Neural Correlates of Gap Detection in Three Auditory Cortical Fields in the Cat

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

Eggermont, Joseph. Neural correlates of gap detection in three auditory cortical fields in the cat. J. Neurophysiol. 81: 2570–2581, 1999. Minimum detectable gaps in noise in humans are independent of the position of the gap, whereas in cat primary auditory cortex (AI) they are position dependent. The position dependence in other cortical areas is not known and may resolve this contrast. This study presents minimum detectable gap-in-noise values for which single-unit (SU), multiunit (MU) recordings and local field potentials (LFPs) show an onset response to the noise after the gap. The gap, which varied in duration between 5 and 70 ms, was preceded by a noise burst of either 5 ms (early gap) or 500 ms (late gap) duration. In 10 cats, simultaneous recordings were made with one electrode each in AI, anterior auditory field (AAF), and secondary auditory cortex (AII). In nine additional cats, two electrodes were inserted in AI and one in AAF. Minimum detectable gaps based on SU, MU, or LFP data in each cortical area were the same. In addition, very similar minimum early-gap values were found in all three areas (means, 36.1–41.7 ms). The minimum late-gap values were also similar in AI and AII (means, 11.1 and 11.7 ms), whereas AAF showed significantly larger minimum late-gap durations (mean 21.5 ms). For intensities >35 dB SPL, distributions of minimum early-gap durations in AAF and AII had modal values at ~45 ms. In AI, the distribution was more uniform. Distributions for minimum late-gap duration were skewed toward low values (mode at 5 ms), but high values (≥60 ms) were found infrequently as well. A small fraction of units showed a response after the gap only for early-gap durations <20 ms. In AI and AII, the mean minimum early- and late-gap durations decreased significantly with increase in the neuron’s characteristic frequency (CF), whereas the lower boundary for the minimum early gap was CF independent. The findings suggest that human within-perceptual-channel gap detection, showing no dependence of the minimum detectable gap on the duration of the leading noise burst, likely is based on the lower envelope of the distribution of neural minimum gap values of units in AI and AAF. In contrast, across-perceptual-channel gap detection, which shows a decreasing minimum detectable gap with increasing duration of the leading noise burst, likely is based on the comparison of on responses from populations of neurons that converge on units in AII.

INTRODUCTION

The auditory system performs a combined spectral and temporal analysis of sound, both with good to excellent resolution. The temporal resolution determines the ability to discriminate rapid changes in the frequency-dependent intensity of a sound over time. This resolution will depend on the ability of individual auditory neurons to represent these changes. The ability to follow rapid sinusoidal AM of a tone or noise decreases from the auditory nerve to auditory cortex. Individual auditory nerve fibers have a low-pass AM-following capacity with a limiting rate of ~1.5 kHz (Joris and Yin 1992). Neurons in the ventral cochlear nucleus (VCN) generally show a band-pass sensitivity with a best modulation frequency (BMF) ~200 Hz (Frisina et al. 1985; Møller 1972). In the central nucleus of the inferior colliculus (ICC) (Rees and Møller 1983), the most frequently found BMF is ~50 Hz. For the medial geniculate body (MGB), the limited data available (Rouiller et al. 1981) suggest BMFs between 25 and 50 Hz. For primary auditory cortex (AI), the BMF is ~10 Hz and limiting rates tend to be <50 Hz (Eggermont 1997; Schreiner and Urbas 1988). However, in awake paralyzed cats, de Ribaupierre et al. (1972b) obtained in about one-quarter of the recordings locking in regular spiking cells up to click repetition rates of 100 Hz. Human psychophysical data, expressed as a modulation threshold, i.e., the smallest modulation depth in the sound that can be perceived, suggest limiting rates ≤150 Hz for pure tone carriers (Fassol and Kohlrausch 1995) and somewhat lower for noise carriers (Viemeister 1979). Clearly, psychophysical performance in detecting AM resembles neural responses in ICC and MGB rather than those in auditory cortex.

More abrupt changes in sound envelope have to be discriminated in gap-detection experiments. Here, a brief gap usually is inserted at the temporal midpoint of a relatively long tone or noise burst. For wideband noise, the minimum detectable gap is 2–3 ms (Plomp 1964). This is a condition where the leading and trailing noise or tone are said to be within the same “perceptual channel.” Under these conditions, the auditory system is most sensitive to perceptual differences, and gaps likely are detected on basis of a discontinuity in the neural representation (Phillips et al. 1997) as a result of the decrement in the noise (Eddins and Green 1995).

The psychophysical performance of detecting AM in noise and short-duration gaps inserted in unmodulated noise (a within channel condition) appear to be both explicable on basis of a model that subjects the noise to half-wave rectification, followed by low-pass filtering with a time constant of 3 ms and energy-detection (Forrest and Green 1987). This model successfully predicts the ~3 dB cutoff frequency of ~50 Hz and the ~3 dB/octave slope for the behavioral temporal modulation transfer function (tMTF) to AM noise, as well as the 3-ms minimum detectable gap. However, for auditory cortex the average high-frequency slope of the neural tMTF is ≥12 dB/octave, and the cutoff frequency is of the order of 20 Hz (Eggermont and Smith 1995). This is clearly different from behavioral performance.
For a gap inserted in noise with a 50-Hz bandwidth, the gap-detection threshold is increased to 10 ms (Eddins et al. 1992). In general, for gaps in noise, the gap thresholds appear to vary as the reciprocal of the square root of the noise bandwidth, an indication that fluctuations in the noise amplitude affect the detection threshold (Eddins and Green 1995). This also can be explained by the energy detection model described in the preceding text.

