Title: Human sensitivity to spatially-patterned amplitude modulation of incoherent noise fields in the horizontal plane

Running Title: Spatial pattern hearing

Authors: Tom S. Lappas and Michael D'Zmura*

Affiliation: Department of Cognitive Sciences, University of California, Irvine

December 9, 2009

*To whom correspondence should be addressed.

Department of Cognitive Sciences

University of California, Irvine

SBSG 2332

Irvine, California 92697-5100 USA

ABSTRACT

Two experiments were performed to investigate human sensitivity to spatially-patterned amplitude modulations of incoherent noise fields in the horizontal plane. A free-field system with 23 speakers produced patterned modulations of noise source amplitude within the front half of the horizontal plane. The amplitude of each speaker's noise waveform was varied independently in real time. A first experiment tested whether the perceived loudness of a central zone depends on the sound intensity of surrounding zones. Results of this simultaneous contrast experiment suggest that there is no effect of surround level modulation on central perceived loudness. This suggests that human mechanisms of spatial hearing lack spatial opponency. Convergent evidence is provided by measurements of the spatial frequency contrast sensitivity function, which has a cutoff frequency of about two cycles per circle and a lowpass filter characteristic. Results suggest that spatially broad, non-opponent mechanisms mediate human sensitivity to spatially-patterned level modulations of incoherent noise fields.

I. INTRODUCTION

A number of results suggest that humans possess one or more banks of auditory cortical neurons which integrate information across different spatial directions, and that these neurons have broadly-tuned, non-opponent receptive fields that vary in peak directional sensitivity. Boehnke and Phillips (1999) interpret results from gap detection experiments to mean that human sound localization acuity in the horizontal plane is consistent with the existence of just two broadly-tuned spatial filters, each associated with a single ear. They suggest that spatial acuity is based on the activation of two spatially overlapping channels, rather than on the selective activation of a larger number of finely tuned channels. A similar notion has been championed by Middlebrooks and colleagues, who model sound localization in terms of patterns of activity across broadly tuned neurons; differences in the responses of broadly-tuned neurons, with response maxima largely along the interaural axis, provide fine localization across the frontal midline (Middlebrooks *et al*., 1994; Middlebrooks *et al*., 1998; Stecker & Middlebrooks, 2003; Stecker *et al*., 2005).

Neuronal mechanisms with properties that are consistent with such a model are found in several species. Work with ferrets suggests that a large majority of spatially-selective neurons in primary auditory cortex linearly combine sound levels in each frequency band and ear (Schnupp *et al*., 2001); such neurons are spatially non-opponent and are quite broadly tuned in azimuth. Spatially-selective neurons in cat AI cortex have receptive fields shaped by excitatory and inhibitory interactions of signals arriving from the two ears' monaural pathways (Brugge *et al*., 1994); their azimuthal tuning is broad, spatially non-opponent and can exhibit the space-time inseparability characteristic of motion-detection mechanisms (Jenison *et al*., 2001). Work with monkeys suggests that processing is divided into "what" and "where" streams, and that the caudal part of the superior temporal gyrus is crucial to spatial localization (Rauschecker & Tian, 2000; Recanzone *et al*., 2000; Tian *et al*., 2001); neurons there are tuned broadly in azimuth and are spatially non-opponent (Woods *et al*., 2006).

While there is considerable evidence in favor of broadly-tuned, spatially-nonopponent neurons serving spatial hearing, there is some evidence for non-opponent mechanisms tuned narrowly. Carlile and colleagues (2001) find that the results of experiments on spatial adaptation to a broadband noise source are consistent with the activity of many narrowly-tuned spatial channels. Schlack and colleagues (2005) found neurons with spatially narrow tuning (sensitivity limited largely to 15-30 deg) in the multimodally-sensitive ventral intraparietal area (VIP).

Common to all of these findings are spatial filters which are not opponent. The response of such a filter is always increased, to varying extent, by sound sources presented within some range of directions matching the receptive field extent. There are no directions from which a source can diminish the filter's response; the directional sensitivity function of a spatially non-opponent mechanism is of a single sign. On the other hand, spatially-opponent filters have responses that can be either increased or decreased in a way that depends on the direction of the sound source. A simple sort of spatially-opponent filter has a center-surround sensitivity. The response of such a filter is increased, to varying extent, by sound sources presented within some central range of directions, but is decreased by sound sources presented to either side of the central range, *viz*. in a lateral inhibitory surround. Spatially-opponent mechanisms would enhance contours in sound fields and act as bandpass filters in the spatial frequency domain. They may be sensitive to the spatial orientation of a pattern, like edge detectors in visual cortex (Hubel & Wiesel, 1962), and, more generally, support a wide variety of sophisticated spatial processing. Do humans possess spatially-opponent mechanisms?

A first way to proceed uses perceptual judgments. Simultaneous contrast is a visual illusion wherein the brightness of a central zone depends on the brightness of surrounding areas (*e.g.*, Helmholtz, 1896; Katz, 1935). This dependence is often taken as evidence for spatially-opponent processes that compare intensities received at some central location to intensities received at surrounding locations (*e.g.*, Ratliff, 1965). An analogous demonstration in hearing would provide evidence in favor of spatially-opponent auditory processing.

A second way to proceed uses more objective measurement methods and has a long history of success in vision research. It was Schade (1956) who first measured visual spatial-frequency contrast sensitivity early on at RCA Laboratories, in the context of work with television displays. Such a spatial frequency contrast sensitivity function is measured by assessing sensitivity to sinusoidal spatial patterns that vary in spatial frequency. One varies pattern contrast when measuring threshold at a particular spatial frequency, and the reciprocal of the contrast at threshold is defined to be the sensitivity. In the '60s, work at Cambridge introduced the techniques to mainstream visual psychophysics (Robson, 1966) and to visual neurophysiology (Enroth-Cugell & Robson, 1966). A flood of spatial frequency domain studies on spatial sensitivity followed over the next twenty years and now underlie technologies like image compression. Common to these studies are spatially-continuous visual displays that are modulated in both space and time.

A fundamental principle of this work is that, if one Fourier transforms a spatially non-opponent directional sensitivity function into the spatial frequency domain, then one will find a lowpass spatial frequency sensitivity function. The simplest types of spatially-opponent filters typically produce bandpass spatial frequency sensitivity functions when Fourier-transformed. These results are wholly analogous to results for the temporal and temporal frequency domains. One can use sinusoidal spatial patterns of varying spatial frequency to measure a spatial frequency sensitivity function in just the same way that sinusoidal temporal patterns may be used to measure the temporal MTF (Viemeister, 1979). If the spatial frequency sensitivity function is bandpass, then the simplest interpretation is that the underlying detection mechanisms function in a spatially-opponent fashion. Likewise, a lowpass spatial frequency sensitivity function is consistent with a lack of opponency.

One is led naturally to consider how to present spatially-continuous fields of sound which may be modulated in both space and time, as it is with such modulated fields that one is most likely able to study these human mechanisms psychophysically. The next section presents methods that use spatially-patterned amplitude modulations of incoherent noise fields in the horizontal plane; these let one explore human sensitivity to spatial and temporal modulations of dense sound fields. With these methods one can conduct behavioral experiments that address the question of spatially-opponent processing. We present the results of experiments on simultaneous contrast and on the spatial frequency contrast sensitivity function, respectively, which suggest that spatially-sensitive mechanisms of human hearing are not spatially opponent.

II. GENERAL METHODS

A. Equipment

A free-field system with 23 speakers was used (see Figure 1). The speakers (Creative Labs Inspire P7800) are arrayed regularly about a semicircle of radius 122 cm (48 in) that is centered on the position of the listener's head and that spans the directions left, through front and center, to right, in the horizontal plane. The speakers are placed regularly at angular separations of 8.18 deg (180 deg / 22 inter-speaker intervals) in a room of dimensions 3.05m (10 ft) by 3.05m by 3.05m. Sonex acoustical foam panels on walls and ceiling reduce reverberation in the carpeted room.

