Neural Correlates of the Pitch of Complex Tones. II. Pitch Shift, Pitch Ambiguity, Phase Invariance, Pitch Circularity, Rate Pitch, and the Dominance Region for Pitch

PETER A. CARIANI AND BERTRAND DELGUTTE
Department of Otology and Laryngology, Harvard Medical School, and Eaton-Peabody Laboratory for Auditory Physiology, Massachusetts Eye and Ear Infirmary, Boston, Massachusetts 02114

SUMMARY AND CONCLUSIONS

1. The neural correlates of low pitches produced by complex tones were studied by analyzing temporal discharge patterns of auditory nerve fibers in Dial-anesthetized cats. In the previous paper it was observed that, for harmonic stimuli, the most frequent interspike interval present in the population of auditory nerve fibers always corresponded to the perceived pitch (predominant interval hypothesis). The fraction of these most frequent intervals relative to the total number of intervals qualitatively corresponded to strength (salience) of the low pitches that are heard.

2. This paper addresses the neural correlates of stimuli that produce more complex patterns of pitch judgments, such as shifts in pitch and multiple pitches. Correlates of pitch shift and pitch ambiguity were investigated with the use of harmonic and inharmonic amplitude-modulated (AM) tones varying either in carrier frequency or modulation frequency. Pitches estimated from the pooled interval distributions showed shifts corresponding to "the first effect of pitch shift" (de Boer's rule) that is observed psychophysically. Pooled interval distributions in response to inharmonic stimulus segments showed multiple maxima corresponding to the multiple pitches heard by human listeners (pitch ambiguity).

3. AM and quasi-frequency-modulated tones with low carrier frequencies produce very similar patterns of pitch judgments, despite great differences in their phase spectra and waveform envelopes. Pitches estimated from pooled interval distributions were remarkably similar for the two kinds of stimuli, consistent with the psychophysically observed phase invariance of pitches produced by sets of low-frequency components.

4. Trains of clicks having uniform and alternating polarities were used to investigate the relation between pitches associated with periodicity and those associated with click rate. For unipolar click trains, where periodicity and rate coincide, physiologically estimated pitches closely follow the fundamental period. This corresponds to the pitch at the fundamental frequency (F_0) that is heard. For alternating click trains, where rate and periodicity do not coincide, physiologically estimated pitches always closely followed the fundamental period. Although these pitch estimates corresponded to periodicity pitches that are heard for F_0s >150 Hz, they did not correspond to the rate pitches that are heard for F_0s <150 Hz. The predominant interval hypothesis thus failed to predict rate pitch.

5. When alternating-polarity click trains are high-pass filtered, rate pitches are strengthened and can also be heard at F_0s >150 Hz. Pitches for high-pass-filtered alternating click trains were estimated from pooled responses of fibers with characteristic frequencies (CFs) >2 kHz. Roughly equal numbers of intervals at 1/rate and $1/F_0$ were found for all F_0s studied, from 80 to 160 Hz, producing pitch estimates consistent with the rate pitches that are heard after high-pass filtering. The existence region for rate pitch also coincided with the presence of clear periodicities related to the click

rate in pooled peristimulus time histograms. These periodicities were strongest for ensembles of fibers with CFs >2 kHz, where there is widespread synchrony of discharges across many fibers.

6. The "dominance region for pitch" was studied with the use of two harmonic complexes consisting of harmonics 3–5 of one F_0 and harmonics 6–12 of another fundamental 20% higher in frequency. When the complexes were presented individually, pitch estimates were always close to the fundamental of the complex. When the complexes were presented concurrently, pitch estimates always followed the fundamental of harmonics 3–5 for F_0s of 150–480 Hz. For F_0s of 125–150 Hz, pitch estimates followed one or the other fundamental, and for F_0s <125 Hz, pitch estimates followed the fundamental of harmonics 6–12. The results are generally consistent with the ranges of component frequencies (500–1,000 Hz) that are found psychophysically to be most important for low pitch.

7. Taken as a whole, the physiological data presented here provide strong evidence that interspike interval information plays an important role in the perception of the low pitch of complex tones. The predominant interval hypothesis for pitch yields surprisingly robust, comprehensive, and unified explanations for a very wide range of pitch phenomena: the missing fundamental, pitch invariance with respect to level, pitch equivalence of spectrally diverse stimuli, the pitch of unresolved harmonics, the pitch of AM noise, pitch salience, pitch shift of inharmonic AM tones, pitch ambiguity, phase insensitivity of pitch, and the dominance region for pitch. Its main weaknesses are its failure to account for the rate pitches of alternating click trains and its underestimation of the salience of low-frequency tones.

INTRODUCTION

The "predominant interval" hypothesis for the low pitch of complex tones postulates that the pitch heard corresponds to the most frequent interval present in the population of auditory nerve fibers at any given time. Explicit models for pitch based on this hypothesis are able to readily account for a wide diversity of pitch phenomena (Licklider 1951; Meddis and Hewitt 1991a,b; Moore 1989; van Noorden 1982).

In the companion paper in this series, we tested the predominant interval hypothesis by examining pooled interspike interval distributions of ensembles of auditory nerve fibers. For many different harmonic complex tones, the most frequent interval in pooled interval distributions invariably corresponded to the period of the pitch that is heard. Peak-to-background ratios in pooled distributions qualitatively corresponded to the strength of the pitch that is heard. Those stimuli that evoke strong, salient pitches produced pooled

TABLE 1. *Stimulus parameters*

Stimulus	Variable	Range, Hz	Duration, ms	Level, dB SPL	Remarks
AM tone	F_m	80–160	480	60	$F_c = 640$ Hz, $m = 1.0$
		160–320			
		240–480			
	F_c	500–700	480	60	$F_m = 125$ Hz, $m = 1.0$
QFM tone	F_m	80–160	480	60	$F_c = 640$ Hz, $m = 1.0$
		160–320			
		240–480			
	F_c	500–700	480	60	$F_m = 125$ Hz, $m = 1.0$
Click trains	F_0	80–160	530.4	80	Uniform-polarity 40-μs clicks
		160–320			Uniform-polarity 40-μs clicks
		80–160			Alternating-polarity 40-μs clicks
Harmonic complexes	F_0	80–160	498	60 per component	Harmonics 3–5
		160–320			
		240–480			
	F_0	96–192	498	60 per component	Harmonics 6–12
		192–384			
		288–576			
Shepard-Risset	F_0	4.7–9.4	5,600	60	Octave harmonics (octaves 1–12)

AM, amplitude-modulated; F_m, modulation frequency; F_c, carrier frequency; m, modulation index; QFM, quasi-frequency-modulated; F_0, fundamental frequency.

distributions with high peak-to-background ratios, whereas those that produce weak pitches produced pooled distributions with low ratios. These psychoneural correspondences suggest that information subserving pitch perception could be represented in the interspike interval statistics of populations of auditory neurons.

In the present paper, we investigate the neural correlates of more complex pitch phenomena, using stimuli that produce multiple, competing pitches rather than one definite pitch. The resulting patterns of pitch judgments were studied by psychophysicists to probe auditory mechanisms underlying pitch perception: the pitch shifts of inharmonic complex tones (de Boer 1956, 1976; Schouten et al. 1962), pitch ambiguity, pitch circularity (de Boer 1956, 1976; Schouten et al. 1962; Shepard 1964), "rate pitch" (de Boer 1956, 1976; Flanagan and Gutman 1960a,b; Schouten et al. 1962; Shepard 1964), and the dominance region for pitch (de Boer 1956, 1976; Flanagan and Gutman 1960a,b; Plomp 1967; Ritsma 1967; Schouten et al. 1962; Shepard 1964). Stimuli having aperiodic waveforms and/or multiple, competing periodicities were used to study these effects: harmonic and inharmonic amplitude-modulated (AM) and quasi-frequency-modulated (QFM) tones, Shepard-Risset frequency-modulated tones, trains of alternating-polarity clicks, and pairs of harmonic complexes with competing fundamentals. The complex patterns of pitch judgments that these stimuli produce can be compared with pitch estimates from pooled interspike interval distributions of the auditory nerve to provide still stronger tests of the predominant interval hypothesis.

METHODS

The experimental and analytical methods used here are described in detail in the companion paper. Responses of single auditory nerve fibers were recorded in Dial-anesthetized cats with the use of glass micropipette electrodes. Stimuli were numerically generated (sampling rate 100 kHz, 16-bit resolution) and presented via a calibrated, closed acoustic assembly.

Basic stimulus parameters are given in Table 1. Most stimuli

were sinusoidally varied in a parameter associated with pitch [e.g., fundamental frequency (F_0), modulation frequency (F_m), or carrier frequency (F_c)], with one cycle of variation occurring over the stimulus duration (~500 ms). AM and QFM stimuli were varied in either F_m or F_c. For AM and QFM tones, the modulation index was always 1.0. For the Shepard-Risset stimulus, F_0 was linearly swept upward over 5.6 s. All stimuli were presented continuously, 40 times per fiber for the Shepard-Risset tones, 100 times for all others.

