Neural Responses in Primary Auditory Cortex Mimic Psychophysical, Across-Frequency-Channel, Gap-Detection Thresholds

JOS J. EGGERMONT
Department of Physiology and Biophysics and Department of Psychology, University of Calgary, Calgary, Alberta T2N 1N4, Canada

Received 7 March 2000; accepted in final form 1 June 2000

Eggermont, Jos J. Neural responses in primary auditory cortex mimic psychophysical, across-frequency-channel, gap-detection thresholds. J Neurophysiol 84: 1453–1463, 2000. Responses of single- and multi-units in primary auditory cortex were recorded for gap-in-noise stimuli for different durations of the leading noise burst. Both firing rate and inter-spike interval representations were evaluated. The minimum detectable gap decreased in exponential fashion with the duration of the leading burst to reach an asymptote for durations of 100 ms. Despite the fact that leading and trailing noise bursts had the same frequency content, the dependence on leading burst duration was correlated with psychophysical estimates of across frequency channel (different frequency content of leading and trailing burst) gap thresholds in humans. The duration of the leading burst plus that of the gap was represented in the all-order inter-spike interval histograms for cortical neurons. The recovery functions for cortical neurons could be modeled on basis of fast synaptic depression and after-hyperpolarization produced by the onset response to the leading noise burst. This suggests that the minimum gap representation in the firing pattern of neurons in primary auditory cortex, and minimum gap detection in behavioral tasks is largely determined by properties intrinsic to those, or potentially subcortical, cells.

INTRODUCTION

Humans can detect temporal gaps as short as 2–3 ms in duration regardless of the position of the gap in a monaurally presented noise, i.e., independent of the duration of the leading or trailing noise burst (Formby et al. 1998; Penner 1977; Phillips et al. 1997; Plomp 1964). In case the frequency content of the leading and trailing burst differs sufficiently so that they activate different frequency channels, the minimum detectable gap becomes dependent on the duration of the leading burst and can reach values of ~40 ms (Phillips et al. 1997).

Gaps in sound play an important role in speech. For instance, the perceptual distinction between the two phonemes /ba/ and /pa/ is largely based on the voice onset time (VOT), the length of the silent interval between the noise burst and the following vowel. The perceptual boundary in this case is usually found for VOT = 35 ms in humans as well as in rhesus monkeys (Morse and Snowdon 1975) and chinchillas (Kuhl and Miller 1978). The same threshold was found in cells of AI stimulated at 45–65 dB SPL with a /ba/-/pa/ continuum in which the VOT was changed from 0–70 ms in 5-ms steps (Eggermont 1995b, 1999a). At more peripheral levels of the auditory pathway, a more or less continuous representation of VOT in the firing rate of auditory nerve fibers (Sinex and McDonald 1988) or cells in the inferior colliculus (Chen et al. 1996) tuned to the dominant frequency in the vowel was found.

The finding of similar thresholds for VOT in a /ba/-/pa/ continuum and for a gap in noise with the same short leading burst duration suggests that the difference in frequency content of the bursts does not determine the gap threshold for the cortical cells. It is more likely that the threshold is determined by the cell properties. If that is the case, how can one explain the differences in gap detection threshold for short leading burst durations obtained perceptually for a within-channel condition and various across-channel conditions? Examples for the two conditions have already been mentioned: detecting a gap in noise and making a perceptual distinction between /ba/ and /pa/on the basis of VOT. Yet these are two different tasks. Phillips et al. (1997) have converted this phoneme categorization task into a gap-detection task by using different frequency content noise bursts leading and trailing the gap. When the frequency content of the leading and trailing noise bursts was the same, the minimum detectable gap was ~5 ms and independent of the duration of the leading burst [a fact well documented in the literature (Forrest and Green 1987), but see Snell and Hu (1999)]. However, when the frequency content of the leading and trailing noise bursts did not overlap, the minimum detectable gap became dependent on the leading burst duration. For a 5-ms burst, the minimum gap was close to 40 ms (similar to the categorical perception boundary), whereas for a 100 ms burst, the minimum detectable gap was ~10 ms with a gradual change in between (Phillips et al. 1997). Similar results were obtained with leading and trailing bursts of the same frequency content presented to different ears or from different spatial locations (Phillips et al. 1998). This across-perceptual-channel behavior is remarkably similar to that suggested by the response of single cortical cells for gaps-in-noise inserted either after 5 ms or after 500 ms following noise onset (Eggermont 1995a, 1999a).

The effect of the leading burst on the probability of a...
response to the trailing burst could be due to a forward masking effect (Eggermont 1995b) and then is expected to depend on the duration of the leading burst and the recovery time since the end of the burst. In the auditory periphery, the onset response to the second burst decreases slightly for increasing duration of the masker for a constant 17-ms gap (Smith 1977). This is not what is observed in auditory cortex since the gap detection threshold decreases and the response, for a constant gap duration, increases with increasing leading burst duration. Cortical cell properties that depend dominantly on the time since the last on response rather than the time since the end of the stimulus are after-hyperpolarization (Schwindt et al. 1988) and potentially lateral inhibitory activity (Brosch and Schreiner 1997; Calford and Semple 1995).