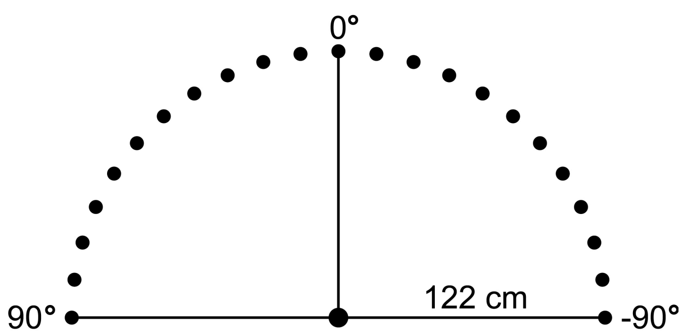


Figure 1. Overview of 23 speakers (small disks) arrayed regularly about a semicircle of radius 122 cm that is centered on the listener's head (larger disk). Speaker azimuths range from 90 deg (leftmost) through 0 deg (front and center) to -90 deg (rightmost) in the horizontal plane.

Speaker output is controlled by four Creative Labs Audigy 4 sound cards installed in a single PC. Subwoofer response for each of the four Creative Labs Inspire P7800 surround-sound systems, one per sound card, is set to zero.

B. Speaker waveforms

Each speaker presents an independent sound pressure waveform at a sampling rate of 48kHz. The presentation is controlled by computer using C++/OpenAL software. The computer controls each speaker's waveform amplitude independently. In these experiments, the speakers present independent narrowband noise waveforms that are created by using a rectangle function to filter white noise waveforms to eliminate energy outside a 2048-4096 Hz passband. Each such waveform has 131,072 samples and so is a source of bandpass noise of duration 2.73 sec. The waveforms were looped to provide noise waveforms of indefinite duration, repeating periodically every 2.73 sec. No aural events at this rate, associated with the looping, were perceived.

The narrowband noise used in the experiments was used also to measure the T_{60} room reverberation time, which is the time taken for the sound to decay to 60 dB below its value when turned off (Sabine, 1922). The reverberation time T_{60} for the room in which the experiments were conducted is 23 msec.

C. Calibration

The system was calibrated using a Brüel & Kjær 2260 Investigator modular precision sound analyzer positioned at head level in the center of the speaker semicircle (see Fig. 1). Master volume controls for each of the four Creative Labs Inspire P7800 surround-sound systems were set to provide single speaker sound pressure levels just under 51.0 dB for an OpenAL software gain (ranging from zero to one) of 0.5 when presenting a bandpass noise waveform described above. A linear relationship between OpenAL gain and output pressure level was verified through measurement for each speaker. For instance, doubling software gain from 0.5 to 1.0

causes an increase in the measured noise waveform sound pressure level of 6 dB ≈ 20 $\log_{10}(2)$. Lines fit to pressure levels, measured as a function of software gain for each speaker, were used to estimate each speaker's output for a gain of one (see Figure 2). These values were then used to create a software lookup table with 23 values, each less than or equal to one, that were used in a multiplicative fashion to cause each speaker to present a level of 50.4 dB for a software gain of 0.5. This level was chosen so that when all 23 speakers present their independent noise waveforms simultaneously at software gains of 0.5, the overall level is 64 dB (≈ 50.4 + 20 $\log_{10}((23)^{1/2})$). Note that level increases as the square root of the number of speakers: (1) each speaker presents an independent noise waveform so that (2) noise variances add and, as a result, (3) sound intensity increases as the number of speakers and (4) sound pressure increases as the square root of the number of speakers. A further series of measurements confirmed linearity when all 23 speakers present their waveforms simultaneously (*e.g.*, 58 dB for a common gain of 0.25, 64 dB at 0.5 and 70 dB at 1.0). The experiments described within involve no manipulation of frequency.

The measurements made to confirm the square-root-law additivity of speaker sound pressure levels were made in the steady state. This confirmation and the short room reverberation time (T_{60} = 23 msec) suggest that there was no untoward build-up of sound caused by reverberation during experiments.

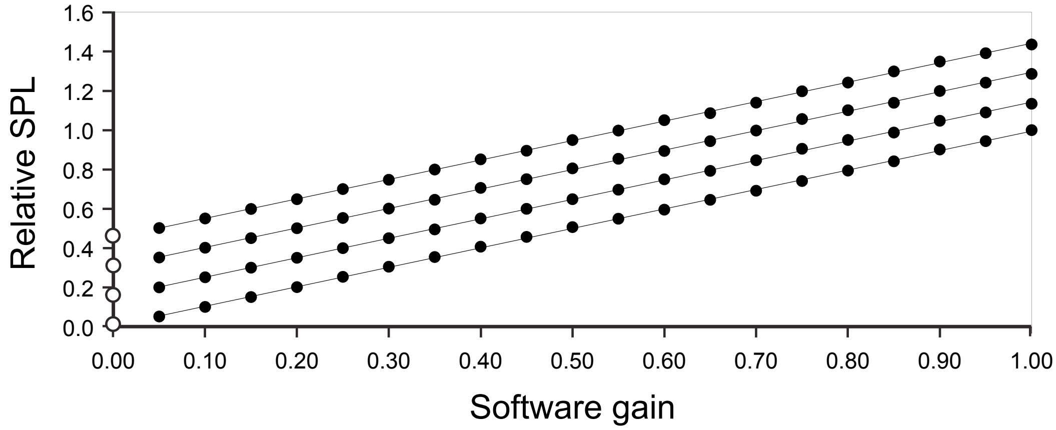


Figure 2. Linear relationship between input software gain value (horizontal axis) and measured output sound pressure level (vertical axis) for four speakers. Open symbols along the vertical axis correspond to measurements made with the speakers off (software gain of zero); absolute SPL values were about 20.1 dB. Filled symbols show data collected with non-zero software gains, indicated along the horizontal axis, using narrowband noise with passband 2048-4096Hz. Data are plotted in a relative fashion to increase visibility; the intercept for the first speaker is set to zero, while the data for the remaining three speakers have been shifted successively upwards by 0.15 units along the vertical axis. Lines through the filled points are the best-fit lines for each of these four speakers; zero-gain measurements were omitted because of a possible floor effect due to ambient noise. These data are wholly representative of the results found for the remaining 19 speakers.

D. Spatiotemporal modulation

Two software applications controlled in real time the sound pressure levels of each of the 23 speakers. The first is a visualization tool that lets one see graphically the time- and space-varying sound pressures and intensities produced by an OpenAL-based modulation engine for a variety of stimulus types. The second uses the same modulation engine, without visualization, to produce stimuli for psychoacoustic experiments.

Speaker levels are modulated as functions of time. Sinusoidal temporal modulations of sound pressure level were used in the first experiment (simultaneous contrast), while Gaussian functions of sound intensity were used to window temporally the spatial Gabor functions presented in the second experiment (spatial frequency contrast sensitivity function). Figure 3 depicts a (truncated) Gaussian function of duration two sec and standard deviation 1/3 sec that provides a smooth temporal window for a stimulus of total duration two sec.

A modulation about a background level may be scaled in several ways. One way is to use a *contrast* scale, where contrast is measured relative to a background level to which the modulation is added. Contrast c is defined in terms of the total pressure t and the background pressure b as $c = (t-b)/b$. For example, consider a single speaker with a background sound pressure level of 50.4 dB corresponding to OpenAL gain 0.5. If the OpenAL gain is now set to 1.0, the sound pressure level will increase 6 dB, which represents a doubling of sound pressure and a contrast of 1.0 = (1.0-0.5)/0.5. An OpenAL gain of 0.0 provides a contrast of -1.0 when presented as a modulation about the background gain 0.5.