Distributions of all-order interspike intervals were analyzed by means of autocorrelation histograms and autocorrelograms (see companion paper). An autocorrelation histogram represents the distribution of all-order intervals, whereas an autocorrelogram represents the running distribution of all-order intervals. Thus the "cross section" of an autocorrelogram at a particular peristimulus time (PST) is an autocorrelation histogram (as in Fig. 3). Interval distributions from 30–90 single fibers were grouped into octave characteristic frequency (CF) bands, weighted to approximate a human distribution of CFs, and summed together to form pooled autocorrelation histograms and autocorrelograms that then served as estimates of the population interval distribution of the human auditory nerve.

Pitches were estimated from smoothed pooled autocorrelation histograms (300-ms rectangular window). The position of the highest interval peak was taken as the estimate of the pitch period. Except for the Shepard-Risset tones, when multiple, harmonically related major peaks of comparable height were present (e.g., 1/F_0, 2/F_0, 3/F_0), the position of the peak representing the shortest interval was chosen as the pitch estimate. Pitch saliences were estimated from peak-to-background ratios in smoothed, pooled autocorrelation histograms (see companion paper). Fiber saliences are peak-to-background ratios in smoothed autocorrelation histograms of single fibers.

RESULTS

Pitch shift and pitch ambiguity of inharmonic complex tones

AM tones with variable F_cs (500–750 Hz) and fixed F_ms (125 or 250 Hz) were used to investigate the neural correlates of pitch shift. F_c varied sinusoidally from 500 to 750

Stimulus waveform

Stimulus autocorrelation
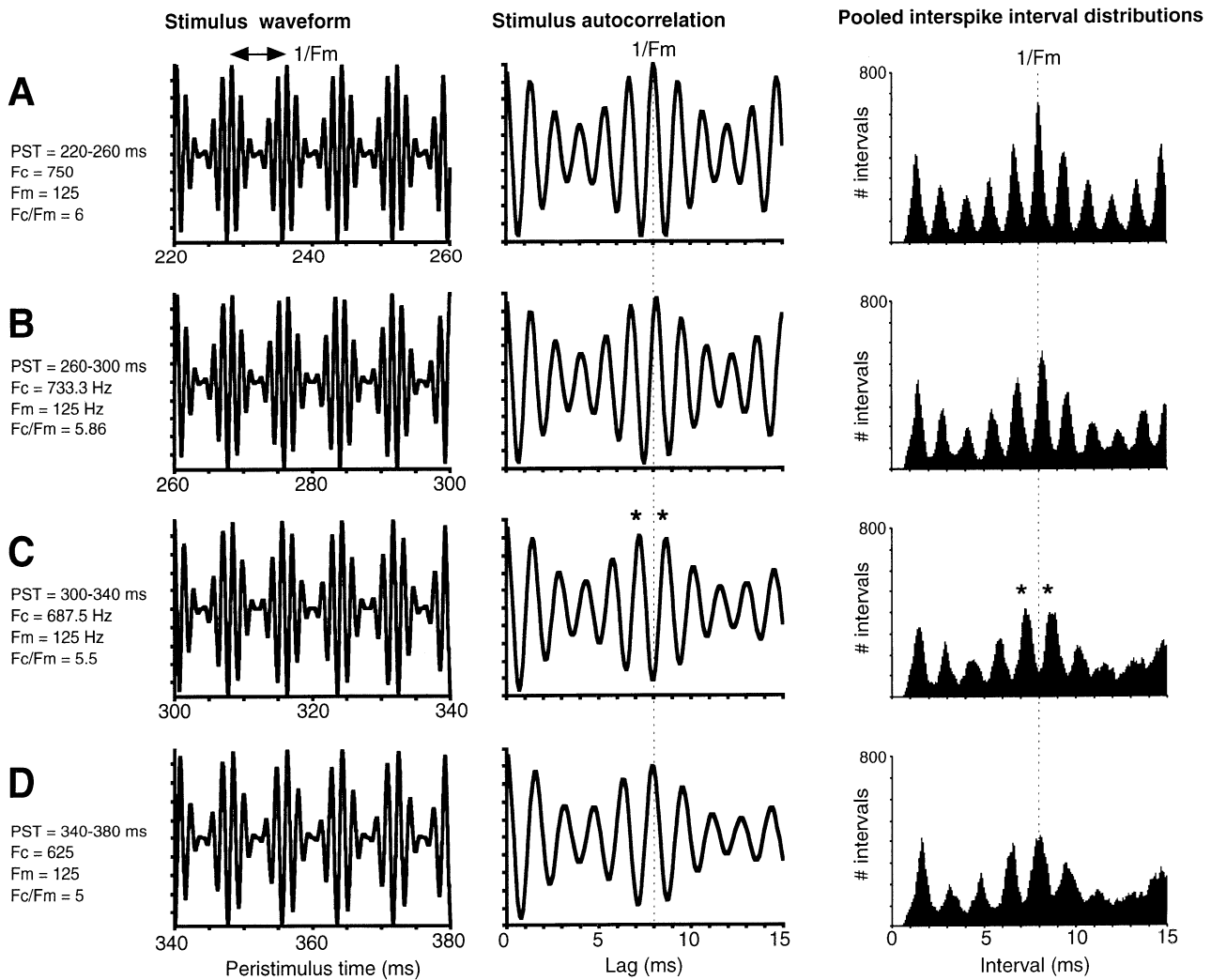
Pooled interspike interval distributions

FIG. 1. Interspike interval correlates of the pitch shift of inharmonic complex tones. *Right*: stimulus waveforms, amplitude-modulated (AM) tones with fixed modulation frequency ($F_m = 125$ Hz) and different carrier frequencies ($F_c = 625-750$ Hz). *A* and *D* are harmonic, whereas *B* and *C* are inharmonic. Peristimulus times (PSTs) listed correspond to those of Fig. 2*A*. *Middle*: stimulus short-term autocorrelation functions. Vertical dotted lines: modulation period $1/F_m$. *Left*: pooled all-order interval (autocorrelation) histograms for 31 auditory nerve fibers. Asterisks: multiple, ambiguous pitches that would be heard.

to 500 Hz over the 480-ms duration of the stimulus. Waveforms for four consecutive segments of this stimulus are shown in Fig. 1, *left*. The interval in the waveform indicated by the arrows is the modulation period $1/F_m$ (8 ms). Autocorrelation functions for each waveform segment are shown in Fig. 1, *middle*. Here the vertical dotted line indicates the modulation period $1/F_m$.

When F_c/F_m is an integer (Fig. 1, *A* and *D*), the stimulus is harmonic and the waveform is periodic. Here F_0 is equal to F_m, and the waveform has a period of $1/F_m$. For such harmonic AM tones, human listeners usually hear a clear, unambiguous pitch at F_m ($F_m = F_0$).

When F_c/F_m is not an integer (Fig. 1, *B* and *C*), the stimulus is inharmonic and aperiodic. Although the envelope of the waveform still has a periodicity of F_m, the fine structure of the waveform contains periodicities that deviate slightly from F_m. For these inharmonic AM tones, human listeners hear pitches at fine-structure periodicities rather than at F_m. This deviation from F_m is the "first

effect of pitch shift," and is described by de Boer's rule (de Boer 1956, 1976). Stated in slightly different form, it is $p = F_c/n$, where p is the frequency of the pitch heard and n is an integer near F_c/F_m. The effect can also be described in terms of stimulus autocorrelation functions, where "pseudoperiods" correspond to maxima near $1/F_m$ (de Boer 1956). Accompanying the shift in pitch is pitch ambiguity, wherein human listeners hear one of two or more pitches in the vicinity of, but not at, F_m (de Boer 1956, 1976; Schouten et al. 1962; Wiersinga-Post and Duifhuis 1995). Pitch ambiguity reaches a maximum when F_c/F_m is halfway between two integers, and the pitch alternatives for this condition are indicated by asterisks in Fig. 1*C*.