Phillips et al. (1997, 1998) recently have shown that when the frequency content of the leading and trailing noise or tone burst is sufficiently different or when they are presented to different ears or from different positions in space and thus require an across perceptual channel comparison, different results are obtained. Under these conditions, the central auditory system must perform a relative timing operation on the firings from different groups of neurons. Specifically, Phillips et al. (1997) showed that the duration of the leading noise or tone burst determines the minimum detectable gap in across channel comparisons but not for within channel comparisons. The shorter the leading burst, the longer was the minimum detectable gap.

For auditory cortical neurons, a discontinuity in the neural representation of the change in noise level has to be defined clearly because these neurons respond largely in a transient way to a noise burst onset. In anesthetized animals, there is initially a short burst, generally less than five spikes and most frequently consisting of one to two spikes. This is followed by a postactivation suppression period produced by a hyperpolarizing afterpotential (AHP) that lasts 60–150 ms (de Ribaupierre et al. 1972a). The post activation suppression usually is followed by one or more rebound spike bursts, after which the firing rate settles at a level that is indistinguishable from the spontaneous activity level. The representation of a discontinuity in the sound by a cortical neuron thus would require the presence of an on response after the discontinuity. The size of this second on response will depend on the amount of forward masking produced by the leading burst (Eggermont 1995a) as well as on its time of occurrence with respect to the magnitude of the postactivation suppression. The amount of forward masking in auditory nerve fibers increases with the duration of the leading burst and is maximal for a burst of ~100-ms duration (Eggermont 1985). The effect of the postactivation suppression will be strongest immediately after the end of the on response and gradually diminish over the next 60–150 ms. Thus postactivation suppression could be more prominent for leading bursts of short duration, whereas forward masking may determine the effect of longer-duration leading bursts.

For SUs in AI, the average minimum detectable gap placed after a 5-ms duration leading noise burst (the “early-gap” condition) was generally >30 ms (somewhat depending on noise level). In case the gap was positioned after a 500-ms burst of noise (the “late-gap” condition), the gap detection threshold was equal to the shortest gap used: 5 ms (Eggermont 1995b). The late-gap findings in AI are thus compatible with the perceptual ones in the classical gap detection experiment with a long-leading noise burst. The early-gap condition also mimics the temporal position of the voice onset time (VOT) for a /ba/-/pa/ continuum, where a short noise burst is followed by a VOT and a vowel. Whereas the early gap-in-noise condition has equal frequency content of the bursts before and after the gap, the frequency content of the leading noise bursts and the trailing vowel in /ba/-/pa/ discrimination is different. The gap-in-noise condition is perceptually a “within channel” condition, and the VOT detection in the /ba/-/pa/ continuum is an “across channel” condition. For primary auditory cortex (AI) neurons, the minimum VOT in the /ba/-/pa/ continuum showed a stronger intensity dependence than the minimum detectable early gap (Eggermont 1995a).

On the basis of spectral content alone, the early and late gap-in-noise conditions present perceptually and neurophysiologically within channel comparison conditions, i.e., the leading and trailing noise bursts likely activate the same neurons. Thus the insensitivity to the duration of the leading tone burst found in the Phillips et al. (1997) experiments for the within channel condition (same frequency of the leading and trailing tone burst and presented to the same ear) at first glance does not appear to have a SU or MU correlate in AI. It is, however, possible that other cortical areas would show a different sensitivity to the early-gap stimulus, for instance, through the existence of a small subgroup of neurons that is sensitive to very small early-gap durations. The lower envelope principle (Barlow 1995; Eggermont 1998a; Parker and Newsome 1998) dictates that sensory thresholds are set by those neurons that have the lowest threshold for the particular stimulus used. As a consequence, even a small subgroup of neurons sensitive to very short gaps will be sufficient to allow these gaps to be detected psychophysically.

In this study, minimum detectable gap values for single neurons and LFP obtained for simultaneous recordings from three auditory cortical areas will be presented. It will be shown that the previously reported large differences in the neural representation of minimum early and late gaps in AI, generalize to the anterior auditory field (AAF) and secondary auditory cortex (AII). It is hypothesized that two subpopulations of neurons may form a neurophysiological basis for the behavioral differences in within and across channel gap detection.

Methods

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 were premedicated with 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 ~0.5 h, they received an intramuscular injection of 25 mg/kg of ketamine (100 mg/ml) and 20 mg/kg of pentobarbital sodium (65 mg/ml). Lidocaine hydrochloride (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 intramuscular injections of 2–5 mg · kg⁻¹ · h⁻¹ of ketamine. The wound margins were infused every 2 h with lidocaine hydrochloride, 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 placed 55 cm in front of the cat’s head. 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&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, gammashape 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 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. Each gap stimulus was presented 10 times. The noise bursts used consisted of “frozen” noise, i.e., they were the same for all conditions.