What are the neural substrates of these two types of detection, within-frequency channel and across-frequency channels, that may be pertinent to gap-detection experiments? A neural frequency channel, in its simplest case, consists of a restricted hierarchical set of neurons all receiving input from a single inner hair cell in the cochlea. The single inner hair cell in cats is innervated by ~10–20 auditory nerve fibers, and, depending on stimulus level, a certain fraction thereof will be activated. At the primary auditory cortex, a strict cochleotopic organization is found so that a single hair cell is likely to affect a single iso-frequency sheet of cortex; the extent of the activation is again dependent on intensity level (Phillips et al. 1994). Neurons within this iso-frequency sheet exhibit a variety of properties. The monaural properties include differential sensitivity to stimulus level, different width of the frequency-tuning curve, different first-spike latency, and different percentages of monotonie or nonmonotonie input-output functions (Schreiner 1995). A single-frequency channel could thus be defined as this divergent/convergent spread of auditory activity from one point on the cochlear partition to activity distributed along, at least in cat, a ventrodorsal sheet transecting the auditory cortex. This definition is congruent with that for a neural representation given by Johnson (1980) and Phillips et al. (1997). Within-channel gap detection has been equated functionally (Phillips et al. 1997) with detecting a discontinuity in the neural activity pattern along its path from cochlea to cortex. This short discontinuity has been demonstrated in the activity of auditory nerve fibers (Zhang et al. 1990), in single cells of the inferior colliculus (Barsz et al. 1998; Walton et al. 1997, 1998), in cells of field LA of the starling (Buchfellner et al. 1989), and in cells in three divisions of auditory cortex (Eggermont 1999a). Surprisingly, the results in auditory cortex were strongly dependent on the duration of the leading burst. For leading burst durations of 500 ms, the minimum gap, a discontinuity in the activity pattern, was 5 ms (the shortest value tested). However, for a leading burst duration of 5 ms, the gap was only represented in an ON response to the burst following the gap (the trailing burst) for gap durations exceeding 35 ms (Eggermont 1995a). The results in auditory cortex were obtained with gaps in noise and thus combined information across nearly all frequency channels. The single cortical cell recorded from is activated by frequencies in the cross section of its frequency response area with all activated frequency channels. This includes all the nonlinear interactions between different simultaneously present frequencies that take place in subcortical nuclei. However, one can maintain that a single cortical cell is only influenced by a single frequency channel, i.e., that determined by the cell’s frequency tuning curve. As a consequence the cortical cell performed a within channel detection task.

Let us assume the simplest case of across-frequency-channel gap detection in which the leading burst and trailing bursts are pure tones with sufficient frequency difference as not to activate, or suppress, the same region on the basilar membrane. Up to the level of the auditory midbrain, with the notable exception of the octopus cells in cochlear nucleus, one expects the activation of neurons by the two bursts to be mutually exclusive, i.e., the response to the trailing burst is the same for any gap (even for 0 gap). In the broadly tuned neurons of the external nucleus of the inferior colliculus (IC) and in some divisions of the medial geniculate body (MGB), these streams may converge onto a single cell. If that single cell responds to the activity induced by both the leading and trailing bursts, one expects this to represent a within-channel detection task. In case the response properties of the cell in secondary auditory cortex (AII) activated by such broadly tuned neurons impose a large minimum detectable time interval, the minimum gap will become dependent on the duration of the leading burst. Of course, because of the divergent/convergent character of feed forward activation of cortical cells, frequency differences larger than those producing interaction on the basilar membrane may still affect activity patterns of cortical cells narrowly tuned for frequency (Schulze and Langner 1999).

In a comparative study of cortical cells in AI, AII, and the anterior auditory field (AAF), no obvious area difference was found in a within-channel design for the minimum detectable gaps positioned after a 5-ms noise burst or after a 500-ms burst (Eggermont 1999a). Thus a study limited to AI is expected to produce representative results for these three cortical areas. The dependence on leading burst duration suggested that the cortical cells in all three areas behaved as in a perceptual across-channel detection task. The exact dependence on leading burst duration should in that case be equal to that observed in behavioral gap-detection tasks. Because the response patterns of cortical cells are largely confined to onset responses to the leading and trailing burst (see also Steinschneider et al. 1994), the dependence on leading burst duration should in effect be a dependence on the summed duration of the leading burst and the minimum gap.

The present study is aimed at extending the study on the dependence of minimum detectable gaps on leading burst duration for cells in AI by using leading burst durations of 5, 20, 50, 200, and 500 ms and to describe and model the recovery properties of the cortical cells. By modeling the response of cortical cells on basis of recovery from synaptic depression and from after-hyperpolarization, the relative importance of these mechanisms can be assessed. Furthermore the more detailed data compared with our previous study (Eggermont 1999a) will allow better comparison with behavioral gap-detection studies.

Meth O d s
The care and the use of animals reported on in this study was approved (No. P88095) by the Life and Environmental Sciences Animal Care Committee of the University of Calgary.

Animal preparation
Cats initially received 0.25 ml/kg body wt of a mixture of 0.1 ml acepromazine (0.25 mg/ml) and 0.9 ml of atropine methyl nitrate (5
mg/ml) subcutaneously. After about one-half hour, they received an intramuscular injection of 25 mg/kg of ketamine (100 mg/ml) and 20 mg/kg of pentobarbital sodium (65 mg/ml). Lidocain (20 mg/ml) was injected subcutaneously and rubbed in gently, then a skin flap was removed and the skull cleared from overlying muscle tissue. A large screw was cemented upside down on the skull with dental acrylic. An 8-mm-diam hole was trephined over the right temporal cortex so as to expose parts of AI and AII. A 4-mm hole was drilled over the AAF. The dura was left intact, and the brain was covered with light mineral oil. Then the cat was placed in a sound-treated room on a vibration isolation frame and the head secured with the screw. Additional acepromazine/atropine mixture was administered every 2 h. Light anesthesia was maintained with intra-muscular injections of 2–5 mg · kg⁻¹ · h⁻¹ of ketamine. The wound margins were infused every 2 h with lidocain and also every 2 h new mineral oil was added if needed. The temperature of the cat was maintained at 37°C. At the end of the experiment the animals were killed with an overdose of pentobarbital sodium.