Pooled all-order interspike interval distributions (pooled autocorrelation histograms) for each of the stimulus segments are shown in Fig. 1, *right*. Most striking is the resemblance of each interval distribution to the autocorrelation function of the corresponding stimulus segment. Where a
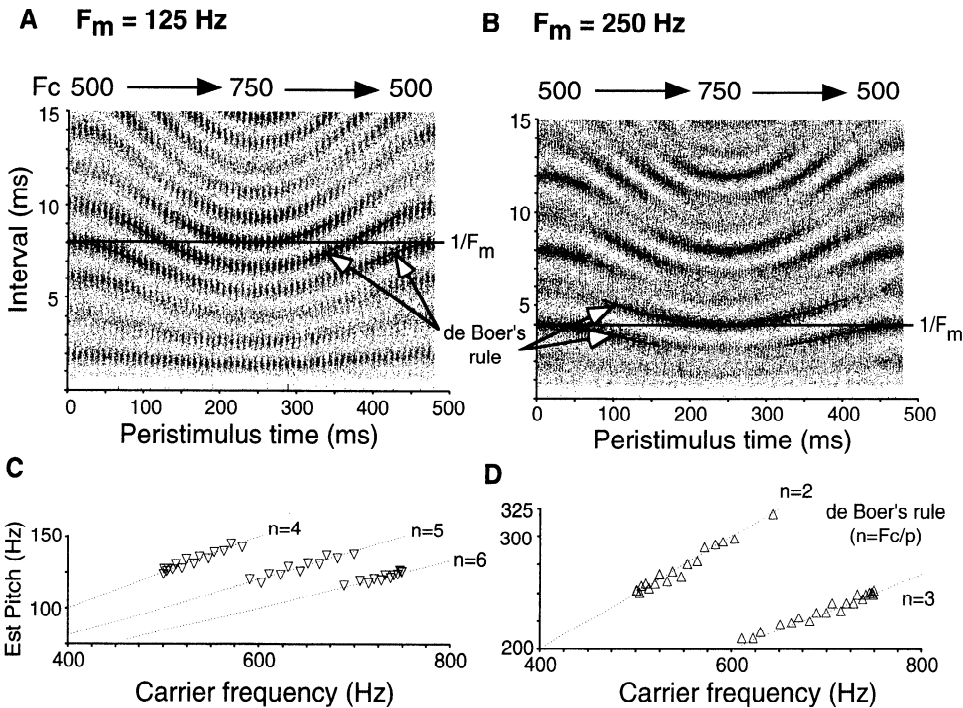
Variable carrier AM Tones Fc = 500-750 Hz

FIG. 2. Neural correlates for pitch shift for variable-carrier AM tones. $F_c = 500-700$ Hz, $F_m = 125$ or 250 Hz. A and B: pooled response autocorrelograms (47 and 53 fibers, respectively). Modulation period $1/F_m$ (straight line) and pitch(es) heard (de Boer's rule, curved lines) are indicated. C and D: physiologically estimated pitches as a function of F_c. Diagonal lines: de Boer's rule.

clear, unambiguous pitch is heard at F_m, in the harmonic segments shown in Fig. 1, A and D, the highest interval peaks are located at $1/F_m$. Where a small pitch shift relative to F_m would be expected, as in Fig. 1B, the interval peak also shows a small deviation from $1/F_m$. Where two ambiguous pitches would be expected, as in the inharmonic segment shown in Fig. 1C, two roughly equal interval peaks appear in the histogram.

Running pooled interspike interval distributions (pooled autocorrelograms) in response to two variable-carrier AM tones ($F_c = 500-750$ Hz, $F_m = 125$ Hz or 250 Hz) are shown in Fig. 2. The modulation period $1/F_m$ is indicated by horizontal, straight lines, whereas curved black lines indicate the pitch(es) that would be heard (de Boer's rule). The two sets of lines converge when the stimulus is harmonic and one pitch is heard. The two sets diverge when the stimulus is inharmonic and one of several pitches can be heard. For inharmonic AM tones, the densest interval bands in the autocorrelograms always coincide with the pitches that are heard (de Boer's rule) rather than the modulation period $1/F_m$.

Pitches estimated from the autocorrelograms showed similar patterns of pitch shift and pitch ambiguity. Pitches estimated from successive 20-ms peristimulus segments are plotted against their respective F_cs in Fig. 2, C and D. Predictions of de Boer's rule are indicated by diagonal lines. All pitch estimates fall near those lines, in close agreement with de Boer's rule. In inharmonic regions where pitch ambiguity is expected (e.g., $F_c \sim 575$ Hz and $F_c \sim 675$ Hz in Fig. 2C; $F_c \sim 625$ Hz in Fig. 2D), the pitch estimation algorithm finds multiple pitches.

The pitch shift experiment was also conducted using AM tones in which the modulator, F_m, was varied while the carrier, F_c, was kept constant (Fig. 3). Here a different

progression of pitches is heard than for the variable-carrier AM tones of Figs. 1 and 2. Instead of continuously shifting pitch values, ambiguities between several discrete pitches wax and wane as probabilities of making particular pitch judgments shift among them. The discrete pitches predicted by de Boer's rule are indicated by sets of straight horizontal lines in the pooled autocorrelograms (Fig. 3, C and D), whereas the continuous, undulating lines show modulation periods $1/F_m$. The stimulus segment pictured in Fig. 3E is a harmonic region ($F_m = 320$ Hz, $F_c/F_m = 2.0$) where one pitch at $1/F_m$ is heard. The pooled autocorrelation histogram for the segment (Fig. 3E) has a maximum very close (+0.5%) to the modulation period. The stimulus segment pictured in Fig. 3G is an inharmonic region ($F_m = 256$ Hz, $F_c/F_m = 2.5$) where maximal pitch ambiguity is expected. Here two interval bands surrounding $1/F_m$ appear to be of roughly equal density. These bands correspond to the two peaks that are indicated by arrows in the corresponding pooled autocorrelation histogram (Fig. 3G). These peaks both lie near, but not at, $1/F_m$ (dashed line), and correspond to 213 Hz ($F_c/3$) and 320 Hz ($F_c/2$).

In regions of maximal pitch ambiguity, when $F_c/F_m = n \pm 0.5$, AM tones have a true F_0 an octave below the F_m. Such "octave drops" are often reported by human listeners if they are allowed to choose pitch matches near the true fundamental (Gerson and Goldstein 1978). The partials present for the stimulus segment shown in Fig. 3G ($F_c - F_m = 384$ Hz, $F_c = 640$ Hz, and $F_c + F_m = 896$ Hz) are odd harmonics 3, 5, and 7 of a fundamental at 128 Hz, which is an octave below F_m (256 Hz). The octave drop pitch can be heard at this true fundamental if listeners consider pitches far from $1/F_m$, and correspondingly, the highest interval peak lies at this true fundamental.
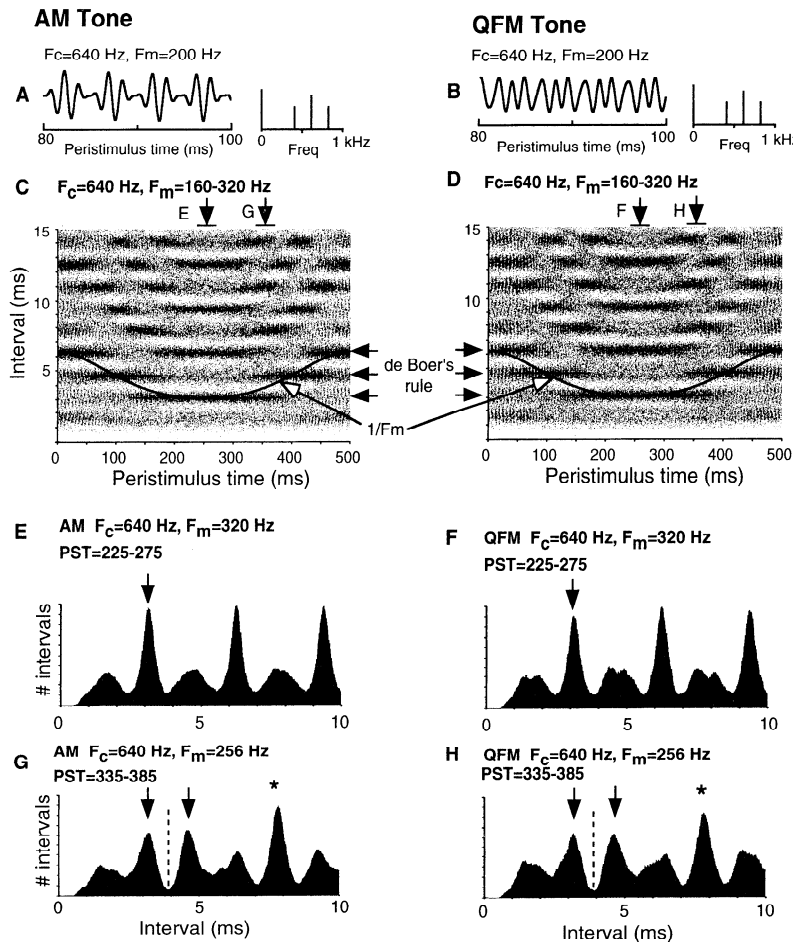
AM Tone

A Fc=640 Hz, Fm=200 Hz

QFM Tone

B Fc=640 Hz, Fm=200 Hz

C F_c=640 Hz, F_m=160-320 Hz

E ↓ G ↓

de Boer's rule

1/Fm

D Fc=640 Hz, F_m=160-320 Hz

F ↓ H ↓

E AM F_c=640 Hz, F_m=320 Hz
PST=225-275

F QFM F_c=640 Hz, F_m=320 Hz
PST=225-275

G AM F_c=640 Hz, F_m=256 Hz
PST=335-385

Interval (ms)

H QFM F_c=640 Hz, F_m=256 Hz
PST=335-385

Interval (ms)

FIG. 3. Neural responses to variable-modulator AM and quasi-frequency-modulated (QFM) tones. *A* and *B*: waveforms and power spectra for AM and QFM tones when F_m = 200 Hz. *C* and *D*: pooled response autocorrelograms. AM tone: 52 fibers; QFM tone: 44 fibers. Pitches heard correspond to de Boer's rule (horizontal lines indicated by triple arrows). *E–H*: pooled interval histograms for 50-ms stimulus segments. *E* and *F*: harmonic case (F_c/F_m = 2) producing 1 clear pitch at $1/F_m$ (arrow). *G* and *H*: inharmonic case (F_c/F_m = 2.5) producing ambiguous pitches (dual arrows) near, but not at, $1/F_m$. Asterisks: lower "octave drop" pitch at $1/$fundamental frequency (F_0). Dashed lines: $1/F_m$.