The time-varying modulation of gain $m_T(t)$ that is used in software to produce the truncated Gaussian temporal window pictured in Fig. 2 is given by:

$$m_T(t) = \begin{cases} = \exp\left(\left(-1/2\right)\left[\left(t - t_c\right)/\sigma\right]^2\right), \text{ for } 0 \le t \le 2, \\ = 0.0, \text{ for } t < 0 \text{ and } t > 2. \end{cases} \tag{1}$$

In the pictured example, the central time t_c = 1 sec and the standard deviation σ = 1/3 sec.

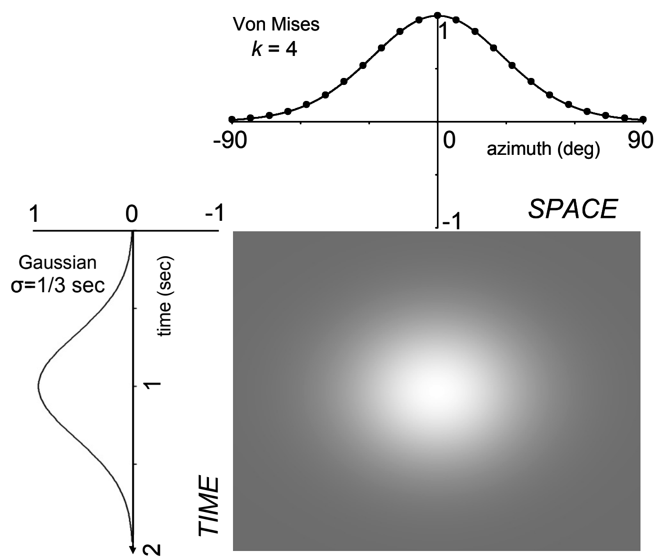


Figure 3. A Gaussian function of time (left) and a Von Mises function of speaker azimuth (top) are used to modulate sound pressure in the space-time separable fashion depicted in the grayscale image. Points in the plot of the Von Mises function indicate sampling by the 23 speaker locations. The background gray in the image represents the background which, in most of the work reported below, corresponds to 50.4 dB for a single speaker set to an OpenAL gain of 0.5 and to 64 dB when all speakers are set to that gain. Increases in sound pressure are coded pictorially as increases in gray value (increased brightness).

Independent control of each speaker's level lets one modulate sound pressure as a function of spatial position. The spatial functions are discrete functions of azimuth in the front half of the horizontal plane; speaker positions sampled the 180 deg range regularly at 8.18 deg intervals.

Figure 3 shows at top a Von Mises function (*e.g.*, Jenison *et al.*, 1998), with shape parameter k set to four, that is sampled regularly on the front semicircle by the speaker array. Von Mises functions are defined on the circle and are analogous to Gaussian functions defined on the real line. The space-varying modulation of gain $m_S[\varphi_i]$ that produces the Von Mises function is given by:

$$m_S[\varphi_i] = n\,\frac{\exp(k\cos[\varphi_i - \varphi_c])}{2\pi I_0(k)}, \quad \text{where } \varphi_i = -\pi/2 + (i-1)(\pi/22) \text{ for i} = 1\ldots 23. \qquad (2)$$

The argument k in Eqn. 2 is the shape parameter of the Von Mises function and controls its width; the shape parameter is set to a value of four in this example. The function I_0 is a modified Bessel function of order zero. We multiply a Von Mises function by a shape-parameter-dependent normalization factor n so that the peak value of the function, found at azimuth $\varphi = \varphi_c$, is one.

A space-time separable modulation $m_{ST}[\varphi_i](t)$ may be produced by multiplying together the space-varying modulation $m_S[\varphi_i]$ and the time-varying modulation $m_T(t)$:

$$m_{ST}[\varphi_i](t) = m_S[\varphi_i]\, m_T(t). \qquad (3)$$

One may produce a space-time modulation of appropriate contrast c about a chosen background by adding to the background an appropriately-scaled space-time modulation. The software gain $g[\varphi_i](t)$ for such a stimulus, depicted by the grayscale image in Fig. 3, is given by:

$$g[\varphi_i](t) = g_b\,(\,1 + c\, m_{ST}[\varphi_i](t)\,)\,, \qquad (4)$$

where g_b is the software gain that produces the desired background sound pressure. Fig. 3 depicts the overall gain for a stimulus with a background gain $g_b = 0.5$ and a contrast $c = 1.0$.

E. Stimulus sound pressure, intensity and level

In the first of the two experiments reported below, the space-time modulations were of the sort described just above. They involve direct modulations of OpenAL gain as per Eqn. 4. To calculate the expected sound pressure $P(t)$ and the corresponding level in decibels $L(t)$ for such a stimulus, calculate first the sound pressure P_b that corresponds to the measured background level L_b in decibels for a single speaker:

$$P_b = 10^{(L_b/20)}. \tag{5}$$

The sound pressure $P[\varphi_i](t)$ produced at time t by the i^{th} speaker is then found using Eqn. 4:

$$P[\varphi_i](t) = P_b\left(1 + c\, m_{ST}[\varphi_i](t)\right). \tag{6}$$

Because the noise waveforms produced by each speaker are independent, their variances add, so that the expected sound intensity $I(t)$ produced by the $N = 23$ speakers is given by:

$$I(t) = \sum_{i=1}^{N} \left(P_b\left(1 + c\, m_{ST}[\varphi_i](t)\right)\right)^2 . \tag{7}$$

The total sound pressure $P(t)$ is the square root of this intensity:

$$P(t) = [I(t)]^{1/2} = \left[\sum_{i=1}^{N} \left(P_b\left(1 + c\, m_{ST}[\varphi_i](t)\right)\right)^2\right]^{1/2}, \tag{8}$$

and the corresponding sound pressure level $L(t)$ in decibels is given by:

$$L(t) = 20\,\log_{10}\left(P(t)\right) = 20\,\log_{10}\left(\left[\sum_{i=1}^{N} \left(P_b\left(1 + c\, m_{ST}[\varphi_i](t)\right)\right)^2\right]^{1/2}\right). \tag{9}$$

Sound intensity rather than sound pressure was modulated directly in the second experiment described below. This second experiment concerns the spatial frequency contrast sensitivity function and, just as with experiments on the temporal MTF (*e.g.*, Viemeister, 1979), is conducted best using sinusoidal modulations of sound intensity. Sound pressure, intensity and decibel level for direct modulations of intensity are described in the section on Experiment 2.

F. Listeners

Three listeners served as subjects in the experiments; all had normal hearing. Individual listeners are indicated by the labels S1, S2, and S3 in the text below. A chin rest was used to help the listeners maintain a steady, aligned head position. Listeners' heads were positioned at speaker height and at the center of the semicircle described by the speaker array (see Fig. 1). Experiments were performed in total darkness.

G. Psychophysical methods

A noise background was present at all times during the experiments. The level of the noise and the number of contributing speakers varied from experiment to experiment. When all speakers are turned on, an incoherent noise field is produced that sounds something like wind rustling through a stand of trees. Two psychophysical procedures were used. These include the method of constant stimuli and two-interval-forced-choice (2IFC) staircases for threshold determinations. Details of these procedures are presented in the methods sections particular to each experiment.

III. EXPERIMENT 1 – SIMULTANEOUS CONTRAST

Does the perceived loudness of a central zone depend on the sound level of surrounding zones? Our results with a simultaneous contrast experiment suggest that central perceived loudness is independent of the levels of surrounding locations. The experiment provides no evidence in favor of spatially-opponent processes.

A. Methods

The seven speakers centered on azimuth zero deg provided the central zone for which judgments of perceived loudness were made (see Figure 4). These central speakers spanned the range of azimuths -26 to 26 deg. The levels of surrounding speakers to left and right were modulated sinusoidally at 100% contrast and at a rate of 0.5 Hz (see Figure 5). What is the effect of modulating surround level on the perceived loudness of the stimulus presented by the central speakers?