Acoustic stimulus presentation

Acoustic stimuli were presented in an anechoic room from a speaker (Fostex RM765) placed 55 cm away from the cat’s head and 45° from the midline into the contralateral field. The sound-treated room was made anechoic for frequencies 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 B and K (type 4134) microphone placed above the animal’s head and facing the loudspeaker. A search stimulus consisting of random-frequency tone pips, noise burst, and clicks was used to locate units. Characteristic frequency (CF) and tuning curve of the individual neurons were determined with 50-ms duration, gamma-shape envelope, tone pips presented randomly in frequency once per second (Eggermont 1996). 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. After the frequency tuning properties of the cells at each electrode were determined, gaps ranging from 5 to 70 ms in duration were placed in two positions in wideband noise bursts of 1 s in duration and presented once per 2 s in random order. The first position was variable, and the gap started 5, 20, 50, or 200 ms after the noise-burst onset. The second position, “the late gap,” was always positioned 500 ms after noise burst onset, thus the leading burst duration for that gap varied depending on the position of the first gap. The leading gap durations for the late gap were 300, 450, 480, and 495 ms. Each stimulus was presented 15 times. The noise bursts used consisted of “frozen” noise, i.e., the pseudo-random noise sequence was the same for all conditions.

Recording and spike separation procedure

Three tungsten micro electrodes (Micro Probe) with impedances between 1.5 and 2.5 MΩ were independently advanced perpendicular to the AI surface using remotely controlled motorized hydraulic microdrives (Trent-Wells Mark III). The electrode signals were amplified using extracellular preamplifiers (Dagan 2400) and filtered between 200 Hz (Kemo VBF8, high-pass, 24 dB/octave) and 3 kHz (6 dB/octave, Dagan roll-off) to remove local field potentials. The signals were sampled through 12-bit A/D converters (Data Translation, DT 2752) into a PDP 11/33 microcomputer, together with timing signals from three Schmitt triggers. In general, the recorded signal on each electrode contained activity of two to four neural units. The PDP was programmed to separate these multi-unit spike trains into single-unit spike trains using a maximum variance algorithm (Eggermont 1996). The spikes from well-separated waveform classes, each assumed to represent a particular neuron, were stored and coded for display. The multi-unit 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 (Eggermont 1996).

The boundary between AI and AAF was explored by taking a series of multi-unit measures from caudal to rostral and assuring that there was a gradual increase in CF, which reversed in direction when advancing to the AAF. The secondary auditory cortex was identified by its location and by the broader tuning curves and different response patterns (latency and bursting activity) compared with those in the central and ventral parts of AI. Because all CFs were all 2 kHz, confusion with the posterior field is unlikely. The three recording-electrode positions in AI were aimed either to be within an iso-frequency sheet or perpendicular to it. Recordings were made between 600 and 1,200 μm below the cortex surface.

Data analysis

The peak of the poststimulus-time histogram (PSTH) in the first 100 ms after each tone-pip onset was estimated for each intensity value used. The peak values for three adjacent frequencies were combined to reduce variability and divided by number of stimuli and presented as a firing rate per stimulus. This resulted in 27 frequencies covering five octaves so that the final resolution was ~0.2 octaves. The results per stimulus intensity were combined into a peak rate-frequency-intensity profile from which tuning curves, rate-intensity functions, and iso-intensity-rate contours could be derived (Eggermont 1996). The frequency-tuning curve was defined for a firing rate at 25% of the maximum firing rate. The threshold at the characteristic frequency (CF) was determined as 2.5 dB below the lowest intensity that produced visible time locked responses to the tone pip, i.e., midway between the stimulus that produced a response and the one that did not.

PSTHs with 2- or 5-ms binwidth were calculated for each of the two “stimuli” in each type of stimulation (e.g., leading and trailing noise burst). From these PSTHs both the peak latency and the spike count were calculated. The minimum value of the gap, for which a “double on response” was obtained, was called the minimum gap. The minimum gap detectable from inspection of the dot displays depends on the number of stimulus presentations and the firing rate of the neurons. We presented 15 repetitions for each gap duration, Δt, and required at least three time-locked spikes per 15 presentations to the trailing burst to detect a gap. Autocorrelograms for lags between 10 and 160 ms were constructed for all single units. Population autocorrelograms were constructed by adding all the individual single-unit ones as described previously (Eggermont 1998).

Forward masking model

The model is based on a combination of fast recovery from synaptic depression and recovery from after-hyperpolarization (AHP). Parts of the model for synaptic depression have been described previously (Eggermont 1985, 1999b) and will be briefly reviewed here.

The model assumes exhaustion of transmitter supply in the immediate release store at the presynaptic terminal at a rate λ, and replenishment at a rate μ. The steady-state fraction of depleted transmitter is λ/(λ + μ) and the steady-state fraction available for release is R₀ = 1 − λ/(λ + μ) = μ/(λ + μ). This steady state is reached with a time constant τᵣₑᵥₑ = (λ + μ)⁻¹. It is assumed that the induced excitatory postsynaptic potential (EPSP) is proportional to the synchronous amount of transmitter released.

Consider next a forward masking experiment with a masking burst long enough to allow the synapse to reach the steady-state release level R₀. The masker is followed after a silent interval, Δt, by a test burst with equal intensity and frequency content as the masker. The transmitter available for release by the test noise burst onset, F(Δt), increases with Δt, with time constant τₑᵥₑ = μ⁻¹.
\[ F(\Delta t) = R_{\text{ss}}(1 - R_{\text{ss}}) \{1 - \exp(-\Delta t/\tau_{\text{recov}})\} \]

(1)

In the model, the duration of the spike-burst was varied between 1 and 20 and 100 ms.