For some recording sites with CFs after 5 kHz, a /ba/-/pa/ continuum with VOTs ranging from 0 to 70 ms in steps of 5 ms was presented in random order. These phonemes were generated with a parallel/cascade Klatt-synthesizer KLSYN88a implemented on a Macintosh II using a 20-kHz sampling frequency. The total duration of the stimulus was 250 ms, regardless of the duration of the VOT; the onset noise-burst was 5 ms in duration. The dominant frequency ranges for the vowel part were F0 = 120 Hz, F1 = 700 Hz, and F2 = 1,200 Hz. The fundamental frequency was 120 Hz. The phonemes were presented once every two seconds. The only parameter that changed in the continuum was the VOT. Twenty repetitions were presented at each VOT value.

Data analysis

The number of action potentials in the first 100 ms after each tone-pip presentation were counted for each intensity. The counts 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 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 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.

Poststimulus time histograms (PSTH) with 1-ms binwidth were calculated for each of the two “stimuli” in each type of stimulation (e.g., leading noise bursts and trailing noise burst). From these PSTHs both the peak latency and the spike count were calculated. The minimum value of the gap or VOT duration, for which a “double on response” was obtained was called the minimum gap (or minimum VOT).

Spike train analysis was done using custom-designed software (Eggermont 1992, 1996) and the Stranger software package. All statistical analyses were performed using Statview 4.5. Graphics were done with Powerpoint and Horizon software.

RESULTS

In 10 cats simultaneous recordings were made with one electrode in each of AI, AAF, and AII; in 9 additional cats, two electrodes were inserted in AI and one in AAF. In total 28 recording sites gave both reliable MU and/or LFP recordings in AI, 18 recording sites resulted in good data from AAF, and for 11 recording sites, both MU and LFP data were recorded from AII. CFs ranged from 2 to 22 kHz in AI, from 2 to 21 kHz in AAF, and from 2 to 18 kHz in AII.

Individual examples

First of all I present, for a recording in AI with a CF = 3.5 kHz, a comparison of the responses to a /ba/-/pa/ continuum
and to the early and late gap in noise conditions at three different intensities (Fig. 1). The individual SUs in these raster plots are indicated by differently colored dots. The oblique lines indicate the onset of the vowel or trailing noise burst. The top row shows the responses to the /bal/-/pal/ stimulus continuum. A clear onset response with a latency of ~15 ms at 60 dB increasing to a little more than 20 ms at 40 dB is present. At 60 dB SPL, the second ON response is limited to VOTs of >30 ms but with the most vigorous response for VOT durations in excess of 45 ms. The response latency to this second ON response, the time difference between the oblique line and the second ON response, is much longer than for the ON response to the leading burst. Essentially the same is found at a level of 50 dB SPL. However, for 40 dB SPL stimulation, double ON responses are found even after the lowest VOT of 0 ms (because of the different spectral content of the vowel and the noise burst), although the most vigorous response remains present for VOTs >40 ms. The early-gap responses are shown in the second row. The onset responses have very similar latency as those for the /bal/-/pal/ stimuli. A big difference is found in the second ON response, which at 60 dB SPL is only found for gaps larger than 50 ms and is very weak as well. There is, however, a strong rebound-like response with ~40 ms longer latency than the second ON response that parallels the end of the gap for durations >30 ms. At lower gap durations, this late response has a constant or slightly increasing latency. Thus there is a double ON response for all gap durations including zero, suggesting that this response represents a rebound from postactivation suppression. At 50 dB SPL, the second ON response is absent, but the rebound response is present for all real gap durations. At 40 dB SPL, the second ON response is weakly visible for all gap values just as for the VOT stimulus, whereas the rebound response is much reduced. The responses to the late gap are shown in the bottom row. For both 60 and 50 dB SPL, the second ON response is visible for gap durations as low as 5 ms and with a latency comparable.
With that for the ON response to the leading burst (second row). At these intensities, a rebound response follows the second ON response with a latency of 80 ms. At 40 dB SPL, only an OFF response to the end of the leading noise burst was found. At higher stimulus levels, this OFF response is absorbed or abolished by the second ON response when its latency becomes shorter than that of the OFF response.

Figure 2 shows a comparison, across the entire intensity range studied, between minimum VOTs and minimum early-gap durations for MU and local field potentials (LFPs) for the recordings shown in Fig. 1 and their dependence on stimulus intensity. With that for the ON response to the leading burst (second row). At these intensities, a rebound response follows the second ON response with a latency of 80 ms. At 40 dB SPL, only an OFF response with a latency of 45 ms to the end of the leading noise burst was found. At higher stimulus levels, this OFF response is absorbed or abolished by the second ON response when its latency becomes shorter than that of the OFF response.

Figure 2A shows a comparison, across the entire intensity range studied, between minimum VOT and minimum early-gap durations for the previous example. The VOT data cluster between 0 and 10 ms for minimum gaps <40 ms and between 25 and 50 ms for minimum gaps >40 ms (with a slight overlap). As a function of stimulus intensity, a similar clustering in minimum VOT is observed (Fig. 2B). For intensities <40 dB, there is generally no ON response to the noise burst, and in this case, the second ON response to the vowel is always present. For higher intensities, the presence of an ON response to the noise burst prevents ON responses to the vowel for the smaller VOT. LFP and MU data generally show the same behavior.