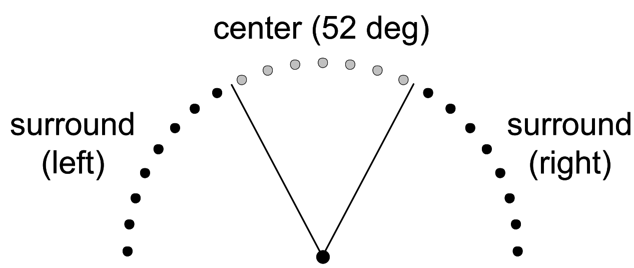


Figure 4. Surround speakers to left and right presented noise waveforms at a level modulated sinusoidally in time. The effect of this surround modulation on the perceived loudness of the zone defined by the seven central speakers was measured. The figure for the center's width, 52 deg, includes six interspeaker intervals and the width in degrees of a speaker cone.

When holding central speaker levels physically constant while modulating the surround, one might expect listeners to perceive that the physically unchanging center is, in fact, changing in loudness (see Fig. 5). If there is a contrast effect, then listeners would perceive center loudness as changing in counterphase to the surround modulation. The center would be perceived to have lesser loudness when the surround is at a higher level and to have greater loudness when the surround is at a lower level. Such a perceived modulation would be induced by the physical modulation of the surround and would be illusory. There could also be an assimilation effect, of opposite sign, that causes center loudness to change in phase with the surround modulation (*e.g.*, Helson, 1963; Hong & Shevell, 2004).

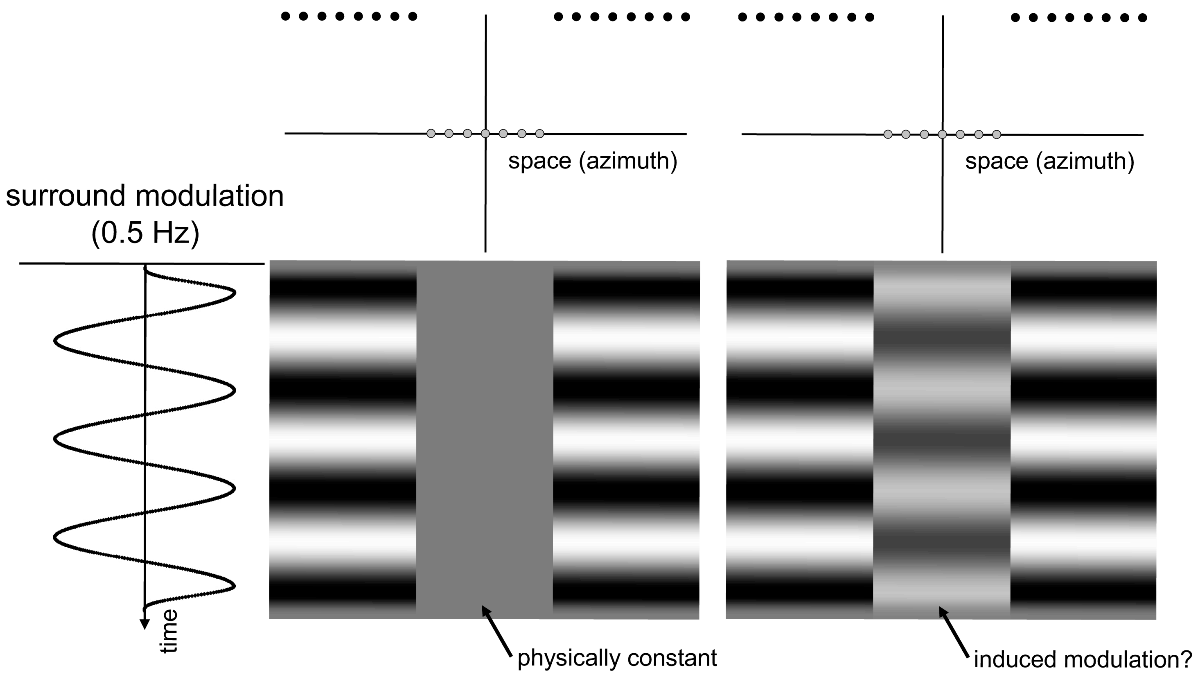


Figure 5. Space-time stimulus diagrams for a simultaneous contrast trial in which the central level is held constant and the surround is modulated sinusoidally at a rate of 0.5 Hz (left). Left diagram: physically constant level in the central zone defined by the central seven speakers with sinusoidal modulation of level for surrounding speakers at left and right. Right diagram: were simultaneous contrast operative in spatial hearing, one would expect the surround modulation to induce a counterphase modulation of perceived central loudness (indicated). Such an induced modulation may be nulled—so that the center is perceived to have a constant loudness—by adding a countervailing physical modulation to the center.

Casual observation suggests that there is little, if any, induced modulation. We used a nulling technique to measure more definitively the strength of induced modulation (McCourt, 1982; Krauskopf *et al*., 1986). With this technique, a nulling modulation is added physically to the center level to try to cancel or null the induced modulation. If the effects of the nulling modulation are equal and opposite to those of the induced modulation, then the central zone should be perceived as having a constant loudness.

The nulling modulation used was a sinusoidal function of frequency identical to that of the surround modulation (0.5 Hz), and was presented either in phase with the surround modulation (a modulation defined to be of positive contrast), or in counterphase (of negative contrast). One measures the contrast of the nulling modulation required for the center be perceived as having a steady loudness, modulating neither "in phase" nor "out of phase" with the surround modulation. If there is a significant induced modulation, one would want to take the further step of measuring both the amplitude and the phase of the nulling modulation (Singer & D'Zmura, 1994), but the results suggest that this further step is not required. Note that a simultaneous contrast effect would induce modulation out of phase with that of the surround, so that an in-phase modulation of positive contrast would be required to null the perceived modulation. An assimilation effect would induce modulation of the center in phase with that of the surround, so that a counterphase modulation (of negative contrast) would be needed to null the perceived modulation.

The method of constant stimuli was used to estimate the center modulation contrast needed to give rise to 50% "in phase" judgments and 50% "out of phase" judgments. If the nulling contrast that produces 50% "in phase" judgments and 50% "out of phase" judgments is positive, then there is evidence for simultaneous contrast. If this nulling contrast is negative-valued, then there would be evidence for assimilation. Finally, if the nulling contrast that produces 50% "in phase" judgments and 50% "out of phase" judgments is of value zero, then the experiment provides no evidence for either simultaneous contrast or assimilation.

Pilot work was used to establish the width of the center, which comprised the central seven speakers and subtended about 52 deg. Center width was chosen to balance (1) the audibility of central modulations, which increases as the number of speakers in the central zone is increased, and (2) the perceived strength of the surround modulation, which increases as the number of speakers assigned to the surround increases. Results presented below suggest that listener sensitivity to modulations of the seven-speaker center is quite high, despite the modulation in level from the 16 surround speakers between 62.4 dB and a nominal 0 dB. Level measurements with speakers set to a gain of zero were never less than about 25 dB.

Each listener started with the largest possible nulling contrast range, [-1,1], to help learn the task. With a nulling contrast of 1, the center is readily perceived as modulating in phase with the surround, while with a nulling contrast of -1, the center is readily perceived as modulating out of phase. Each experimental run presented a number of nulling contrasts, within such a range, ten times apiece in a block-randomized fashion.

The nulling contrast range was narrowed progressively as a listener grew more sensitive. Initially, each trial was of duration seven sec, as shown in Fig. 5. Three complete cycles of the 0.5 Hz sinusoidal modulation of surround level were presented (six sec), and raised-cosine functions of half-cycle duration 500 msec ramped the sinusoidal modulation on and off (0.5 sec apiece). This trial duration was decreased after several sessions of practice to five sec: two complete cycles of the 0.5 Hz modulation with raised-cosine flanks of half-cycle duration 500 msec. The listener responded after each presentation interval with a keypress to indicate whether the center modulated in phase or out of phase with the surround. The unmodulated noise background of 64 dB was present during the intertrial intervals of duration one sec.