Medium-duration AHP has been demonstrated to follow a burst of spikes in thalamic or cortical cells (Schwindt et al. 1988) and recovers during the course of the leading noise bursts and the gap following it. In the model, the duration of the spike-burst was varied between 1 and 10 ms but the effects were small, so a value of 5 ms is used throughout. The amount of AHP decreases exponentially from a level \( -H \) as

\[ AHP = -H \exp(-\Delta t + D - 5)/\tau_{\text{AHP}} \]

(5)

In this equation \( \tau_{\text{AHP}} \) is the recovery time constant. The time course of the AHP (for \( H = 1 \)) is illustrated in Fig. 1, top, and indicated by AHP.

Thus the membrane potential reflecting the recovery from the synaptic depression and the AHP is given by

\[ V(\Delta t, D) = F(\Delta t, D) + AHP(\Delta t, D) \]

(6)

This combination of forward masking due to synaptic depression and AHP produced recovery functions that were increasing much slower as a function of \( \Delta t \) than the cortical data. This is illustrated by the \( (F + AHP) \) curves in Fig. 1, top. A much better fit was obtained by assuming a sigmoidal recovery of firing rate based on the membrane potential \( V(\Delta t, D) \) and with adjustable slope of the curve (Koch 1999)

\[ A(\Delta t, D) = 1/[1 + \exp(-\alpha(V(\Delta t, D) - T))] \]

(7)

Here \( \alpha \) determines the slope of the recovery function and \( T \) is the value of \( V \) at the 50% recovery point. By assuming that the recovery of the firing rate starts from the level of (spontaneous) activity at the end of the masker, \( \psi \), the modeling function reads:

\[ r_m(\Delta t, D) = (1 - \psi)A(\Delta t, D) + \psi \]

(8)

where \( r_m(\Delta t, D) \) is the peak firing rate to the probe. These sigmoidal recovery functions, starting at the end of the leading burst, are illustrated in Fig. 1, bottom. The curves marked by \( F \) indicate recovery from depression only, the curves marked \( F + AHP \) indicate recovery from depression and AHP combined. One observes that for leading bursts of 100 ms the difference is small indicating that most of the recovery in this case is determined by the depression.

Spiking train analysis was done with custom designed software (Eggermont 1996) and with the Stranger software package. All statistical analyses were performed using Statview5. Graphics, modeling, and systems analysis were done with Matlab, Powerpoint, and Horizon software.

RESULTS

Recordings were made in 13 cats from three sites in AI simultaneously for a total of 51 recording sites. For each site, four different gap-in-noise combinations were used with leading burst durations of 5, 20, 50, and 200 ms and a late gap inserted 500 ms after the onset of the leading burst. The noise intensity was 65 dB SPL (RMS) for all conditions.

Individual example

Figure 2 shows the multi-unit responses for one particular recording site with a CF = 5 kHz and a recording depth of 880 μm below the cortex surface. Results for the three shortest leading burst duration conditions are shown in Fig. 2, left, and for the “late-gap” inserted 500 ms after leading burst onset in Fig. 2, right. The leading burst durations for the late gap conditions (after a varying leading burst and early gap duration) were long enough (380–500 ms) to give nearly identical recovery results. The recording consisted of two single units, with similar response properties. The vertical axes represent the 15 gap durations, ranging from 0 to 70 ms. For each gap duration the stimulus was repeated 15 times, and as a consequence one observes the staircase pattern of the response after the gap. The horizontal axis represents the time (0–150 ms) after leading burst onset (left) or after leading burst onset plus 500 ms (right). The late-gap condition shows onset responses to the trailing noise burst for all gap values different from zero. For the early gap, the minimum value for which a second ON response is present depends on the duration of the leading noise burst. The estimated minimum early gap values are from top to bottom row: 55, 40, and 10 ms. After the onset responses to the leading and trailing bursts, spontaneous activity was almost completely suppressed and rebounded after ~125 ms as shown for the 5- and 20-ms leading burst duration. Note that the ON response to the trailing bursts also suppresses the rebound activity. After the end of the long trailing burst (right), the spontaneous activity during the gap is largely preserved. The PSTHs for the three early gap conditions, calculated for a binwidth of 2 ms, are shown in Fig. 3. The change in the
response to the trailing burst with increasing gap duration is quite abrupt. The response to the trailing burst very quickly reaches the value for that to the leading burst. The difference in onset latencies for the responses to the leading and trailing noise bursts is, depending on the duration of the leading noise burst, only 2–4 ms longer than the sum of leading burst plus gap duration.

To assess whether the duration of the leading burst plus the duration of the gap, i.e., the time interval between the onset responses to the leading and trailing bursts, was represented in the interspike intervals of the individual single units, autocorrelation histograms were calculated. Figure 4 shows the sum of the two single-unit autocorrelation histograms using the same analysis window length (but omitting the very short intervals that are found in the burst like response). Peaks indicate the number of firings at that time interval that followed a previous spike in the same neuron, conditional of an onset response to the leading burst. The dependence of the major histogram peaks on the gap duration parallels that of the PSTH and suggests that a large fraction of the spikes in the second ON response is the result of repeated firing with an interval approximately equal to the leading burst duration plus the gap. The detection threshold for a minimum gap on basis of interspike interval values in this example is the same as that estimated from the PSTH or dot display. For Fig. 4, left, leading burst duration of 5 ms, one observes for gap durations

FIG. 2. Multi-unit dot displays for gap in noise stimuli. Two gaps are inserted in the 1-s noise burst: 1 early in the noise burst (left) and one 500 ms after noise burst onset (right). The duration of the burst preceding the early gap was either 5 ms (top left), 20 ms (middle left), or 50 ms (bottom left). The multi-unit record consisted of 2 well-separated single units with identical properties. Along the horizontal axis, time since noise burst onset is presented. Left: time runs from 0 to 150 ms; right: time runs from 500 to 650 ms. The vertical axis represents gap duration. This was advanced in 5-ms steps, which is reflected in the step-wise response change to the trailing burst. Every gap condition was presented 15 times. One observes (left) that the time between the onset response to the leading burst and that for the lowest gap to the trailing burst is approximately constant. When no onset response to the trailing burst is present a rebound response occurs ~130 ms after the onset response.