An example of clustered SU responses to early and late gaps simultaneously recorded in three cortical areas is shown in Fig. 3. At each electrode, the cluster consists of three separated SUs with similar response patterns. The oblique lines again indicate the onset of the trailing noise burst. The CFs for the recording sites were 7 kHz for AI and 4 kHz for AAF. Frequency-tuning-curve bandwidths at 20 dB above threshold were less than one octave for AI, about two octaves for AII, and about three octaves for AAF. The intensity of the noise burst was 60 dB SPL. Looking at the ON responses (Fig. 3, left), one observes that the response in AAF has the shortest latency and that those in AI and AII were similar. This was a general finding, latencies in AAF for noise bursts were on average ~3 ms shorter than in AI or AII. The response to the early gap is sharpest in AII and broadest in AAF, but irrespective of the area the minimum gap is ~45 ms. In contrast, the responses to late gaps are obtained for all gap durations. No OFF response to the leading 500-ms noise burst was found. The latencies of the ON response to the trailing noise burst are again shorter for the late gap. Spontaneous activity (right, showing the last 200 ms of the 2-s silent period after the 1-s duration noise burst) was also comparable for the three areas.

For the combination of the individual SU data per electrode from the previous example, Fig. 4 presents PSTHs (1-ms bins) for three selected gap durations (5, 50, and 60 ms) and each area. The time base is 700 ms, so onset, early gap, and late-gap responses are all potentially visible. For the 60-ms gap duration (Fig. 4, right), all three recordings show the onset, early gap, and late gap albeit with some differences in relative response strength between areas. For the 50-ms gap, only the recording in AII shows a sharp early-gap response, whereas for AI a sustained combination of a second ON response and a rebound response is visible (cf. Fig. 3). For the 5-ms gap condition, responses to the late gap are present for all three areas.

An example of a MU recording in AI for a 60 dB SPL noise for which the late-gap response was combined with a well-defined OFF response to the leading noise burst is shown Fig. 5, top. This combination of OFF and ON responses was a general finding, latencies in AAF for noise bursts were on average ~3 ms shorter than in AI or AII. The response to the early gap is sharpest in AII and broadest in AAF, but irrespective of the area the minimum gap is ~45 ms. In contrast, the responses to late gaps are obtained for all gap durations. No OFF response to the leading 500-ms noise burst was found. The latencies of the ON response to the trailing noise burst are again shorter for the late gap. Spontaneous activity (right, showing the last 200 ms of the 2-s silent period after the 1-s duration noise burst) was also comparable for the three areas.

For the combination of the individual SU data per electrode from the previous example, Fig. 4 presents PSTHs (1-ms bins) for three selected gap durations (5, 50, and 60 ms) and each area. The time base is 700 ms, so onset, early gap, and late-gap responses are all potentially visible. For the 60-ms gap duration (Fig. 4, right), all three recordings show the onset, early gap, and late gap albeit with some differences in relative response strength between areas. For the 50-ms gap, only the recording in AII shows a sharp early-gap response, whereas for AI a sustained combination of a second ON response and a rebound response is visible (cf. Fig. 3). For the 5-ms gap condition, responses to the late gap are present for all three areas.

An example of a MU recording in AI for a 60 dB SPL noise for which the late-gap response was combined with a well-defined OFF response to the leading noise burst is shown Fig. 5, top. This combination of OFF and ON responses was a general finding, latencies in AAF for noise bursts were on average ~3 ms shorter than in AI or AII. The response to the early gap is sharpest in AII and broadest in AAF, but irrespective of the area the minimum gap is ~45 ms. In contrast, the responses to late gaps are obtained for all gap durations. No OFF response to the leading 500-ms noise burst was found. The latencies of the ON response to the trailing noise burst are again shorter for the late gap. Spontaneous activity (right, showing the last 200 ms of the 2-s silent period after the 1-s duration noise burst) was also comparable for the three areas.
relatively frequent finding (24% of the cases); in most cases there was no clear OFF response (56%), whereas in 20% there was only an ON response. In this example, the OFF response and subsequent ON response in AI to the late gap (Fig. 5, right) are present for gap durations of 15–70 ms; for shorter gap durations, the second ON response is not distinguishable from the OFF response. The early gap (Fig. 5, left) shows a second ON response for gap durations >40 ms in both AI and AAF (bottom). The late-gap response for AAF is not very pronounced but, except for an interruption in the 20- to 35-ms duration range, still discernible in the spontaneous activity. Most of the time, the strong postactivation suppression of spontaneous activity after the leading noise bursts (see also Figs. 1 and 3) makes identification of the gap response easy.