Insensitivity of pitch to changes in phase spectrum and waveform envelope

To investigate the sensitivity of population interval distributions to changes in phase spectra, the pooled neural responses of AM tones were compared to those of their QFM counterparts (Fig. 3). QFM tones had the same carrier and variable F_ms as the variable-modulator AM tones (F_c = 640 Hz, F_m = 160–320 Hz). These AM and QFM tones thus differ only in their phase spectra, specifically in the phase of their F_c. Although their waveform envelopes are quite different (Fig. 3, *A* vs. *B*), these stimuli are subjectively very hard to tell apart.

Neural responses for the QFM tones (Fig. 3, *right*) are shown next to their AM counterparts (Fig. 3, *left*). The pooled autocorrelogram for the QFM tone (Fig. 3*D*) is strikingly similar to its AM counterpart (Fig. 3*C*) in both inharmonic and harmonic regions. Likewise, the pooled autocorrelation histograms for corresponding stimulus segments resembled those of their AM counterparts (i.e., Fig. 3, *E* and *F, G* and *H*). Interval correlates of pitch ambiguity and octave pitch drop were thus also seen for QFM stimuli. The only discernible difference between pooled interval histograms for AM and QFM lies in subtle splits in side peaks for the harmonic QFM case (Fig. 3*F*). The similarity of pooled interval distributions is consistent with the psychophysical observation that low pitches produced by complex

tones with low-frequency harmonics are highly invariant with respect to phase manipulations.

Pitch circularity

Harmonic stimuli with many components spaced at octave intervals can also evoke a multiplicity of pitches ("octave confusions"). Shepard constructed a series of such ambiguous stimuli to illustrate the relativity of pitch intervals (Shepard 1964). Harmonics at octaves 1–11 comprise the stimulus. The waveform itself (Fig. 4*A, top*) has a fractal form that is invariant with respect to changes in time scale (Schroeder 1986). Shepard attenuated frequency components according to a fixed, log-cosine spectral envelope (Fig. 4*A*), so that F_0 and all of the octave harmonics could then be slowly swept upward with minimal changes in timbre. Shepard used discrete versions of this timbre-invariant stimulus to demonstrate that pitch has a locally transitive but globally circular structure analogous to the visual staircase illusions of Penrose and Penrose (1958) and M. C. Escher. Subjectively, one hears a continuously rising pitch that at some ill-defined point seamlessly drops down an octave, only to continue rising again.

Pitch circularity was investigated with the use of a continuous Shepard-Risset tone (Risset 1971; Shepard 1964) in which an entire octave cycle of F_0s was continuously traversed over a 5.6-s period. The pooled response autocorrelo-

A Shepard-Risset stimulus

B

C F0 = 6.125 Hz, 1.4 s PST

Stimulus components

● Estimated pitch

F0 = 4.9 Hz, 0 s PST

F0 = 9.8 Hz, 5.6 s PST

Radius = frequency (Hz)

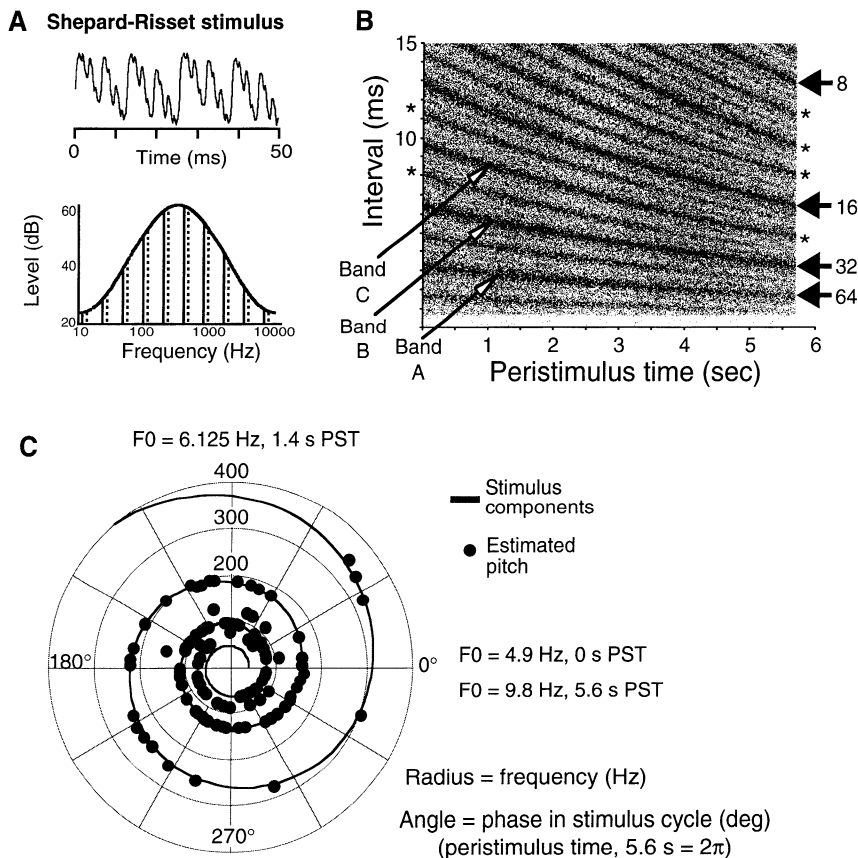Angle = phase in stimulus cycle (deg)
(peristimulus time, 5.6 s = 2π)

FIG. 4. Neural correlates of pitch circularity. A: waveform and power spectrum for the Shepard-Risset stimulus. All (octave) components are swept upward in frequency and attenuated according to the spectral envelope shown, with F_0 = 4.9–9.8 Hz. B: pooled response autocorrelogram (45 fibers). Harmonic numbers: interval bands associated with each octave harmonic. Asterisks: interval bands associated with component subharmonics. C: physiologically estimated pitches plotted on polar coordinates. PST increases as angle increases (counterclockwise). Frequency increases as radius increases.

gram (Fig. 4B) shows many interval bands of comparable density. The densest bands (←) correspond to stimulus components, whereas less dense ones (asterisks) correspond to component subharmonics. Following the interval bands marked "A" and "B" from the beginning of the stimulus to its end, the two bands begin as roughly equal, with band A gradually becoming less dense than band B as the stimulus ends. If the autocorrelogram is projected onto a cylinder with interval length as the vertical axis, and PST as its circumference, the sets of bands form spirals like those of a barber pole. These physiological results are comparable with those produced by autocorrelation models (Slaney and Lyon 1993).

Pitches estimated for each portion of the stimulus (Fig. 4C) show similar progressions. A polar coordinate system is useful for visualizing the spiral structure and multiplicity of pitch estimates. Because the stimulus traverses an octave cycle of F_0s, and then repeats, angle (0–360°) indicates where in the 5.6-s stimulus cycle a particular PST lies. Frequency is represented by radius. All frequency components present at a given time are therefore represented by a set of points having the same angle. Because F_0 is continually increasing, the points sweep out a spiral figure over time (———).

When pitch estimates for each stimulus time are plotted on these polar coordinates, most estimates fall on the octave spiral (i.e., at one of the octave components). A few estimates in 50- to 100-Hz range (inner circle) not on the octave spiral line trace out a spiral path corresponding to subharmonics of higher-frequency components (band C in the auto-

correlogram). Following the octave spiral counterclockwise (outward), it can be seen that there are fewer and fewer pitch estimates as frequencies increase above 200 Hz (outer 2 rings). The probability of "dropping down" an octave in pitch therefore increases as one proceeds around the spiral (as time goes on). If a probabilistic model for pitch based on relative heights of interval peaks is combined with a "proximity principle" for maintaining pitch continuity in time (so that one only rarely jumps from 1 turn of the spiral to another), then the sequence of (mostly rising) pitch estimates resembles those heard by human listeners (Shepard 1964).