FIG. 3. Post stimulus time histograms (PSTHs) for the data presented in Fig. 2, left. The organization is similar as in the previous figure, the time window since noise burst onset is from 0 to 150 ms, and vertically histograms are shown for the 15 different gap durations. Bin width was 2 ms.

FIG. 4. Autocorrelograms for the same data as in Fig. 3. Lag times between 10 and 160 ms are shown on the horizontal axis. Binwidth is 2 ms. One observes that the preferred intervals are showing the same pattern as the response to the trailing burst in the PSTHs. In the absence of intervals that follow the duration of the gap, one notices a preference for firing at intervals ~130 ms. These are the result of rebound activity after postactivation suppression, likely as a result of after-hyperpolarization (AHP).
<55 ms that there is an increased firing probability with intervals that are ~130–140 ms. This increased firing probability after 130 ms is likely the result of a recovery from the AHP.

**Group data**

All individual single-unit autocorrelograms were added to obtain population autocorrelograms (binwidth = 5 ms) for the different leading burst durations. Contour line plots (Fig. 5) are used to represent the summed autocorrelation histograms. Results are shown for lag times of 10–160 ms for leading burst duration of 5, 20, and 50 ms and for lag times of 10–300 ms for the 200-ms leading burst length. Contour lines were set at 50% (dark shaded area) and 25% of maximum response. One can detect clear deviations from the background activity down to gap durations of 45 ms for leading burst length of 5 ms. The minimum gap represented in the autocorrelogram is 30 ms for a 20-ms leading burst duration, 10 ms for a 50-ms leading burst, and equal to the shortest gap duration of 5 ms for the 200-ms leading burst. In the latter condition, one observes considerable rebound activity between 80 and 150 ms after the initial spikes to the leading burst.

The minimum gap durations for which an on response to the trailing burst could be detected (minimally 3 time locked spikes per 15 presentations) from the dot displays for individual multi-unit recordings are shown in Fig. 6 as a function of leading burst duration. The late-gap responses (gap inserted 500 ms after the leading burst onset) have a leading noise burst duration (from 300 to 495 ms) depending on the duration of the initial leading burst and the value of the gap. An extended range of minimum gap values is found for the shortest leading bursts e.g., for a 5-ms burst the range is from 10 to 65 ms. The

The minimum detectable gaps based on the population autocorrelograms are indicated by ▫. These values are representative of the group of neurons with the highest onset firing rates because only those allowed calculation of reliable autocorrelograms. Minimum detectable gap values obtained in psychoacoustical experiments by Phillips et al. (1997), using a wideband noise leading burst and a narrowband noise trailing burst, with a varying duration of the leading noise burst are indicated by ×. These values are, as expected, consistently below the mean values for the neuron data for the shorter leading burst durations, but converge to a similar asymptotic value. They are closer to the population autocorrellogram values for leading burst durations <100 ms, but longer than those values for leading bursts of greater length.

The minimum detectable gap’s dependence on leading burst duration in the example (Figs. 2–4), suggested that the onset to the trailing burst occurred at a fixed time since the leading burst onset. This is evaluated in Fig. 7, which shows that the sum of the leading burst duration and the minimum gap is nearly constant for leading bursts ≤50 ms in duration. For the neuron data the sum is ~55 ms, and for the psychoacoustical data (Phillips et al. 1997) the sum is ~40 ms. The model prediction will be discussed in the appropriate section.

**Recovery functions**

Average recovery functions were calculated from PSTHs with 5-ms binwidth for those conditions in which the average peak onset response was either larger [35 multi-unit (MU) recordings] than or smaller (16 MU recordings) than 10 spikes in a 5 ms bin. Because all conditions were repeated 15 times this translates into a peak value of >0.75 or <0.75 spike.
bin$^{-1}$ · burst$^{-1}$. The response to the trailing bursts was normalized to the average response at a gap of 70 ms. The results for the high response group ($\geq 0.75$ spike · bin$^{-1}$ · burst$^{-1}$) are shown in Fig. 8 for four leading burst durations (the superimposed model curves will be discussed in the next section). Mean values are shown and a considerable scatter around the full recovery value is observed for leading burst durations of 50 and 200 ms and gap durations 40 ms. Note that the values for zero-duration gaps sample the ongoing activity at the end of the leading noise burst.

An ANOVA for the effect of leading burst duration was done per gap duration value. For the group of recordings with onset activity 0.75 spike · bin$^{-1}$ · burst$^{-1}$, a significant effect of leading burst duration ($P < 0.05$) was found for gap durations from 10 to 55 ms. In contrast, for the group with onset activity <0.75 spike · bin$^{-1}$ · burst$^{-1}$, no significant effect at the $P < 0.05$ level for leading burst duration was found. Paired comparisons with the response for gap duration zero, i.e., activity at the end of the leading burst, were made for each leading burst duration to obtain another estimate of the average minimum gap duration. Table 1 shows minimum gap durations for which the response to the trailing burst becomes larger ($P < 0.05$) than for zero gap duration. The results suggest that neurons with firing rates >0.75 spikes · bin$^{-1}$ · burst$^{-1}$ have, on average, shorter minimum gap durations for long leading bursts. In addition, for the shortest leading bursts larger minimum gap durations are found for the less responsive neurons. These results suggest that the behavioral responses (as in Phillips et al. 1997) are likely made on basis of the most responsive neurons.