An infrequent finding (4% of the MU recordings, but only in AI and AAF) was a response preference for short minimum early gaps over long ones. In these cases all SUs in the MU record showed this same preference. Figure 6, top, shows the first 400 ms after the noise onset for such a MU recording in AI. Clearly an ON response to the gap is found for gap durations of 5–25 ms and, albeit much weaker, also for gap durations >60 ms. The late-gap response for the same recording (Fig. 6, bottom) consists of a combination of a late OFF response (latency >50 ms but only for gap durations >40 ms) and an onset response to the end of the gap. When the ON response arrives earlier than the OFF response (for gap durations <40 ms), the latter is abolished. A late weak rebound of spontaneous activity is seen after 160–180 ms after the second ON response.
Group data for gap stimuli

For AI, all minimum early- and late-gap durations for which there was a clearly discernible ON response after the gap are plotted as a function of the noise burst level for both MU spikes and LFP triggers (Fig. 7). Not all units responded at all intensities because of nonmonotonic rate-intensity functions. For 25 of 28 recordings, MU data allowed estimation of the minimum early gap either at 50 or 60 dB SPL, and for the minimum late gap, this was the case for 27 of 28 recordings. At lower intensity levels, the number of useful MU and LFP recordings decreased. The 70 dB stimulus was tested for only nine recordings. Because the various SUs in each MU recording all gave the same estimate, I only report the value for the MU recording. The minimum late-gap values are clustered 5–10 ms but with occasionally values as high as 60 ms. The minimum early-gap values, especially for intensities >35 dB SPL, clustered in a broadband between 30 and 70 ms but show occasionally values as low as 5–10 ms.

For AAF, fewer data points are available, especially at low intensity values, but the same trend is visible as in AI (Fig. 8). In all 18 recordings, minimum early-gap estimates were obtained at either 50 or 60 dB SPL for both MU and LFPs. For the late gap, only 13 MU recordings allowed an estimate of the minimum gap at 50 or 60 dB SPL; however, all LFP data allowed an estimate. Again the minimum late-gap values are dominantly found at 5–10 ms, but the proportion of much higher values is larger than found in AI. This makes the mean value in AAF distinctly higher than in AI. The minimum early-gap values again could be as low as 10–15 ms for intensities of 40–70 dB. In AII (Fig. 9) in 8 of 11 recordings reliable MU estimates of the minimum early and late gap were obtained at 60 dB SPL and for 7 recordings also at 50 dB SPL. It is noted that a few minimum early-gap values were as low as 10 ms for the higher intensities.

The distributions of the minimum early and late gaps estimated for MU or LFPs were not significantly different (Kolmogorov-Smirnov test, \( P > 0.1 \)) for the three areas. Mean values and standard deviations (SD) for minimum early and late gaps are listed per area and separately for MU and LFP in Table 1. There was no significant difference between minimum gap durations estimated from MU spikes and from LFP triggers. It is noted that the mean minimum late gap in AAF was about twice as large as in the other two areas.

For the early gap, there were a total of 49 estimates in AI (for 25 MU recordings), 33 in AAF (for 18 MU recordings), and 22 in AII (for 13 MU recordings). For the late gap, the numbers were 33 in AI, 18 in AAF, and 16 in AII. This combination of all minimum early-gap durations for MU responses across areas for intensities >35 dB SPL (Fig. 10) shows that the majority is found with durations between 30 and 50 ms, but that also a substantial number is present for durations of 5–15 ms especially in AI. As a result, the distribution of minimum values for AI is nearly uniform, whereas for the other two areas the 30–50 ms values dominate. The most frequent minimum late-gap duration in all areas was 5 ms, the shortest duration used, but values between 35 and 60 ms were found occasionally (Fig. 11).

For the three auditory areas combined, the intensity dependence of minimum early-gap durations (Fig. 12A), shows a
significant positive correlation \((P < 0.02)\) with intensity. No intensity dependence \((P > 0.25)\) was found for the minimum late-gap duration (Fig. 12B). Note that the regression lines for early- and late-gap conditions are nearly identical for MU spikes and LFP-trigger-based estimates.

For MU spikes the minimum early and late gap for intensities in the range of 30–60 dB SPL decreased with CF. The dependence was significant for AI and AII but not for AAF. Figure 13 shows the dependence for recordings in AI and AII combined, the late minimum gap decreases significantly \((P < 0.05)\) with CF from \(\sim 15\) ms at 2 kHz to \(\sim 5\) ms at 20 kHz. The early minimum gap decreases significantly \((P < 0.0001)\) from \(\sim 55\) ms at 2 kHz to \(\sim 20\) ms at 20 kHz.

**DISCUSSION**

The estimate of the minimum detectable early and late gaps was based on the occurrence of the on response to the trailing noise burst. We have seen that there are at least two types of response in the latency region of interest for this on response. The first is the proper on response, which follows the onset of the trailing noise burst with a nearly constant latency, whereas the other is a rebound response that follows the on response to the leading noise burst with a latency of \(\sim 100–140\) ms. I have only used the proper on response in the analysis. The justification for using only the proper on response to the trailing noise burst is that the rebound response is initiated by the on response to the leading noise burst. Such rebound responses are a manifestation of stimulus synchronized LFP spindling, which only occurs when the animal is in a state of light sleep or drowsiness (Steriade et al. 1993). An alert animal will therefore likely not show this rebound response and could make the behavioral decision unambiguously on the basis of the on responses to leading and trailing noise burst in case of early gap and on the basis of the off response to the leading burst and the on response of the trailing burst in the late-gap condition. From the examples shown, it became obvious that the off response to the leading noise burst was abolished by the on response to the trailing noise burst whenever the latency of the latter became shorter than that of the off response. If the off response contributes to the perceived loudness of the leading noise burst, this abolishment could be a neural correlate of perceptual backward masking.

