level (65 or 75 dB SPL) used in these experiments.

In our previous paper (Noreña et al., 2003), we showed that neural activity was more strongly suppressed following the presentation of the tone burst for the post-noise trauma condition than for the pre-trauma condition. This suggests that the acoustic trauma may result in an increase of post-activation suppression, and hence longer MNVOT and MNGT values.

4.3. Potential perceptual correlates of the central changes that follow an acoustic trauma

Glasberg et al. (1987) showed that hearing loss was associated with a slower rate of recovery from forward masking. The authors suggested that larger gap thresholds in hearing-impaired subjects might be associated with this. This result is consistent with the hypothesis, suggested by our results (described above), that hearing loss increases the strength of the post-activation suppression. Namely, in preventing the occurrence of the second “on-response” related to the presentation of the trailing stimulus or voicing, the increased post-activation suppression could alter the temporal resolution.

We showed that the acoustic trauma changed the slope of the fitting function (Fig. 8). Indeed, for a /ba/-/pa/ continuum, the slopes were shallower in the post-trauma condition. Then, if the presence of the VOT in the /ba/-/pa/ continuum is actually encoded through a population average of the FRmax related to the presentation of the trailing stimulus, the ambiguity between /ba/ and /pa/ should consequently be increased. It is important to note that the slope derived from the psychometric function (Kuhl and Miller, 1978; Strouse et al., 1998) defines the accuracy of the categorical perception or, in other words, the ambiguity between the two extremes. It is indispensable for accurate coding of a signal-like speech to minimize these ambiguities. That may be the reason why the slope in the psychometric function is relatively steep. The decreased slopes of the fitting function of our data in the post-trauma condition may be a neural correlate of an impaired psychoacoustic temporal resolution. In addition, in low signal-to-noise ratio conditions, this decrease in temporal resolution may even aggravate the perceptual performances.

This study shows for the first time that a hearing loss induced by an acute acoustic trauma decreases the temporal resolution of cortical neurons, both in the region of the hearing loss and up to 1.5 octave below it. We suggest that this impairment in the ability to code VOT might be related to an increase in the post-activation suppression.

Acknowledgements

This work was supported by the Alberta Heritage Foundation for Medical Research, the National Sciences and Engineering Research Council, the Canadian Language and Literacy Research Network, and the Campbell McLaurin Chair for Hearing Deficiencies.

References