**Modeling**

The parameter estimation for the recovery from depression and AHP model was done for those 35 recordings with an average onset response of $\geq 0.75$ spikes · bin$^{-1}$ · burst$^{-1}$. The parameters to be estimated in Eq. 8 are $r_{\text{adap}}$, $\tau_{\text{recov}}$, $\tau_{\text{AHP}}$, $T$, and $\alpha$. $H$ was set equal to 1. The data for the 500-ms leading bursts (Fig. 8) were used to estimate the parameters for the synaptic depression mechanism (because the recovery from AHP is likely complete at the end of these long bursts), resulting in $r_{\text{adap}} = 12$ ms and $\tau_{\text{recov}} = 20$ ms. For the shorter leading bursts, a least-mean-square curve fit on the data shown in Fig. 8, with an initial choice of $\tau_{\text{AHP}} = 50$ ms (the literature value), resulted in the remaining three parameter estimates: $\tau_{\text{AHP}} = 55$ ms, $T = 0.61$ and $\alpha = 19.6$. The estimate for $\tau_{\text{AHP}}$ is slightly larger than the value of 50 ms estimated for layer IV pyramidal cells from sensorimotor cortex in vitro (Schwindt et al. 1988). The estimate for $T$ suggests that when the membrane potential is recovered to 61% of its (normalized) maximum depolarization value, the peak firing probability becomes 50%. For spontaneous activity $\psi$, the value obtained for the zero gap duration was used. The resulting curve fits are shown in Fig. 8. Acceptable results were obtained for leading burst durations of 5, 50, 200 (not shown), and 500 ms, but the curve for 20-ms burst duration was $-5$ ms displaced to the shorter gap values.

This model (the only difference being the estimate of the $\psi$ values) was then tested on the individual multi-unit recording data (Figs. 2–4), and the results are shown in Fig. 9. Acceptable curve fits were obtained for leading burst durations of 20, 50, and 500 ms using the same parameters as estimated for the mean values of the entire high responsive group. For the 5-ms

![Image](http://jn.physiology.org/Downloadedfrom/10.220.33.2/0nSeptember27,2016)

**FIG. 7.** Leading burst plus minimum detectable gap duration as a function of leading burst duration shows an approximate constancy for short burst durations that suggests a critical minimum interval between the onset responses to the leading and trailing bursts. Individual data and mean values from psychoacoustic experiments are indicated as well as threshold values from the population autocorrelogram, and the prediction from the model.

**FIG. 8.** Mean recovery functions for units with $>0.75$ spikes · bin$^{-1}$ · burst$^{-1}$ for 4 lengths of the leading burst. The responses are normalized to the level at full recovery. Curve fits from the model based on Eq. 8 and combining recovery from depression and after-hyperpolarization are drawn in.

### Table 1. Minimum gap duration for which the response was larger ($P < 0.05$) than for gap duration zero

<table>
<thead>
<tr>
<th>Leading Burst Duration, ms</th>
<th>Onset Response $&gt; 0.75$ spikes/burst</th>
<th>Onset Response $&lt; 0.75$ spikes/burst</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>50</td>
<td>40</td>
</tr>
<tr>
<td>20</td>
<td>40</td>
<td>40</td>
</tr>
<tr>
<td>50</td>
<td>25</td>
<td>35</td>
</tr>
<tr>
<td>200</td>
<td>5</td>
<td>25</td>
</tr>
<tr>
<td>500</td>
<td>5</td>
<td>10</td>
</tr>
</tbody>
</table>
suggest that this dependence on leading burst duration is a

gap detection but no causal relationship to it. Modeling results
have the same time course as the behavioral between-channel

categorical perception boundaries, or it may be accidental, i.e.,

ilarity may underlie a basic mechanism for establishing cate-

gorical perception. This simultaneously high firing rates was comparable to that for human
gap on leading burst duration in cat cortical cells with rela-
tionship that could be based on the set of very responsive units.

In an across-perceptual-channel task, the cortical (or poten-
tially thalamic) neurons are located central from the conver-
gence point of the two perceptual channels implied in the behav-
ioral task. The neuron populations in these perceptual
channels process the leading and trailing burst activity inde-
pendently (by definition). Thus the neurons in the two sets of
frequency channels will respond with an on response to either
the leading burst or the trailing burst regardless of the duration
of the gap. Only through the convergence of activity of neurons
from these two populations onto neurons with properties such as the cortical neurons in the present study can the observed
dependence of gap threshold on leading burst duration emerge.
This could be neurons with broad frequency tuning curves
located in the dorsal part of AI or in secondary auditory cortex.
We showed previously (Eggermont 1999a) that neurons in AI,
AAF, and AII share the same gap detection properties and also
show similar VOT response properties (provided the CFs al-
low a response to the vowel part of the phoneme).

One could assume that in the absence of an AHP \((H = 0)\),
short minimum gap values would be found for all leading burst
durations. In fact, model results for the minimum gap (set at
25% recovery) under this condition show a slight increase with
duration. For a 20-ms leading burst, the minimum gap is 2 ms,
and it reaches a 5.7-ms value for a 50-ms burst and increases
only to 6 ms for the longest leading burst duration. For a 5-ms
gap, the recovery function is already 0.97 at the end of the
burst, but some gap will be required to be able to detect a
discontinuity. These predicted values are comparable to the
situation for a within-frequency-channel gap-detection task. In
case the AHP is present, the minimum detectable gap for 25%
recovery is indicated in Fig. 7. The model prediction matches
the cortical data quite well but is higher than obtained psycho-
physically. Reducing the recovery criterion to only 10% makes
the prediction to coincide with the population autocorrelogram
values, which are still above the psychophysical ones.