**Comparison with other studies**

There are only a few electrophysiological studies on gap detection: two for inferior colliculus in young and aged mice...
Walton et al. (1997, 1998), one in mouse inferior colliculus pointing out that small changes in the stimulus envelope have an effect on gap detection (Barz et al. 1998), and one for AI (Eggermont 1995b). The three inferior colliculus studies were of the late-gap type, and the results, albeit including smaller gaps than in the present study, were comparable to ours. The emphasis of the following comparisons will be with psychoacoustical findings in humans.

In mouse inferior colliculus, minimum detectable gaps after a 100-ms noise burst, and thus in the late-gap category, are in the 1- to 2-ms range with only two units having a minimum detectable gap of 11 ms. The similarity of minimum gap durations based on SU, MU or LFP data within each cortical area was a consistent finding. This, taken together with the previous observation of strong similarity of the temporal modulation transfer functions for click train stimulation for LFP triggers and SU spikes (Eggermont and Smith 1995), suggests that LFP triggers are good predictors for the neural representation of temporal stimulus features in auditory cortex. The difference observed between the mean minimum early- and late-gap durations in AI confirm those from a previous report (Eggermont 1995b). In general, AI and AII showed similar late-gap representation properties, whereas AAF presented significantly longer minimum late-gap durations. This difference was largely the result of a relatively large number of long minimum late-gap durations (cf. Fig. 11). In all areas, the distributions of minimum late-gap duration distributions were skewed toward the lowest value (5 ms). The overall distributions for the minimum late-gap values were not significantly different, thus the finding of higher mean minimum late-gap durations in AAF may be not meaningful. Slightly bimodal distributions of minimum early-gap durations were found in AAF and AII areas, whereas in AI the distribution was closer to uniform. Previously (Eggermont 1998b), I showed that temporal modulation transfer functions for periodic click trains and amplitude modulated noise bursts were also very similar in AI, AAF, and AII. Combined with the minimum gap data, this suggests that there is, at first sight, no clear specialization or differentiation in the processing of AM and amplitude discontinuities in these three cortical areas.

Minimum early and late gaps in AI and AII decreased with increasing CF. For the late gap, this finding parallels the psychoacoustical observations that the gap detection threshold in pure tones decreases with increasing pure tone frequency (Eddins et al. 1992). This finding, however, has been attributed to spectral splatter, and with proper control thereof gap thresholds were, psychoacoustically, found to be independent of frequency (Eddins and Green 1995). It has been observed previously (Eggermont 1996; Schreiner and Mendelson 1990) that frequency tuning curves in AI have a fairly constant range of $Q_{10\text{dB}}$ or $Q_{40\text{dB}}$ values as a function of CF, but these values decreased toward the dorsal and ventral border. Thus the bandwidth increases approximately proportional to CF, and as a result the dependence of the neural minimum gap duration on CF thus could be a dependence on bandwidth. However, frequency-tuning curve bandwidths in AII are at least twice as broad than those in AI (Eggermont 1998b). This dependence on bandwidth would have to show up as a cortical area dependence in minimum gap duration. In that case, one would expect shorter minimum gap durations in the cortical areas with broader frequency tuning, i.e., in AAF and AII, and that does not appear to be the case. In the energy detection model (Forrest and Green 1987) that I described in INTRODUCTION, a

![FIG. 12. Dependence of minimum gap duration across all 3 areas on stimulus intensity for levels >25 dB SPL and with linear regression lines drawn in. Both MU and LFP data are shown. Linear regression lines for MU and LFP are very similar for both the early and late gap. To avoid complete overlap of data points especially at the 5-ms, 10-ms, etc., gap durations in this plot, the plot values were randomized uniformly so that those for gap values of 5 ms appear between 3 and 7 ms, for 10 ms gap appear between 8 and 12 ms, etc., and also for the intensity axis so that the 50-dB points are uniformly scattered between 47.5 and 52.5 dB, etc. This was only done for graphic purposes and not used in any of the analysis. Regression lines are identical to those for the actual values.

![FIG. 13. Dependence of early and late minimum gap duration for MUs on CF for AI and AII for intensities >25 dB SPL. For both, there is a significant decrease in minimum detectable gap with increasing CF. Bottom envelope for late gaps is independent of CF.]
frequency filtering was present before the rectifier stage, but it was so broad (bandwidth is 4 kHz) that psychoacoustically both sensitivity to AM of noise and gap detection in noise appear to rely on information from a wide range of frequencies, well beyond those within a single critical band. Looking again at Fig. 13, it is noted that the lower boundary of the minimum late gap is independent of CF. If the psychoacoustic minimum detectable gap in the within channel condition is determined by the lower envelope principle, then psychoacoustic and physiological results are in agreement. For the early gap, the lower boundary of the data also shows a dependence on CF.