Flanagan and Gutman rate pitch

Three click trains with variable F_0s were used to study the neural correlates of rate pitch. Waveforms and power spectra for corresponding segments of the three trains are shown in Fig. 5. Two of the click trains (labeled A and B) consisted of pulses having the same polarity (condensation clicks), whereas the third train (labeled C) consisted of pulses alternating in polarity (condensation alternating with rarefaction). Train A and train C share the same fundamental, whereas trains B and C share the same click rate, interclick interval, and harmonic spacing. Trains A and B consist of consecutive harmonics, whereas train C consists of odd harmonics. Relations between the time patterns and spectra of the three trains were maintained as their fundamentals were sinusoidally varied over 530 ms. The fundamental of click train A varied from 80 to 160 Hz, that of
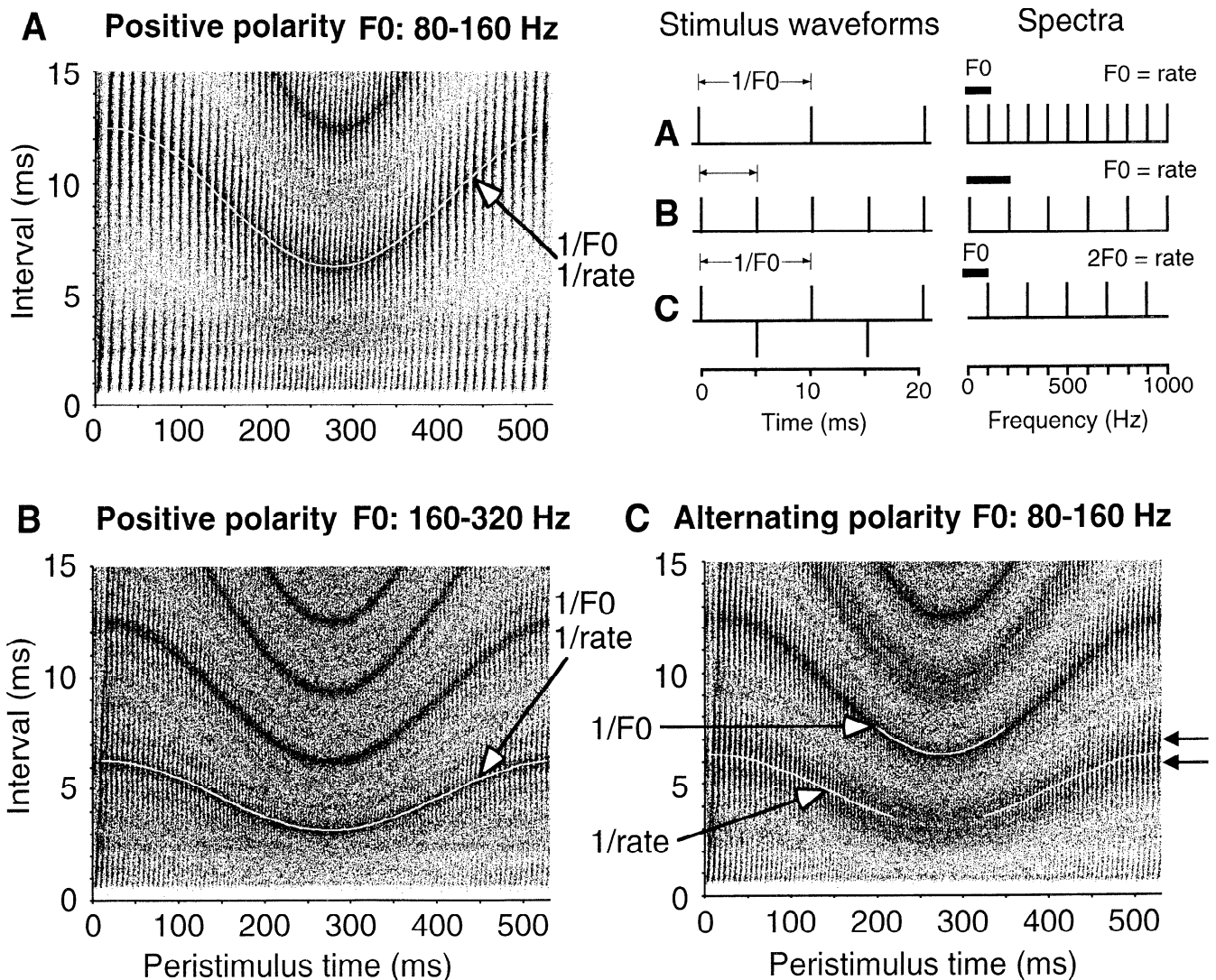
FIG. 5. Pooled interval distributions for click trains with variable fundamentals. *Top right*: representative stimulus waveforms and power spectra for click trains labeled A–C. *A–C*: pooled response autocorrelograms. *A*: unipolar click train with $F_0 = 80$–160 Hz. *B*: unipolar click train with $F_0 = 160$–320 Hz. *C*: alternating-polarity click train with $F_0 = 80$–160 Hz. White curved lines: periods of pitches heard (F_0 and/or rate pitch).

train B varied from 160 to 320 Hz, and that of train C varied from 80 to 160 Hz.

Unipolar click trains produce clear unambiguous pitches at their fundamentals, which equal their click rates. Alternating click trains can produce rate pitches associated with interclick interval or a periodicity pitch associated with the period of the alternating click pattern, the rate pitch being an octave higher than the periodicity pitch. Whether a rate pitch or a periodicity pitch is heard depends on F_0 (or click rate) (Flanagan and Gutman 1960a,b, 1964). F_0s <150 Hz (click rates < 300 Hz) tend to produce rate pitches, whereas F_0s above 150 Hz (click rates > 300 Hz) tend to produce periodicity pitches at the fundamental.

Pooled autocorrelograms for the two unipolar click trains are shown in Fig. 5, *A* and *B*. In both cases, the densest interval bands closely follow the fundamental period $1/F_0$ and the interclick interval, corresponding to the pitches that are heard in each case (white curved lines). The pooled autocorrelogram for the alternating train is shown in Fig.

5C. PSTs when rate and/or periodicity pitches could be heard are indicated by two sets of white curved lines. Here the densest interval band always follows the fundamental period $1/F_0$ throughout the stimulus duration. Although the most frequent interval corresponds to the pitch that is heard for F_0s >150 Hz (top curved line), the correspondence does not hold for F_0s <150 Hz (bottom curved lines). In place of a clear interval band at the rate pitch, there are two diffuse bands.

These patterns are seen more clearly in Fig. 6, where autocorrelation histograms for $F_0 = 80$ Hz (*right*) and $F_0 = 160$ Hz (*left*) are presented. Histograms for 50 single fibers arranged by CF are shown in the neurograms of Fig. 6, *A* and *D*. The corresponding pooled autocorrelation histograms, which are CF weighted but left unsmoothed, are shown in Fig. 6, *B* and *E*. Asterisks indicate the periods of pitches that would be heard. For both $F_0 = 80$ Hz and $F_0 = 160$ Hz, the highest peak in the pooled interval distributions is clearly at $1/F_0$. For $F_0 = 80$ Hz there are also two promi-
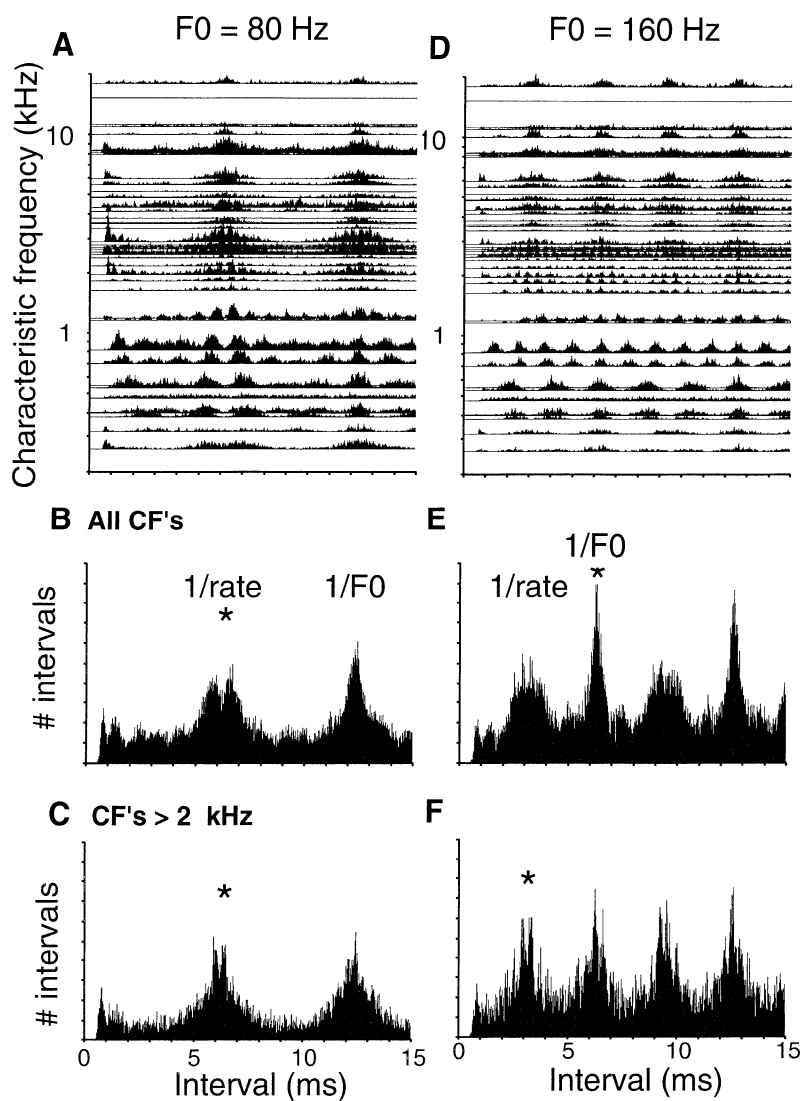
FIG. 6. Single-fiber and pooled autocorrelation histograms in response to alternating-polarity click trains. A and D: single-fiber autocorrelation histograms for 50 fibers arranged by characteristic frequency (CF), when $F_0 = 80$ Hz and $F_0 = 160$ Hz. B and E: pooled autocorrelation histograms. C and F: pooled autocorrelation histograms for 30 fibers with CFs >2 kHz. Asterisks: periods of pitches heard.

nent peaks near the interclick interval (1/click rate), but neither of these peaks is as high as that at $1/F_0$. Thus, because the most frequent interval in the pooled distribution does not follow the rate pitch that is heard, there is a clear discrepancy between human pitch judgments and the predominant interval hypothesis for pitches produced by alternating-polarity click trains with low F_0s.