When listener sensitivity plateaued, psychometric functions were measured using a fixed nulling contrast range that was centered approximately on the nulling contrast which produced 50% “in phase” judgments. Each run included nine (S1), 13 (S2) or 11 (S3) nulling contrast values. The data were fit, for each observer, by a Weibull function of form

$$F(c) = 1 - \exp\left(-\left((c-\gamma)/\lambda\right)^{k}\right), \tag{10}$$

which relates the fraction F of in-phase judgments to nulling modulation contrast c, position γ, scale λ and shape k parameters.

B. Results

Results for three listeners (see Figure 6) provide no evidence for either simultaneous contrast or assimilation under the tested circumstances.

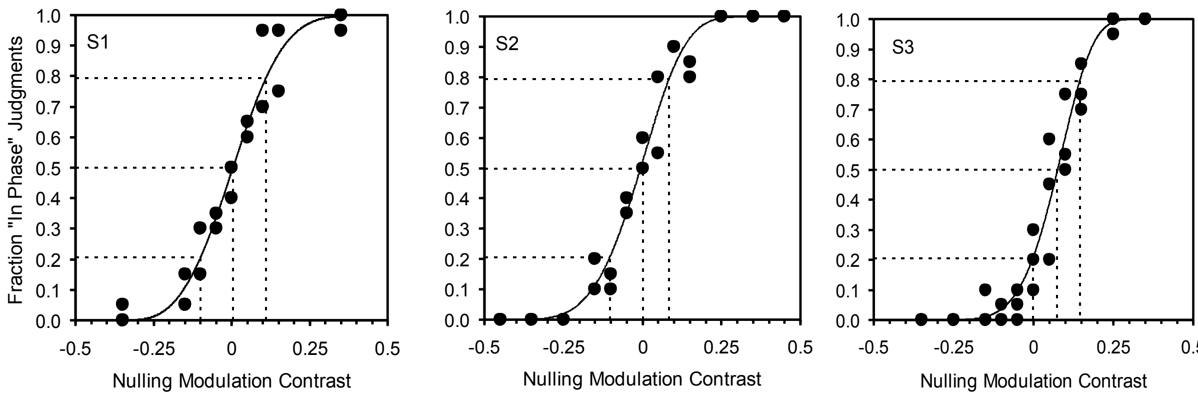


Figure 6. Results of a simultaneous contrast experiment for three listeners S1 (left), S2 (center) and S3 (right). The fraction of "in phase" judgments is plotted as a function of nulling modulation contrast, which varies along the horizontal axis from physically out of phase (negative-valued) through physically in phase (positive-valued). Each data point represents the result of twenty trials for a particular nulling modulation contrast (40 total trials per nulling modulation contrast for S1 and S2; 60 trials for S3). Weibull functions are fit to the data (solid curves) and account for 0.996, 0.987 and 0.997 of the variance for listeners S1, S2 and S3, respectively. Dotted lines indicate the nulling modulation contrasts that correspond to 20.6%, 50% and 79.4% "in phase" judgments.

Nulling modulation contrasts for 50% "in phase" judgments, estimated using the Weibull function fits, are 0.0045, -0.005, and 0.075 for listeners S1, S2, and S3, respectively. One way to gauge the significance of these values' departure from zero is by using the Weibull function fits to determine intervals of uncertainty. We define these intervals in terms of the change in nulling modulation contrast required to increase the fraction of "in phase" judgments from 50% to 79.4% (as with the three-down-one-up staircases used in the second experiment; Levitt, 1971) or to decrease the fraction from 50% to 20.6%. As is clear from the data (see Fig. 6), the nulling modulation contrast of zero lies squarely in the center of the interval of uncertainty for listeners S1 and S2. The nulling modulation contrast of zero lies just within the interval of uncertainty for listener S3; the results for listener S3 suggest that counterphase modulation of volume at a nulling modulation contrast of zero is detected about 80% of the time.

By comparing the peak level presented by the central seven speakers that gives rise to 50% "in phase" judgments to the peak level presented that gives rise to 79.4% "in phase" judgments, one may generate an estimate of the change in level required for the center modulation to be audible. Again relying on the Weibull function fits, one finds for S1 the contrasts 0.0045 and 0.11 for the two center contrasts that correspond to 50% and 79.4% "in phase" judgments, respectively. These figures are -0.005 and 0.085 for S2, respectively, and 0.075 and 0.145 for S3, respectively. Equations 5-9 may then be used to find the corresponding levels in dB and, by finding the difference in level for each listener, one estimates the change in level required for the center modulation to be audible. These estimated changes in level are 0.87 dB, 0.75 dB and 0.55 dB for listeners S1, S2 and S3, respectively, and are found at a space-time average level of 64 dB, considering all 23 speakers, or of 58.9 dB, considering just the central seven speakers. The values of the estimated changes in level suggest that the listeners are very sensitive to sound pressure level modulations of the seven-speaker center, despite the simultaneous modulation of the surround between levels 62.4 dB and a nominal 0 dB. Part of this sensitivity may be due to the fact that center stimuli are presented as modulations about a steady background, present at all times; that the stimulus has a long time-course (five sec) and is presented at a low temporal frequency presumably contribute also.

While there is a hint of simultaneous contrast in the results of listener S3 (see Fig. 6), any induction perceived by this listener at a nulling modulation contrast of zero is reported only about 80% of the time, which suggests that the perceived change in center loudness is minimal. The results are very clear for listeners S1 and S2: no hint whatsoever of simultaneous contrast (or of assimilation). The absence of simultaneous contrast for spatial pattern hearing when modulating incoherent noise fields suggests that mechanisms which mediate judgments of perceived loudness lack spatial opponency. Yet the null result of the present experiment provides indirect evidence for this suggestion at best. Convergent evidence for this suggestion is provided by the next experiment, which uses more objective psychophysical methods to measure the spatial frequency contrast sensitivity function for amplitude-modulated incoherent noise field stimuli.

IV. EXPERIMENT 2 – SPATIAL FREQUENCY CONTRAST SENSITIVITY

We measured spatial frequency contrast sensitivity functions in this second experiment to learn more about spatial pattern detection mechanisms. Spatial frequency contrast sensitivity functions are informative in at least two ways. First, the highest detectable spatial frequency provides information about the receptive field size of the smallest contributing detection mechanism. Second, if the spatial frequency contrast sensitivity function has a bandpass filter characteristic rather than a lowpass one, then there is evidence for spatial opponency. Results suggest that detection mechanisms are broad and non-opponent.

A. Methods

We used spatial Gabor functions, each composed of a spatially-sinusoidal modulation in cosine phase that was windowed spatially by a Von Mises function with shape parameter set to two. The Von Mises function window localizes the sinusoidal modulation while minimizing the spread of energy about the central frequency. The formula for the Von Mises function of space was shown earlier in Eqn. 2; multiplying the right-hand side of Eqn. 2 by $\cos\left(2\pi f \varphi_i\right)$, in which f refers to frequency, provides the formula for the Gabor functions used here.

The Gabor functions are used to create spatially-patterned amplitude modulations of incoherent noise fields in the horizontal plane. Fig. 7 shows Gabor function stimuli with sinusoidal components of frequency 2, 4, 6 and 8 cycles per circle (cpc). Note that only half the number of cycles per circle could be presented; the 23 speakers were positioned to span only half a circle. Each spatial Gabor function stimulus is windowed temporally by a truncated Gaussian function of duration 2 sec and standard deviation 1/3 sec. Stimulus contrast is varied to determine pattern detection threshold for each spatial frequency.