Changes in the total stimulus duration from 5 to 400 ms (gap in the middle) did not have much effect on the psychoacoustical detection threshold (Forrest and Green 1987). This is in stark contrast to our neural findings. Forrest and Green (1987) also showed that the gap detection threshold was largely independent of the temporal position within the noise as confirmed by the within channel results of Phillips et al. (1997). Obviously, the neuronal representation of early and late minimum gap durations in auditory cortex at first glance suggests otherwise. Can physiology and psychophysics be reconciled in this respect as well?

Neural substrate of psychophysics

Psychophysical decisions are based on the neural coding of the sensory activity. Johnson (1980), in a discussion of the neural processes that precede the perceptual discrimination decisions, distinguishes "neural representation" from "neural code," and it is relevant to review his definitions. A neural representation was defined as "the complete spatiotemporal array of neural activity of all the neurons passing through a transverse plane across a sensory system." Neural representations thus exist at the level of the auditory nerve, at the level of the inferior colliculus, etc. These neural representations differ as a result of transformations taking place in each nucleus or part thereof. For instance, it is widely believed that the temporal representation present in the activity pattern of the auditory nerve has been transformed into a firing rate representation at the level of the inferior colliculus (Palmer 1995). A neural code was considered to be a parameter of the afferent discharge in a neural representation on which discrimination is based. This parameter, an N-dimensional vector (with N the number of neurons) contains all the information needed for the decision process. Examples of such parameters are the instantaneous firing rates, the integrated firing rates, or mean interspike interval durations (Eggermont 1998a). The usefulness of the neural representation for constructing a neural code depends on the sensitivity of the neurons to detect the changes to be discriminated and the variability in the individual neuron's response combined with the correlation between the responses of the individual neurons. Correlations between the firings of individual neurons generally diminish the representational capacity of a population when a firing rate code is used (Zohary et al. 1994), whereas correlations are relevant in temporal or synchrony codes (Stevens and Zador 1998).

Phillips et al. (1997) have argued that gap detection is based largely on the detection of discontinuities in the neural signal representation. Thus the first problem concerns the neural representation of such a discontinuity. It cannot just be an interruption in the firing pattern of the pyramidal cell because of the transient nature of these cell's responses to sound. So it must be sought in the occurrence of a second on response to the start of the trailing noise burst. What is it that allows a second on response of a cortical neuron in case of a long leading burst, even after only 5 ms of interruption of the sound? Is this the result of an instantaneous release of inhibition from the non-adapting interneurons? What is likely relevant here is in the neural representation of the transition from the leading to the trailing burst (Eggermont 1998a).

Perceptual channels generally are defined in terms of critical bands. When the leading and trailing tone bursts are more than one critical bandwidth apart, they belong to different perceptual channels. In case noise is used for both leading and trailing bursts, one cannot speak of within or across perceptual channels as defined for pure tones. In fact, Forrest and Green (1987) argue, on basis of their energy detection model, that the detection in such a case is made across a large number of critical bands. For recordings of SU responses to gap-in-noise stimuli, the leading and trailing signals' frequency content is the same and determined only by the frequency-tuning curve bandwidth of the neuron. The bandwidth of the SU frequency-tuning curves in auditory cortex at levels >10 dB above threshold is generally wider than the critical bandwidth. The critical bandwidth roughly equals the bandwidth of auditory nerve fiber tuning curves at 10 dB above threshold (Moore 1995) and distinguishes within- and across-channel processing for narrowband signals. This implicates that cortical neurons at normal ambient sound levels operate on inputs from units with frequency sensitivity more than one critical band apart. One thus could argue that cortical neurons integrate across channel information just as the listener does in a psychoacoustic experiment. It is thus even more surprising that the local neuronal group and the listener behave so differently for short leading burst gap detection.

Does a single neuron or a small group of neurons recorded on the same electrode belong to one perceptual channel? The neuronal group might behave differently for a /ba/-/pa/ continuum and for an early gap-in-noise continuum (cf. Figs. 1 and 2). It is thus likely that a local neuronal group belongs to various perceptual channels depending on the task. This is not unlike the neural assembly concept, which calls for dynamically forming and disbanding groups of neurons into assemblies (Abeles 1991). Is a perceptual channel composed of all the neurons with the same CF in the sample or only of the most sensitive neurons? As Phillips et al. (1998) remark, similar CF does not mean topographic proximity in the cortical map at suprathreshold levels. Thus at each level, several noncontiguous patches of neurons respond to that pure tone or critical bandwidth sound (Phillips et al. 1994). As a consequence, a perceptual channel is not a fixed neuroanatomical entity.