The group of neurons with onset response \(<0.75\) spike \cdot
bin\(^{-1}\) \cdot\ burst\(^{-1}\) showed shallower, and more variable, recovery
functions. These could be approximated with the model by
reducing the slope of the recovery function to at least half its
value used for the high response group and reducing the
strength of the AHP to \(\leq 0.5\).

**Discussion**

It was found that the dependence of the minimum detectable
gap on leading burst duration in cat cortical cells with rela-
tively high firing rates was comparable to that for human
behavioral, across-frequency-channel gap detection. This sim-
ilarity may underlie a basic mechanism for establishing cate-
gorical perception boundaries, or it may be accidental, i.e.,
have the same time course as the behavioral between-channel
gap detection but no causal relationship to it. Modeling results
suggest that this dependence on leading burst duration is a
consequence of an AHP induced by strong responses to the
leading noise burst. Modeling also suggests that in the absence
of AHP there is only a modest dependence on leading burst
duration and that minimum detectable gaps are predicted to be
between 2 and 6 ms. The findings suggest that behavioral gap
thresholds in humans, for leading and trailing noise bursts with
different frequency content, have a correlate in the population
gap threshold of a group of high firing rate cortical neurons for
leading and trailing noise bursts with identical frequency con-
tent. These cortical data have to be considered as resulting
from a within-frequency-channel task. The AHP responsible
for the dependence on leading burst duration is likely intrinsic
to cortical cells. A temporal constraint as found here for within
channel gap detection is required for a between-channel mech-
anism to account for the dependence on leading burst duration.
This between-channel mechanism likely resides in single cells
(as suggested by the present analysis) either as a result of
convergence of two channels or as a result of an operation on
a neural assembly that represents activity of the two channels.

In an across-perceptual-channel task, the cortical (or poten-
tially thalamic) neurons are located central from the conver-
gence point of the two perceptual channels implied in the behav-
ioral task. The neuron populations in these perceptual
channels process the leading and trailing burst activity inde-
pendently (by definition). Thus the neurons in the two sets of
frequency channels will respond with an on response to either
the leading burst or the trailing burst regardless of the duration
of the gap. Only through the convergence of activity of neurons
from these two populations onto neurons with properties such as the cortical neurons in the present study can the observed
dependence of gap threshold on leading burst duration emerge.
This could be neurons with broad frequency tuning curves
located in the dorsal part of AI or in secondary auditory cortex.
We showed previously (Eggermont 1999a) that neurons in AI,
AAF, and AII share the same gap detection properties and also
show similar VOT response properties (provided the CFs al-
low a response to the vowel part of the phoneme).

A group of neurons with low firing rates showed recovery
patterns that were not significantly dependent on leading burst
duration, although the trend was similar to that for more
responsive neurons: longer leading burst result in faster re-
covery and shorter minimum gaps. Such neurons with little or
no AHP may be able to detect the short (\(<5\) ms) within-
frequency-channel minimum gaps. This would require a “se-
lective listening” hypothesis. The between-channel tasks that
are occurring frequently in natural sounds (as e.g., in a /ba/-/pa/
continuum) could be based on the set of very responsive units.
In contrast, the detection of discontinuities that are in the 2- to
5-ms range in a within-channel task (a phenomenon that does
not occur frequently in natural sounds) would require the
recruitment of less responsive units that show little or no AHP.

On the basis of the population autocorrelograms for neurons
with high firing rates, one may conclude that the minimum
detectable gaps are represented in the interspike intervals of
individual cortical neurons and neuronal populations. This
coding allows time interval information to be transmitted to
other cells, for instance in speech areas, that use this informa-
tion to construct, e.g., categorical perception boundaries. A
potential problem with this is that the time interval transmitted
is that between the onsets to the leading and trailing noise
bursts, and not the duration of the gap. However, for speech the

![Figure 9](http://jn.physiology.org/figures/fig_9.png)
leading noise burst duration in consonant-vowels such as /pa/ is <10 ms, so the transmitted interval is at most 10 ms longer than the VOT.

**AHP or another mechanism?**

The behavioral and neural data suggest a critical minimum interval between the onsets of the leading and trailing burst of ~40–55 ms, i.e., no onset to the trailing burst occurs for smaller intervals. This interval has its proposed neural basis in the AHP or any other hyperpolarizing process that only depends on time since leading burst onset and has an ~55-ms recovery time constant. Potential other mechanisms that could cause the post activation suppression after the onset response to the leading burst are feed-forward and feed-back inhibition via slowly or nonadapting interneurons (Douglas and Martin 1998). These inhibitory inputs result in inhibitory postsynaptic potentials (IPSPs) mediated by GABA_A receptors. These IPSPs could also affect the cortical neuron’s response to the trailing burst. On our basis of our modeling results, IPSPs resulting from GABA_A receptor activation cannot be distinguished from those resulting from Ca^{2+}-gated K^+ channels because they have approximately the same time constants and both start after a few milliseconds after the excitatory response of the neuron. However, if slowly adapting inhibitory interneurons are producing the IPSPs, one expects that the duration of the leading burst will have a stronger effect than what we observed.