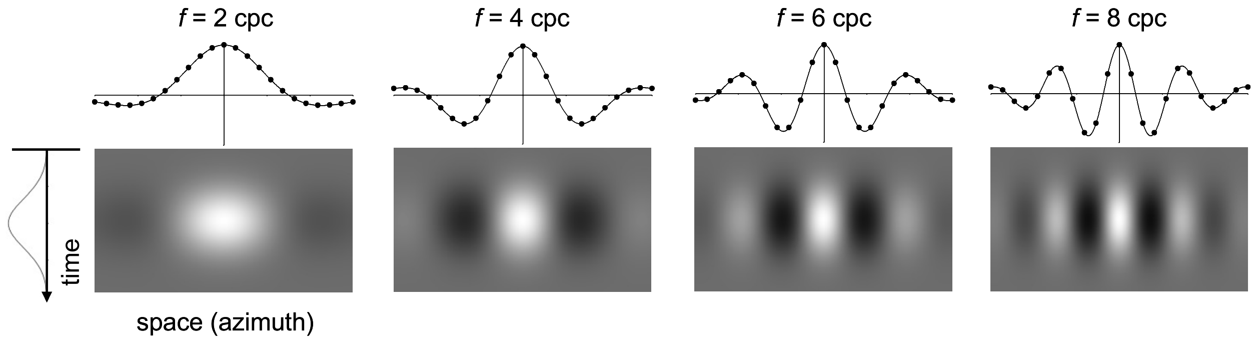


Figure 7. Space-time stimulus diagrams of the Gabor functions at spatial frequencies f = 2, 4, 6 and 8 cycles per circle (cpc) used to measure spatial frequency contrast sensitivity. These diagrams differ from those for the earlier experiments in that they depict modulations of intensity rather than of pressure. See text for details.

The space-time modulations used to measure spatial frequency contrast sensitivity are defined as modulations of intensity rather than pressure; the spatial Gabor functions and the Gaussian temporal windows modulate sound intensity directly. As a result, sound pressure and software gain are related to the modulation as the square root. In particular, the software gain $g[\varphi_i](t)$ applied to a single speaker with spatial index i at time t is given by (compare to Eqn. 4):

$$g[\varphi_i](t) = g_b\ (\ 1 + c\ m_{ST}[\varphi_i](t)\ )^{1/2}\ . \qquad (11)$$

Both the spatio-temporal modulation $m_{ST}[\varphi_i](t)$ and the contrast c have the range [-1,1]. If the background gain is set to $1/\sqrt{2}$, as in the experiment reported here, then the resulting software gain $g[\varphi_i](t)$ has the range [0,1].

The sound pressure level L_b from a single speaker, which corresponds to a background gain of $1/\sqrt{2}$ in this experiment, is 53.4 dB. Eqn. 5 provides the corresponding sound pressure P_b. Using this information, one can then determine the sound pressure from a single speaker for stimuli described by Eqn. 11:

$$P[\varphi_i](t) = P_b \left(1 + c\, m_{ST}[\varphi_i](t)\right)^{1/2} . \tag{12}$$

The resulting sound intensity produced by the $N = 23$ speakers is thus given by:

$$I(t) = P_b^2 \sum_{i=1}^{N} \left(\left(1 + c\, m_{ST}[\varphi_i](t)\right)\right). \tag{13}$$

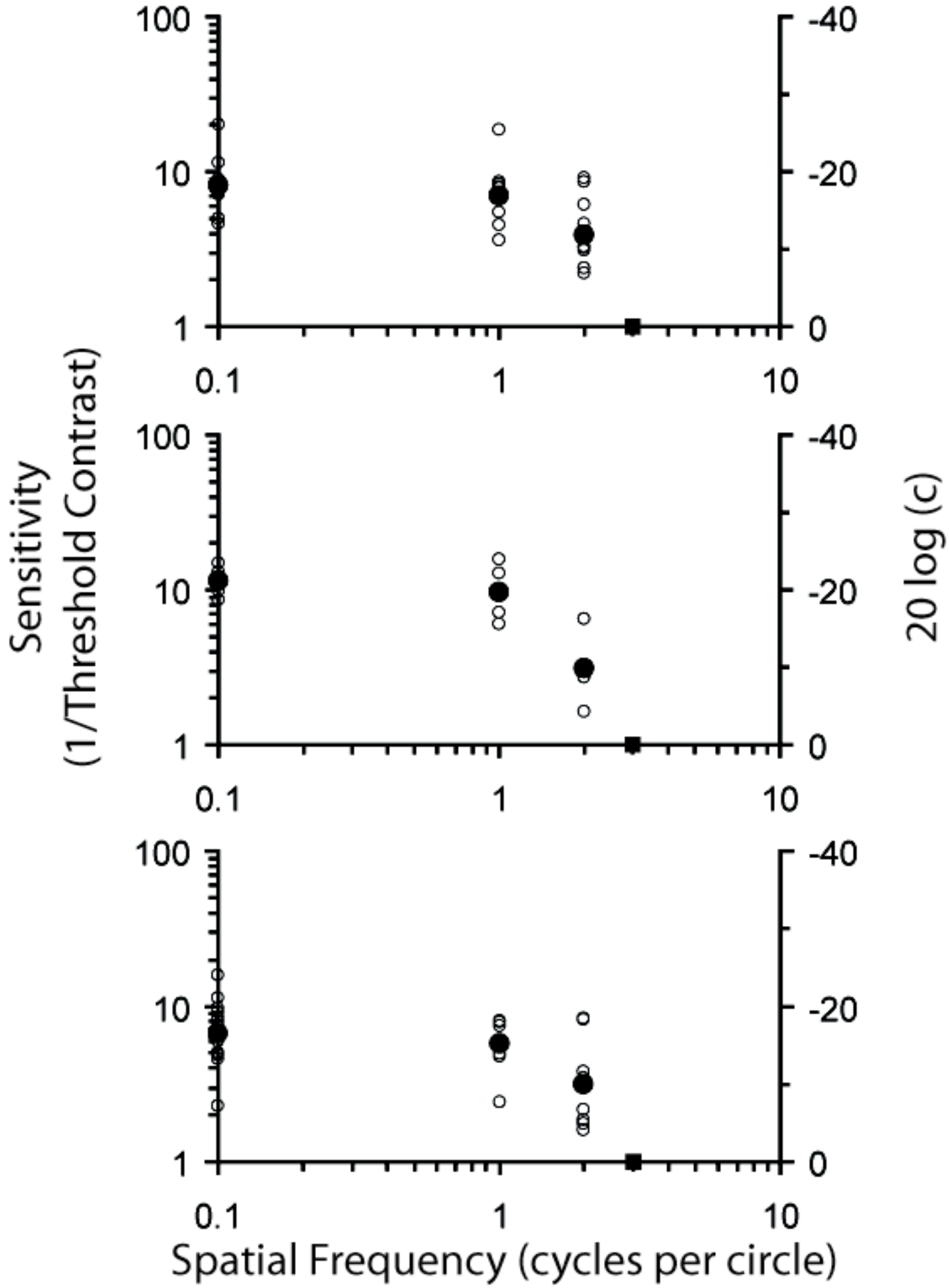
The sound pressure level from the 23 speakers, when each is held constant at their background level, is 67 dB. From Eqn. 13, one sees that if the sum of the modulations $m_{ST}[\varphi_i](t)$ across the spatial index i is equal to zero, which is the case for a sinusoid presented with an integral number of cycles, then the total sound intensity from all speakers is identical to that from the background. This "isolevel" property found when modulating intensity is desirable when measuring a spatial frequency contrast sensitivity function, just as it has proved useful when measuring temporal MTFs (*e.g.*, Viemeister, 1979). To measure spatial frequency contrast sensitivity in as frequency-specific manner as possible, one should avoid stimuli which are not isolevel, as these cause an overall change in sound pressure level, relative to the background. A listener can detect the overall level change of a pressure modulation stimulus—a change at spatial frequency zero—rather than its spatial pattern. This is not the case for isolevel intensity modulations.

Two interleaved three-down-one-up staircases, each of length 40 trials, were used to estimate contrast at a threshold level corresponding to a probability of correct detection of 0.794 (Levitt, 1971). The random interleaving of trials from two staircases was used to reduce the likelihood that a listener would be able to track staircase progress. Termination of a staircase after 40 trials produced sufficient numbers of turnarounds for threshold estimates to be generated. After running a practice pair of staircases, listeners ran either two (S2), three (S3) or four (S1) pairs of staircases per stimulus spatial frequency. The geometric means of the staircases' turnarounds are used to estimate thresholds.