Within-channel gap detection

In case neurons respond differentially to the duration of the leading noise burst, e.g., by their different sensitivity to forward masking or by a strong synaptic facilitation effect, one would expect the distribution of minimum gaps that we obtained. For the early gap, this presents a slightly "bimodal" distribution with a small group of neurons showing minimum gaps close to the perceptual minimum in a within-channel comparison. Some of these neurons are unique: they preferen-
tially respond to short minimum gaps only. This response pattern indicates synaptic facilitation for short gaps or a delay in the onset of the AHP, allowing ON responses for short gap durations but producing strong depression for longer gaps. The independence of the leading burst duration in within-channel psychophysical experiments (Phillips et al. 1997) may be determined by the relatively few neurons that prefer short minimum early-gap durations. Individual neurons may be the dominant coding units for near threshold stimuli and other just noticeable differences in psychophysical (behavioral) tasks. This has been formalized in the lower-envelope principle (Barlow 1995; Parker and Newsome 1998), which states that sensory thresholds are set by those sensory units that have the lowest threshold for the particular stimulus used. Thus behavioral thresholds, such as those for pure tones, may be represented by a limited set of sensitive neurons and are little influenced by the presence or absence of responses in the larger number of other neurons that are less sensitive to that stimulus. For discontinuity representation, the same principle may hold. The data presented here suggest that for minimum detectable early-gap representation, AII and AAF may be best suited because of the larger abundance of short minimum gaps and especially the presence of a small sample of neurons specialized in representing short minimum gaps.

**Across-channel gap detection**

In across-channel detection, the release of inhibition at the end of the leading burst cannot be the physiological cue for gap detection because it only occurs in one channel. Instead comparisons have to be made between the ON responses, adapted firing rates, or OFF responses of neurons responding to the leading burst and the ON responses of other neurons responding to the trailing burst. For efficiency reasons, the OFF responses would provide the best marker for the end of the leading burst provided that its duration is long enough to permit one. In this study, OFF responses to the end of the 500-ms burst preceding the late gap were present in 44% of the neurons, and in 24%, there was both an OFF and an ON response. For across-channel comparisons with short-duration leading noise bursts, e.g., by stimulating left and right ear, respectively, with the leading and trailing noise burst, the comparison will likely have to be between the onset responses to both the leading and trailing noise burst.

Does the bimodality in the minimum early-gap representation across neurons, or the different type of neurons in short versus long gap threshold found in this study, indicate at least two neurophysiological channels for early-gap detection? Is the across-channel gap detection based on different groups of neurons than those that allow detection of a discontinuity in a sound with a constant frequency content? As argued before, discontinuity within a channel may be based on fast cessation of neural inhibition, whereas discrimination across channels as in categorical perception may require timing the difference between OFF and ON responses, or between two ON responses, in potentially nonoverlapping populations of neurons.

From our previous (Eggermont 1995a) and present (cf. Fig. 1) studies of the intensity dependence of the second ON response for a /ba/-/pa/ continuum, it is evident that there is always a response to the vowel onset when the leading noise burst does not activate the neuron. Thus at low to moderate intensity values, the minimum VOT is close to the minimum gap-in-noise for a within channel condition. Across channel conditions using different frequencies for the leading and trailing tone burst thus should, when the frequencies are sufficiently different, always produce ON responses in the two populations of neurons responding, respectively, to the leading tone burst and the trailing tone burst, independent of the duration of the gap. Thus to account for the psychophysical findings, the activity of the two populations of neurons must converge on neurons with properties as described in this paper. The neurons receiving this convergent input should produce a double ON response only when the time encompassed by the duration of the leading burst and the gap is sufficiently long. If these neurons would compare the OFF response to the leading burst with the ON response to the trailing burst, a dependence on the leading burst duration would not be found. This convergence assumption requires that the receiving neurons have broad frequency-tuning curves. This rules out units in AI but suits units in AII that have on average frequency-tuning curve bandwidths at 20 dB above threshold that are at least one octave larger than in AI (Eggermont 1998b). The recipient neurons should also have a minimum early-gap distribution with a modal value ~40 ms and should not be able to encode very short-duration early gaps. This also fits the unit’s properties in AII. So it is likely that AII is a candidate area where the physiological substrate for the across channel gap detection is located. The psychophysically observed inverse relation between gap threshold and leading burst duration must have a basis in a slowly changing modal value in the number of neurons sensitive to smaller and smaller gaps. This would require that for increased durations of the leading bursts, the modal minimum gap value that is represented in cortical neurons as a double ON response will gradually shift toward lower values, i.e., from ~40 ms for a leading burst duration of 5 ms to ~5 ms for leading burst durations in excess of 50 ms. This can be experimentally tested.

**Conclusions**

Minimum gap measurements based on SU, MU, or LFP data in each cortical area were not significantly different. For minimum early gap, AAF and AII showed very similar distributions with a modal value ~45 ms, whereas AI showed a more uniform distribution. For minimum late gap, AAF showed larger minimum late-gap durations compared with the other two areas. The observed difference between the mean minimum early- and late-gap durations parallels the findings in across channel psychoacoustic measurements but not those based on within channel comparisons. However, a small percentage of neurons also showed responses after very short (5 ms) early gaps, and these may be responsible for the within channel results that show independence of the minimum detectable gap and the duration of the leading noise burst. The across-channel perception results can be explained by a convergence of ON response in different neural populations to units in the broadly tuned AII.

M. Kenmochi and M. Kimura assisted in the data collection.

This work was supported by the Alberta Heritage Foundation for Medical Research, the Natural Sciences and Engineering Research Council, and the Campbell McLaurin Chair for Hearing Deficiencies.
Address for reprint requests: Dept. of Psychology, University of Calgary, 2500 University Dr. N.W., Calgary, Alberta T2N 1N4, Canada.

Received 30 October 1998; accepted in final form 1 February 1999.

REFERENCES