The modeling suggests that the recovery of these neurons from their activity during stimulation with the leading bursts can only be explained if recovery from an AHP is incorporated. Fast recovery from depression also plays a role and fully explains the recovery following long leading bursts (cf. Fig. 1). The medium-duration AHP is characteristic for neurons with Ca^{2+}-gated K^+ channels found in the principal cells in the thalamus and pyramidal cells in cortex (Schwindt et al. 1988). AHP potentials have been described in some stellate cells in the ventral cochlear nucleus and in various cell types of the dorsal cochlear nucleus, but they last typically not longer than 20–40 ms (Oertel 1991). AHPs appear to be absent in the medial nucleus of the trapezoid body and the lateral superior olivary complex ( Wu and Kelly 1993). Longer duration, ≲200 ms, AHP potentials are present in the dorsal nucleus of the lateral lemniscus ( Wu and Kelly 1995). However, these cells are able to fire at relatively high repetition rates (up to ~167/s). In the IC, long-lasting AHP potentials are visible in patch clamp recordings (Covey et al. 1996), but IC neurons typically also show following rates up to ≥100 Hz. Thus it is not likely that they will be the cause of the results observed in AI or psycho-acoustically. Simultaneous recording of thalamic principal cells and cortical pyramidal cells in awake guinea pigs in response to sinusoidally amplitude modulated sounds (Creutzfeldt et al. 1980) showed that thalamic neurons could be driven to modulation frequencies up to ~100 Hz, whereas cortical neurons ceased following at ~20 Hz. This is incompatible with a strong AHP in the thalamus and suggests that the rate limiting mechanism is present in cortex. Thus the characteristics that determine the minimum gap dependence on leading burst duration are likely those of cortical neurons.

**Effects of anesthesia**

The ketamine anesthesia used in these experiments likely does not have an effect on the strength of the AHP. The main effect of ketamine is blocking of N-methyl-d-aspartate channels (Orser et al. 1997) and increasing acetylcholine release (Kikuchi et al. 1997). Ketamine has little or no effect on GABA receptors and does not appear to affect such K^+ channels as I_{Na}, I_{K}, and I_{GABA} (Kress 1997). The medium AHP invoked in the explanation of the recovery functions peaks within 1–5 ms, lasts for several hundreds of milliseconds and is based on voltage-insensitive Ca^{2+}-activated K^+ (SK_{Ca}) channels that are blocked by apamin. These channels are activated by Ca^{2+} influx via N-type Ca^{2+} channels (Sah 1996). Whereas an inhibitory effect of ketamine on L-type Ca^{2+} channels has been reported (Wong and Martin 1993), its effect on N-type channels manifests itself only at a dose much higher than the anesthetic one (Kress 1997). The AHP is dependent on the strength of the onset response (Schwindt et al. 1988), but ketamine has little effect on the strength of onset responses.

**Comparison with other studies**

Peripheral adaptation in the auditory system occurs with time constants of ~40–50 ms, and recovery from there has a time constant of 60–80 ms, albeit that there is also a very fast adaptation component with a time constant of 5–15 ms (Eggermont 1985; Meddis 1986). The fast adaptation and recovery process in cortex is characterized by an adaptation time constant of 12 ms and a recovery time constant of 20 ms. This short recovery time constant makes it unlikely that the depression that we observe results from peripheral mechanisms. The value for τ_{adap} = 12 ms used in this study is comparable to the adaptation time constant of 11.5 ± 1.3 ms found by Ahmed et al. (1998) for intracellular recordings from superficial layers of cat visual cortex. Most of our recordings were from depths corresponding to layers III and IV.

In agreement with Brosch and Schreiner (1997), it is found that the recovery after a fixed gap (e.g., 20 ms) is dependent on the masker duration as long as this exceeds 20 ms. The masking effect decreased for longer intervals between the onsets of masker and probe. Our results for leading burst lengths 20 ms suggest that recovery was, on average, complete after gap durations of 50 ms. The average maximal duration of tone-on-tone forward masking found by Brosch and Schreiner (1997) was 143 ms (range 53–430 ms), but masker levels were generally higher than the probe level, and the masking effect increased with level.

For the shortest leading burst duration, a wide distribution of minimum detectable gaps is found with values as low as 10 ms. Previously (Eggermont 1999a) we reported values as low as 5 ms (the shortest gap used). Such values are approaching the minimum gap that can be detected perceptually when the leading and trailing burst have the same frequency content. This within-perceptual-channel condition is obviously the same as that used in the experiments reported here. It is likely that this small subset of neurons that respond to very small interruptions of a sound with an on response to the trailing burst can account for this finding. From our previous analysis (Eggermont 1999a), it could be inferred that this group of neurons has onset firing rates of <0.75 spike · burst^{-1} · bin^{-1}. 

```
and that they fire either to the leading or trailing burst. Thus these short gaps would not be represented in single-unit inter-spike intervals. Recent results of Fitzpatrick et al. (1999) in auditory cortex of the unanesthetized rabbit using double clicks found recovery times as short as a few milliseconds for a substantial fraction of neurons. A large group of neurons had recovery times comparable to ours and some had longer recovery times. This could represent a species difference. Similar to our results, multi-unit activity recorded in awake monkeys showed a double onset response to a /ta/ phoneme that tracked the duration of the VOT (Steinschneider et al. 1994).

In conclusion, our findings suggest that the firing patterns of neurons in auditory cortex can potentially account for within-channel and across-channel perceptual detection tasks. It is likely that the within-channel task is based on population activity that represents both the onsets to the leading and trailing burst in neurons with low firing probability using a rate representation. In contrast, the across-perceptual-channel detection task would access a temporal representation, as reflected in the autocorrelogram, of both onset responses in the same neurons. Such a representation can only occur for onset-onset intervals >50 ms and is limited by intrinsic neuron properties specific to cortical cells. Intrinsic properties of cells in auditory cortex can likely fully account for the behavioral findings of across perceptual channel gap detection in humans.

M. Kimura, H. Komiyi, and P. Valentine assisted with the data recording. G. Shaw assisted with the data analysis.

This investigation was supported by grants from the Alberta Heritage Foundation for Medical Research and the Natural Sciences and Engineering Research Council of Canada and by the Campbell McLaurin Chair for Hearing Deficiencies.

REFERENCES