Trials had two intervals. A signal of nonzero contrast was presented in one of these intervals; the signal interval was chosen randomly. The signal stimulus was presented at zero contrast during the other interval. Each interval was of duration 2 sec, while the period between the two intervals was 1 sec. A truncated Gaussian function of duration 2 sec and standard deviation 1/3 sec was used to window the stimulus. A very brief level increase was presented 1 sec prior to each interval to aid the listener in determining trial time-course. Listeners used a keyboard to signal in which interval they believed the signal was presented (left-hand keys, first interval; right-hand keys, second interval). Feedback was provided immediately after the response. One brief level increase indicated that the signal had been presented in the first interval, while two brief level increases indicated that the signal had been presented in the second. Each speaker presented noise at the background level 53.4 dB during inter-interval and inter-trial periods, so that the signal was presented as a modulation about a background level of 67 dB.

B. Results

Spatial frequency contrast sensitivity functions for the three listeners have a lowpass characteristic and a surprisingly low maximum resolvable spatial frequency.



100
10
1
0.1
1
10
-40
-20
0
Sensitivity
(1/Threshold Contrast)
20 log (c)
Spatial Frequency (cycles per circle)

Figure 8. Sensitivity to isolevel spatial Gabor function patterns decreases monotonically as spatial frequency increases. Small unfilled circles represent results of single staircases for listeners S1 (top), S2 (middle), and S3 (bottom). Filled disks represent the geometric means of individual staircase results. At spatial frequency three cycles per circle and higher, stimuli were inaudible at full strength (100% contrast). Results for the stimulus at spatial frequency zero, indicated by the stimulus diagram for $f = 0$ at top, are plotted along the logarithmic horizontal axis at a spatial frequency value of 0.1. All three listeners failed repeatedly to detect modulations of 100% contrast at 3 cycles per circle; these failures are indicated by filled squares placed along the abcissas at 3 cpc. Stimuli at higher frequencies were also undetectable.

Results for the three listeners are shown in Figure 8 (S1, top; S2, middle; S3, bottom). Sensitivity is defined to be the reciprocal of contrast at threshold: a higher sensitivity corresponds to a lower contrast at threshold. It is plotted on a logarithmic axis as a function of spatial frequency, also plotted logarithmically. Small, unfilled data points show results from single staircases, while filled data points show their geometric means. Sensitivity is greatest for each listener at a spatial frequency of zero cycles per circle and has a value close to ten (S1, 8.18; S2, 11.3; S3: 6.7). Threshold contrast is thus about 0.1. Sensitivity is slightly lower for all three listeners at spatial frequency one, lower still at spatial frequency two, and is zero at spatial frequency three and beyond. Spatial patterns at a spatial frequency of three cpc or higher are inaudible when presented at 100% contrast under the conditions studied here. The maximum resolvable spatial frequency in these experiments is two cpc, while the spatial frequency contrast sensitivity function itself has a lowpass characteristic.

V. DISCUSSION

Two experiments on sensitivity to spatial amplitude modulation of incoherent noise fields have been presented. Their results, when taken together, suggest that mechanisms of spatial hearing are not designed to detect or enhance spatial discontinuities in sound level.

The first experiment used a nulling method to pursue the observation that the perceived loudness of a central sector in the horizontal plane does not depend on the level in surrounding areas. Psychometric functions for the amount of nulling modulation, required to offset any induced center modulation caused by surround physical modulation, show that no perceived modulation is induced. This is a null result. It differs substantially from that of analogous work in the visual modality, where the term simultaneous contrast is used to describe the dependence of perceived central gray level on surrounding light intensity. Simultaneous contrast has long been interpreted in the vision literature as a perceptual consequence of spatial edge detection and enhancement. The present result suggests that human spatial hearing is not organized to detect or enhance spatial edges in level.

The second experiment measured the contrast required for detection at threshold for spatial Gabor patterns that varied in spatial frequency. Gabor patterns with sinusoidal components at frequencies beyond two cycles per circle were inaudible at the maximum possible modulation depth. Thresholds measured at zero, one and two cycles per circle increase monotonically, indicating that pattern sensitivity decreases monotonically with increasing spatial frequency. The lowpass characteristic of the measured spatial frequency contrast sensitivity functions suggests that spatial sensitivities of the mechanisms that mediate detection are not spatially opponent. This suggestion agrees with receptive field measurements in a variety of mammalian species.

The configuration of the employed speaker system limits the precision with which sensitivity to spatial modulations of low frequency can be measured. The system spans half a circle rather than a whole one. Yet the number of speakers, 23, is more than adequate to the task. One may use the cutoff frequency measurements from the spatial frequency contrast sensitivity experiments to deduce the corresponding Nyquist sampling rate (Nyquist, 1928/2002). The speakers must sample azimuthal variation in the horizontal plane at a rate equal to or greater than the number of peaks and troughs in the sinusoid at the highest frequency to which one is sensitive. This sampling rate has twice the value of the cutoff frequency. If we assume that this cutoff frequency lies between two and four cycles per circle, then between four to eight speakers are required to reproduce audible spatial patterns of the sort investigated here.

Our results for spatial frequency contrast sensitivity are limited in several further ways. First, measurements with Gabor functions using sine-phase sinusoidal components were not performed; only cosine-phase Gabor functions were used. Sensitivities to Gabor functions of identical frequency but of differing phase need not agree. Indeed, pairs of sensitivities obtained at single frequencies for Gabor functions with sine-phase and with cosine-phase sinusoidal components provide the information required to inverse Fourier transform the spatial frequency contrast sensitivity function results into the space domain. The inverse-transformed function, in the space domain, would be an estimate of detection mechanism spatial sensitivity at threshold. This style of analysis is most useful if the spatial hearing mechanisms recruited in such tasks behave in a linear fashion (obeying spatial pattern scaling and superposition rules), an assumption for which there is no evidence at this time.

Second, one fully expects spatial frequency contrast sensitivity to vary as a function of pattern position in the horizontal plane. Our stimuli were always centered on the front-and-center direction at azimuth zero deg. Were spatial frequency contrast sensitivity to vary as a function of pattern position in the horizontal plane, then the simplest possible model of spatial pattern sensitivity would be a space-varying (rather than a space-invariant) linear system.

Third, the spatial frequency contrast sensitivity function measurements made here are for amplitude modulations of an incoherent noise field. It may be the case that spatial pattern sensitivity to modulations of incoherent noise fields differs significantly from sensitivity to modulations of more nearly coherent fields. Furthermore, it may be the case that spatial frequency contrast sensitivity measurements made for patterns that vary in frequency across space, namely FM spatial patterns, differ substantially than the AM patterns studied here.

Fourth, the experiments manipulate only level differences. Any conclusions concerning the properties of spatial hearing mechanisms are thus limited to the properties of spatial hearing mechanisms sensitive to level differences. Further mechanisms, like those sensitive to interaural timing differences or to spectral differences, may well have spatial properties wholly different from those suggested here.

One psychophysical method likely to prove useful in addressing a number of these issues is spatial profile analysis, which is the spatial analog of profile analysis methods currently in use in the study of hearing (Richards *et al*., 1989; Berg & Green, 1990). Introduced by Ahumada and Lovell (1971), these methods have been used in the study of spatial vision quite extensively (*e.g.*, Ahumada, 2002). The idea is to study the effects of externally-added space-varying noise on spatial pattern signal detectability. Noise patterns that tend to cause false alarms can be used to generate a spatial profile of the detection mechanisms: a spatial perceptive field analogous to a neuron's spatial sensitivity profile. One would expect spatial profile analysis experiments for Gabor pattern detection to produce spatially very extensive perceptive fields, in line with the low cutoff frequency measured here. Of course, spatially broad sensitivities are not at variance with results on auditory localization blur, which suggest that one can discriminate source position differences on the order of a single degree (*rev*. Blauert, 1997). It was Helmholtz (1891), working in color vision, who showed that fine wavelength discrimination across the visible spectrum is possible using three broadly-tuned filters; the trick is to look at differences in their signals in regions of sensitivity overlap (Middlebrooks *et al*., 1994, 1998; Jenison, 1998; Boehnke & Phillips, 1999).