The reason that most frequent interval does not follow the rate pitch for $F_0 < 150$ Hz is that interval peaks for 1/rate are split in the pooled autocorrelation histograms. These split peaks correspond to the two diffuse interval bands in the autocorrelogram (Fig. 5C). The origin of these two interval bands is evident in the autocorrelation neurogram of Fig. 6A. Whereas fibers with CFs >2 kHz show one interval peak near the interclick period (1/rate), those with lower CFs show two peaks that diverge with decreasing CF. These pairs of interval peaks are created by different spike latencies ($\pm \frac{1}{2}$CF) for clicks of positive and negative polarity. When a fiber responds to clicks of opposite polarity, intervals related to the interclick interval $\pm$ the CF-dependent latency are generated (Kiang et al. 1965). Consequently, when the responses of many fibers with different CFs are summed, a

pair of interval peaks is created in the pooled autocorrelation histogram, with neither peak being as high as the interval peak at the fundamental period.

When alternating-polarity click trains are high-pass filtered, rate pitches are strengthened and their existence region is extended upward to F_0s ≥ 400 Hz (Flanagan and Gutman 1964). Because low-CF fibers respond poorly to higher-frequency stimuli, the population interval distribution for these high-pass-filtered click trains can be estimated by summing the responses of fibers with CFs above the cutoff frequency. From the neurograms of Fig. 6, A and D, it can be seen that high-CF fibers produce comparable interval peaks at 1/click rate and at $1/F_0$. When only fibers with CFs > 2 kHz are used to construct pooled autocorrelation histograms (Fig. 6, C and F), for F_0s of both 80 and 160 Hz, interval peaks at 1/click rate become roughly equal in height to those at $1/F_0$. Thus the predominant interval hypothesis does appear to account for rate pitch in the high-pass-filtered case.

Other features of neural responses also correspond with rate pitch. Examining the autocorrelogram of Fig. 5C, prominent vertical striations can be seen for PSTs 0–200 ms and 330–530 ms, when fundamentals are <150 Hz. These verti-
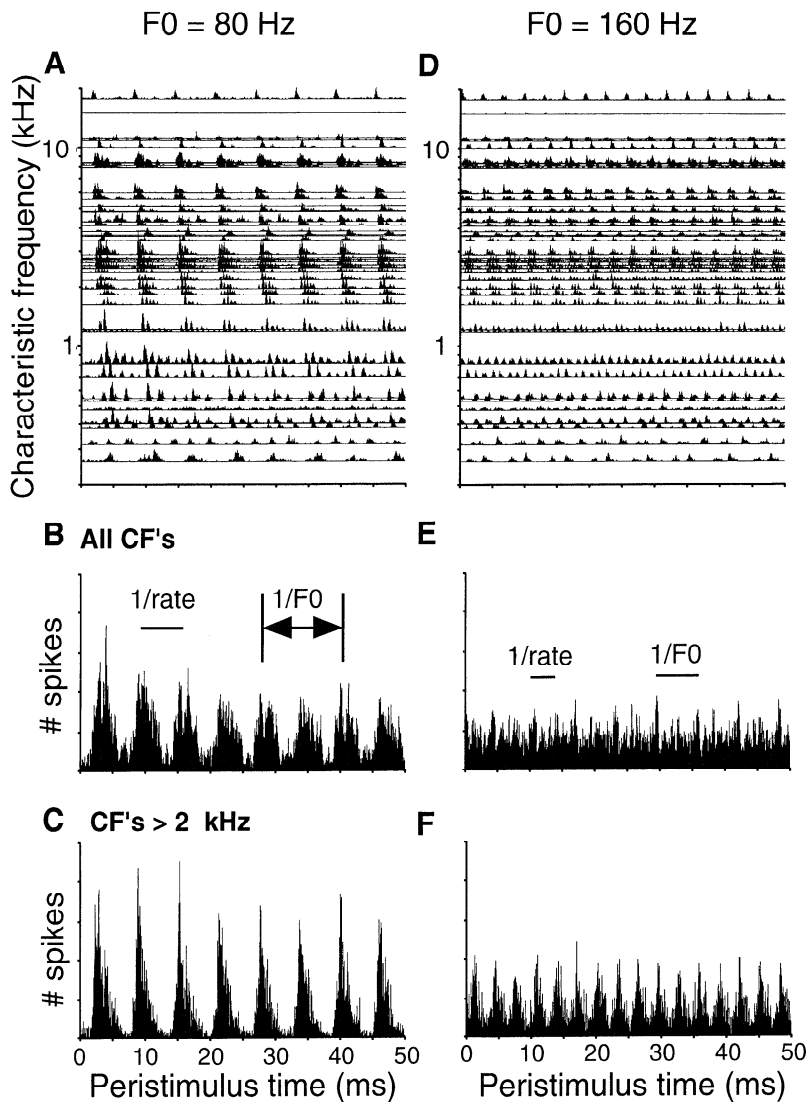
FO = 80 Hz   FO = 160 Hz

FIG. 7. Single-fiber and pooled PST histograms in response to alternating-polarity click trains. A and D: single-fiber PST histograms for 50 fibers arranged by CF, when $F_0 =$ 80 Hz and $F_0 = 160$ Hz. B and E: pooled PST histograms. C and F: pooled PST histograms for 30 fibers with CFs >2 kHz.

cal striations indicate a high degree of interneural synchrony across the auditory nerve. This can be seen very clearly in the neurogram and pooled PST histograms of Fig. 7. For $F_0 = 80$ Hz, when the rate pitch is heard, many fibers across the auditory nerve are discharging in rough synchrony (Fig. 7A), thereby producing regular, distinct peaks in the pooled PST histogram (Fig. 7B) that are separated by the interclick interval (or 1/click rate). In contrast, when $F_0 = 160$ Hz, and the rate pitch is not heard, there are no clear peaks in the pooled PST histogram (Fig. 7E). This apparent loss of interneural synchrony on a population level is not due to changes in single-fiber discharge patterns, but to the smearing out of peaks and troughs in population PST histograms as CF-dependent latency differences reach significant fractions of interclick intervals. When pooled PST histograms of fibers with CFs >2 kHz are examined (Fig. 7, C and E), clear peaks separated by the interclick interval are apparent for both $F_0 = 80$ Hz (Fig. 7C) and $F_0 = 160$ Hz (Fig. 7F). Thus, when a rate pitch is heard, there exists widespread interfiber synchrony across the entire population that gives rise to clear periodicities in population discharge

rates. When rate pitches are absent, no clear, periodic patterns of population discharge rates are apparent.

Dominance region for pitch

Two sets of harmonic complexes were used to investigate the dominance region for pitch. One complex consisted of harmonics 3–5 of one fundamental ($F_{0\ 3-5}$), whereas the other consisted of harmonics 6–12 of another fundamental ($F_{0\ 6-12}$) that was always 20% higher in frequency. This F_0 separation, also employed by Meddis and Hewitt (1991a), was made greater than those used in comparable psychophysical studies (Plomp 1967, 1976), to ensure that interval peaks could be clearly identified and separated. All frequency components had the same magnitude (60 dB SPL) and were in cosine phase. The waveforms and magnitude spectra of the individual complexes and their combination are shown in Fig. 8. Three sets of waveforms were constructed to cover the ranges of 80–160 Hz, 160–320 Hz, and 240–480 Hz. For the pitch dominance experiment, the two complexes were presented together. Thus for each range
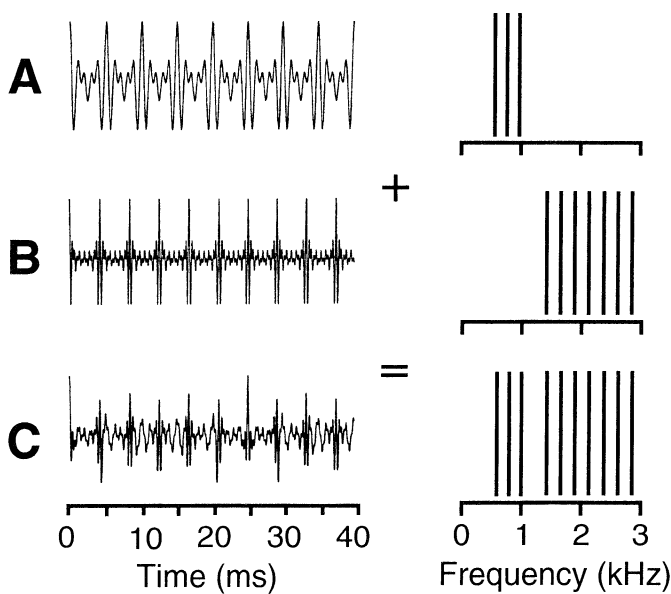
FIG. 8. Stimuli used to study the neural correlates of the dominance region for pitch. *Left*: waveforms. *Right*: power spectra. *A*: harmonics 3–5. *B*: harmonics 6–12 of a fundamental 20% higher in frequency. *C*: harmonics 3–5 and harmonics 6–12 presented concurrently.

of F_0s, there were three stimulus conditions: harmonics 3–5 presented alone, harmonics 6–12 presented alone, and both complexes presented together. From psychophysical studies (Moore and Glasberg 1985; Plomp 1967, 1976; Ritsma 1967), human listeners would be expected to hear the $F_{0\,3-5}$ over $F_{0\,6-12}$ throughout the range of fundamentals from 100 to 500 Hz.