ACKNOWLEDGMENTS

We thank Dr. K. Saberi for his kind loan of the Brüel & Kjær sound meter used in calibration. We thank also Dr. J. Hagedorn for his work in early phases of this project. This work was supported by UCI CORCLR SIIG-2006-07-7 and by ARO 54228-LS-MUR.

REFERENCES

Ahumada, A.J., Jr. (**2002**). "Classification image weights and internal noise level estimation." J. Vision **2**, 121-131.

Ahumada, A. J., Jr., and Lovell, J. (**1971**). "Stimulus features in signal detection." J. Acoust. Soc. Am., **49**, 1751-1756.

Berg, B.G. and Green, D.M. (**1990**). "Spectral weights in profile listening." J. Acoust. Soc. Am. **88**, 758-766.

Blauert, J. (**1997**). *Spatial hearing. The psychophysics of human sound localization*. *Rev. Ed.,* Allen, J. S., Trans. (Cambridge: MIT).

Boehnke, S. E., and Phillips, D. P. (**1999**). "Azimuthal tuning of human perceptual channels for sound location." J. Acoust. Soc. Am. **106**, 1948-1955.

Brugge, J.F., Reale, R.A., Hind, J.E., Chan, J.C.K, Musicant, A.D., and Poon, P.W.F. (**1994**). "Simulation of free-field sound sources and its application to studies of cortical mechanisms of sound localization in the cat." Hearing Research **73**, 67-84.

Carlile, S., Hyams, S., and Delaney, S. (**2001**). "Systematic distortions of auditory space perception following prolonged exposure to broadband noise." J. Acoust. Soc. Am. **110**, 416-424.

Enroth-Cugell, C., and Robson, J.G. (**1966**). "The contrast sensitivity of retinal ganglion cells of the cat." J. Physiology (London) **187**, 517-552.

Helmholtz, H. von (**1891**). "Versuch einer erweiterten Anwendung des Fechnerschen Gesetzes im Farbensystem." ("A broader application of Fechner's Law to Color Vision.") Zeitschrift fuer Psychologie und Physiologie der Sinnesorgane **2**, 1-30.

Helmholtz, H. von (**1896**). *Handbuch der physiologischen Optik (Handbook of physiological Optics)*, 2nd Edition (Hamburg: Voss).

Helson, H. (**1963**). "Studies of anomalous contrast and assimilation." J. Opt. Soc. Am. **53**, 179-184.

Hong, S.W., and Shevell, S.K. (**2004**). "Brightness contrast and assimilation from patterned inducing backgrounds." Vision Research **44**, 35-43.

Hubel, D.H., and Wiesel, T.N. (**1962**). "Receptive fields, binocular interaction and functional architecture in the cat's visual cortex." J. Physiology **160**, 106-154.

Jenison, R.L. (**1998**). "Models of sound location acuity with broad cortical spatial receptive fields." In *Central Auditory Processing and Neural Modeling*, Poon, P. & Brugge, J., Eds. (New York: Plenum), pp. 161-174.

Jenison, R.L., Reale, R.A., Hind, J.E., and Brugge, J.F. (**1998**). "Modeling of auditory spatial receptive fields with spherical approximation functions." J. Neurophysiology **80,** 2645-2656.

Jenison, R.L., Schnupp, J.W.H., Reale, R.A., and Brugge, J.F. (**2001**). "Auditory space-time receptive field dynamics revealed by spherical white-noise analysis." J. Neuroscience **21**, 4408-4415.

Katz, D. (**1935**). *The World of Colour*. MacLeod, R.B. & Fox, C.W., Trans. (London: Kegan Paul, Trench, Trubner & Co., Ltd.).

Krauskopf, J., Zaidi, Q., and Mandler, M.B. (**1986**). "Mechanisms of simultaneous color induction." J. Opt. Soc. Am. *A* **3**, 1752-1757.

Kelly, D.H. (**1979**). "Motion and vision. II. Stabilized spatio-temporal threshold surface." J. Opt. Soc. Am. **69**, 1340-1349.

Levitt, H. (**1971**). "Transformed up-down methods in psychoacoustics." J. Acoust. Soc. Am. **49**, 467-477.

McCourt, M.E. (**1982**). "A spatial frequency dependent grating-induction effect." Vision Research **22**, 119-134.

Middlebrooks, J.C., Clock, A., Xu, L., and Green, D. (**1994**). "A panoramic code for sound location by cortical neurons." Science **264**, 842-844.

Middlebrooks, J.C., Xu, L., Eddins, A.C., and Green, D.M. (**1998**). "Codes for sound-source location in nontonotopic auditory cortex." J. Neurophysiology **80**, 863-881.

Nyquist, H. (**1928**). "Certain topics in telegraph transmission theory." Trans. AIEE **47**, 617-644, reprinted IEEE ***90***, 280-305, 2002.

Ratliff, F. (**1965**). *Mach Bands: Quantitative Studies on Neural Networks in the Retina*. (San Francisco: Holden-Day).

Rauschecker, J.P., and Tian, B. (**2000**). "Mechanisms and streams for processing of "what" and "where" in auditory cortex." Proc. Natl. Acad. Sci. U.S.A. **97**, 11800-11806.

Recanzone, G.H., Guard, D.C., Phan, M.L., and Su, T.-I.K. (**2000**). "Correlation between the activity of single auditory cortical neurons and sound-localization behavior in the macaque monkey." J. Neurophysiology **83**, 2723-2739.

Richards, V.M., Onsan, Z.A., and Green, D.M. (**1989**). "Auditory profile analysis: potential pitch cues." Hearing Research **39**, 27-36.

Robson, J.G. (**1966**). "Spatial and temporal contrast-sensitivity functions of the visual system." J. Opt. Soc. Am. **56**, 1141-1142.

Sabine, W.C. (**1922**). *Collected Papers on Acoustics* (Harvard U.P., Cambridge).

Schade, O.H. (**1956**). "Optical and photoelectric analog of the eye." J. Opt. Soc. Am. **46**, 721-739.

Schlack, A., Sterbing-D'Zngelo, S.J., Hartung, K., Hoffmann, K.-P., and Bremmer, F. (**2005**). "Multisensory space representations in the macaque ventral intraparietal area." J. Neuroscience **25**, 4616-4625.

Schnupp, J.W.H., Mrsic-Flogel, T.D., and King, A.J. (**2001**). "Linear processing of spatial cues in primary auditory cortex." Nature **414**, 200-204.

Singer, B. , and D'Zmura, M. (**1994**). Color contrast induction. Vision Research **34**, 3111-3126.

Stecker, G.C., Harrington, I.A. , and Middlebrooks, J.C. (**2005**). "Location coding by opponent neural populations in the auditory cortex." PLoS Biology **3**, 520-528.

Stecker, G.C. & Middlebrooks, J.C. (**2003**). "Distributed coding of sound locations in the auditory cortex." Biological Cybernetics **89**, 341-349.

Tian, B., Reser, D., Durham, A., Kustov, A. , and Rauschecker, J.P. (**2001**). "Functional specialization in rhesus monkey auditory cortex." Science **292**, 290-293.

Viemeister, N.F. (**1979**). "Temporal modulation transfer functions based upon modulation thresholds." J. Acoust. Soc. Am. **66**, 1364-1380.

Woods, T.M., Lopez, S.E., Long, J.H., Rahman, J.E., and Recanzone, G.H. (**2006**). "Effects of stimulus azimuth and intensity on the single-neuron activity in the auditory cortex of the alert macaque monkey." J. Neurophysiology **96**, 3323-3337.