Pooled autocorrelation histograms for F_0s near 80, 160, 240, 320, and 480 Hz are shown in Fig. 9. The two vertical dotted lines running through each histogram indicate the fundamental periods of harmonics 3–5 ($1/F_{0\,3-5}$) and harmonics 6–12 ($1/F_{0\,6-12}$). The *top row of histograms* shows the pooled interval distributions for harmonics 3–5 presented alone, and in each case the highest interval peak is very close to the fundamental period of the complex ($1/F_{0\,3-5}$). The *middle row of histograms* shows a similar result for harmonics 6–12. The *bottom row of histograms* shows the pooled distributions for the two complexes presented together. In the *leftmost histogram*, when $F_{0\,3-5} = 80$ Hz, the highest interval peak lies near the fundamental period of harmonics 6–12, whereas for all other F_0s shown (160, 240, 320, and 480 Hz) the highest peak lies near that of harmonics 3–5. For all F_0s the pooled distribution for the complexes presented together resembles the sum of the distributions of the complexes presented individually. Thus the complex producing the highest salience (peak-to-background ratio) when presented alone is the one that produces the highest salience (the dominant pitch) when the complexes are presented together. In all three cases, as F_0 increases, peak-to-background ratios decrease, paralleling the decline in pitch salience that is observed when complex tones have higher harmonic numbers.

Algorithms for pitch and salience estimation were used to provide a more quantitative description of the results. Pitch estimates for the individual and combined complexes are plotted as a function of F_0 in Fig. 10, *A* and *B*. The solid

lines indicate the F_0s of the two complexes. Over the entire range of F_0s, the estimated pitch of each complex presented alone (Fig. 10A) closely followed its corresponding fundamental (the corresponding solid line). For harmonics 6–12, as the F_0 increased beyond 300 Hz and component frequencies increased beyond 2 kHz, pitch estimates showed wider deviations from the fundamental. For these F_0s, individual harmonics would not be psychophysically resolved.

When the two complexes were presented together (Fig. 10B), estimated pitches followed either one or the other fundamental, depending on the F_0. Pitch estimates for a given PST are plotted against $F_{0\,3-5}$ at that time. For F_0s > 160 Hz, estimated pitches consistently followed $F_{0\,3-5}$. For F_0s from 120 to 160 Hz, estimates included roughly equal numbers of pitches near each fundamental. For F_0s < 120 Hz, the estimated pitches consistently followed $F_{0\,6-12}$.

Estimated saliences for the harmonic complexes are plotted as a function of F_0 in Fig. 10, *C* and *D*. Estimated salience curves for the fundamentals of each harmonic complex presented alone are shown in Fig. 10C. The salience curve for harmonics 6–12 has been shifted leftward (downward in frequency) by 20% to facilitate direct comparison with the salience curves in the concurrent case (Fig. 10D). The salience for harmonics 3–5 reaches a maximum when F_0 is between 150 and 200 Hz, when the frequency components of the complex are between 450 and 1,000 Hz. If the maximum of the salience curve for harmonics 6–12 is taken to be in the vicinity of 100 Hz, then the maximum saliences for both complexes would be obtained when the frequency components are in the 600- to 1,000-Hz range. Thus salience estimates appear to depend more on the frequencies of stimulus components than on F_0 or harmonic number.

Reinforcing this interpretation, estimated salience curves for the individual complexes undergo roughly parallel declines as the F_0 increases, so that the salience of harmonics 3–5 when $F_0 \approx 400$ Hz is roughly similar to that of harmonics 6–12 when $F_0 \approx 200$. The salience curves of the two complexes cross when the F_0 ($F_{0\,3-5}$) is 160 Hz. Thus for F_0s < 160 Hz, the estimated salience of harmonics 6–12 is greater than that of harmonics 3–5, whereas at higher F_0s the situation is reversed.

Estimated salience curves for the two fundamentals when both complexes are presented concurrently are shown in Fig. 10D. Estimated saliences for the two fundamentals are computed from the same pooled autocorrelation histograms with the use of the same background interval density value for both cases. Estimated saliences in the concurrent case have consistently lower values than those for separate presentation, indicating competition between the different complexes for production of intervals related to their fundamentals ("interval suppression"). Despite this difference, the forms of the salience curves and the F_0s at which they cross are very similar for the separate and competing conditions. Thus the fundamental that is most salient when presented in isolation is the one that dominates when the two are presented together.

Although all of the saliences discussed thus far have been estimates of interspike interval peak-to-background ratios for the entire auditory nerve, peak-to-background ratios were also computed from interval distributions of single fibers.
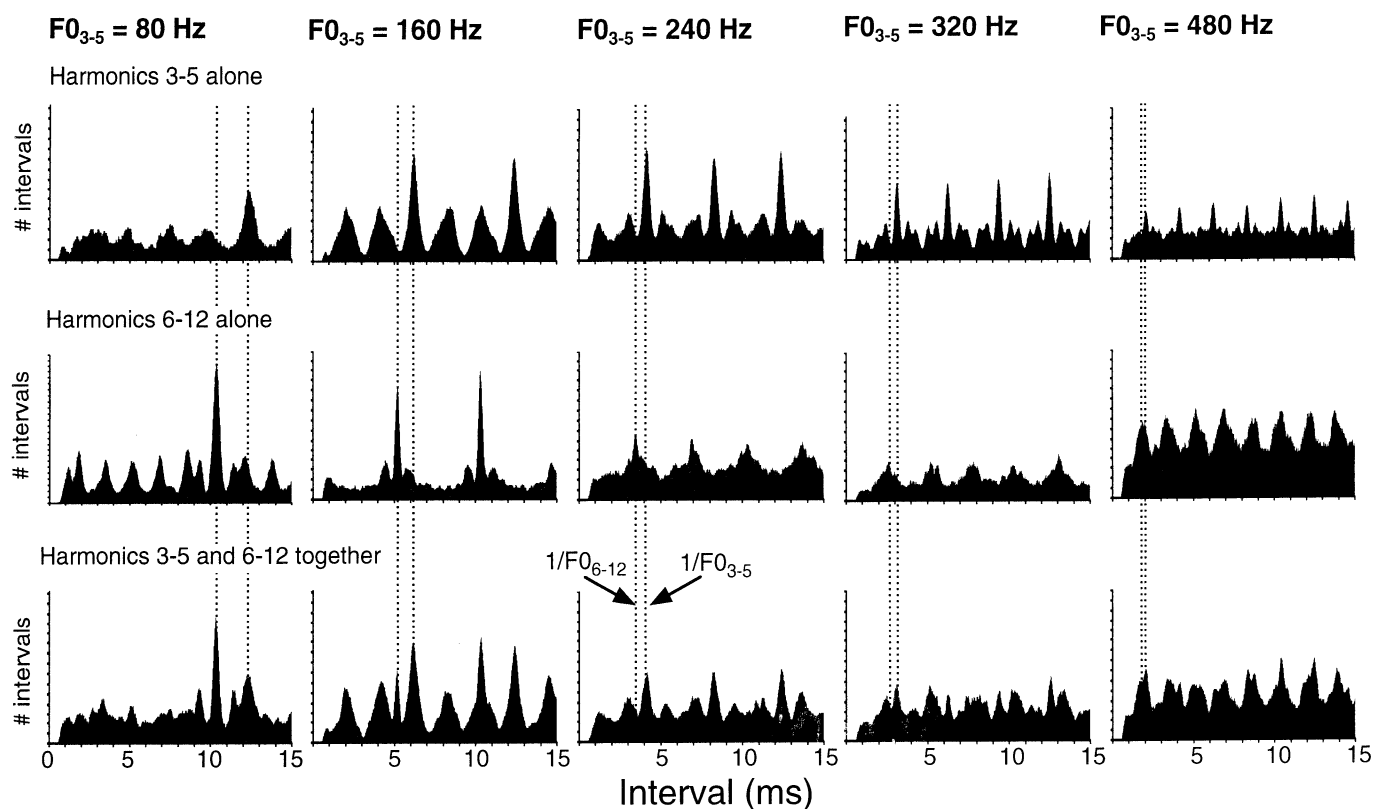
FIG. 9. Neural correlates of the dominance region for pitch. F_0s indicated are for harmonics 3–5. F_0s for harmonics 6–12 were always 20% higher. *Top row*: pooled interval histograms for harmonics 3–5 presented alone (47, 83, 49, 83, and 49 fibers, respectively). *Middle row*: pooled interval histograms for harmonics 3–5 presented alone (43, 70, 44, 70, and 44 fibers, respectively). *Bottom row*: pooled interval histograms for harmonics 3–5 and harmonics 6–12 presented together (60, 79, 72, 79, and 72 fibers, respectively). Vertical dotted lines: intervals corresponding to the respective fundamentals, harmonics 3–5 of one fundamental ($F_{0\,3-5}$) and harmonics 6–12 of another fundamental ($F_{0\,6-12}$) that was always 20% higher in frequency, with $1/F_{0\,6-12}$ always on the left and $1/F_{0\,3-5}$ always on the right. Ordinal axis range for histograms is 0–150 spikes per bin, except for the leftmost middle plot, where it is 0–250 spikes per bin.

These fiber saliences are useful for assessing relative contributions of different CF regions to overall salience estimates. Fiber saliences for the harmonic complexes presented both alone and concurrently are shown in Fig. 11. To facilitate comparisons between separate and concurrent presentations, fiber saliences were computed from the same histogram (i.e., the same peristimulus segment) when $F_{0\,3-5} = 150$ Hz and $F_{0\,6-12} = 180$ Hz. Fiber saliences for harmonics 3–5 presented alone were roughly equal for all CF regions (Fig. 11*A*, ●). Because stimulus components ranged from 450 to 750 Hz, there were very few fibers that did not respond to the stimulus (i.e., fiber saliences ≤ 0.1). In comparison, the fiber saliences for harmonics 6–12 presented alone (○) are noticeably smaller for CFs <600 Hz. Because stimulus components ranged from 1,080 to 2,160 Hz, this complex lies outside the response area of many low-CF fibers. Fiber saliences for each of the two fundamentals for the concurrent case are shown in Fig. 11*B*. Fiber saliences for $F_{0\,3-5}$ (●) drop off dramatically for CFs >800 Hz, whereas fiber saliences for $F_{0\,6-12}$ begin to rapidly increase. The frequency boundaries of the two complexes (450–750 Hz vs. 1,080–2,160 Hz) are consistent with the observed sharp crossover between 700 and 1,000 Hz. For higher F_0s, similar patterns of fiber saliences were seen, but with correspondingly higher crossover regions situated between the frequency boundaries

of the two complexes. Thus, for the concurrent case, different CF regions contribute intervals predominantly related to one or the other fundamental, giving rise to different temporal representations in different parts of the auditory nerve. Which CF regions contribute the most intervals can be determined from the stimulus spectrum.
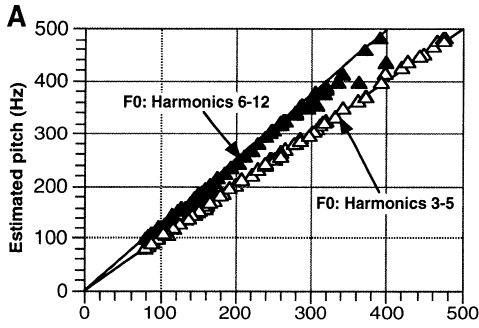
DISCUSSION

Predominant interval hypothesis for the low pitch of complex tones

The data presented in the preceding paper demonstrate that simple neural correlates of the low pitch of harmonic complexes exist in pooled interval distributions of the auditory nerve. Data presented here further demonstrate that neural correlates of more complex patterns of pitch judgments are also directly observable in pooled interval distributions of the auditory nerve.

The basic assumptions of the predominant interval hypothesis are that *1*) the low pitch of complex tones corresponds to the most frequent interval in the population interval distribution of the auditory nerve and *2*) the strength of this low pitch is related to the ratio of intervals associated with the pitch to the mean number of intervals present in

Harmonics 3-5 and 6-12 presented separately

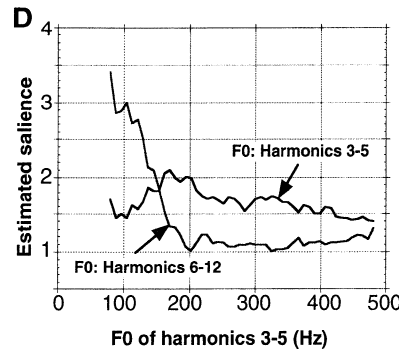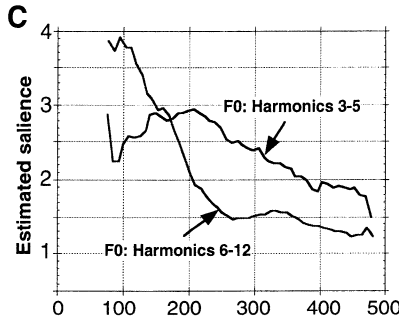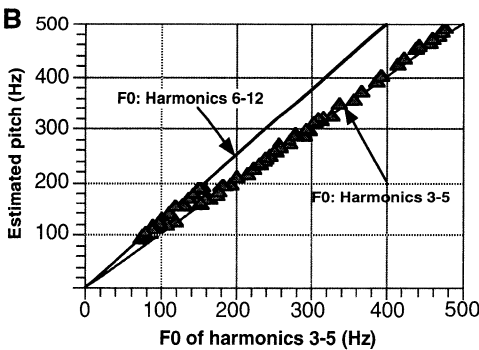**Harmonics 3-5 and 6-12 presented concurrently**

FIG. 10. Physiologically estimated pitches and pitch saliences for the dominance region experiment. All F_0s are plotted in terms of corresponding $F_{0\,3-5}$ (i.e., $F_{0\,6-12}$ is always 20% higher than plotted). A and B: pitch estimates as a function of F_0, $F_{0\,3-5}$. Estimates were made from consecutive 20-ms peristimulus response segments. C and D: estimated salience as a function of F_0 $F_{0\,3-5}$. Curves are least-squares fits of 9th-order polynomials to estimates made from 600 overlapping 20-ms peristimulus response segments.

the population distribution. Taken together with the basic assumptions, the data presented here provide explanations for pitch shift, pitch ambiguity, pitch circularity, the phase insensitivity of pitch, the pitch of unresolved harmonics, and the dominance region for pitch. The only pitch phenomenon studied that is not readily explicable in these terms is the rate pitch of alternating-polarity click trains with low fundamentals.

Pitch shift

Psychophysical experiments using inharmonic AM tones were originally carried out to probe the nature of the auditory pattern-recognition mechanisms responsible for the pitch of complex tones (de Boer 1976; Schouten 1940a,b; Schouten et al. 1962; Warren 1982). If auditory pitch mechanisms used local frequency spacings between spectral peaks (''dif-

ference frequencies''), or waveform envelopes to make pitch discriminations, then pitch should remain constant as all frequency components are shifted by a constant amount. When frequency components were all shifted, rather than hearing a constant pitch at the difference frequency F_m, listeners heard a small shift in pitch away from F_m. These experiments decisively ruled out 1) simple spectral pattern analyzers that detected absolute spacings between neighboring frequency components and 2) simple temporal pattern analyzers that detected periodicities of the waveform envelope. To account for these shifts, a spectral pattern processor must carry out an harmonic analysis of spectral patterns rather than simply recognizing constant frequency spacings, or alternately a temporal pattern processor must analyze the temporal fine structure of the stimulus rather than its envelope (de Boer 1956, 1976).

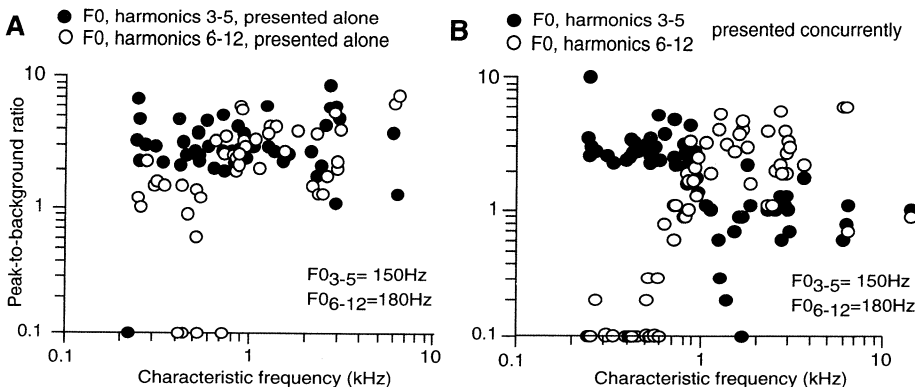Studies have previously shown that the fine structure of

FIG. 11. Fiber saliences as a function of CF in the dominance region experiment. $F_{0\,3-5}$ = 150 Hz, $F_{0\,6-12}$ = 180 Hz. A: fiber saliences (peak-to-background ratios) for $F_{0\,3-5}$ and $F_{0\,6-12}$, each presented individually. B: fiber saliences for $F_{0\,3-5}$ and $F_{0\,6-12}$ presented concurrently.

inharmonic complex tones is reflected in the temporal discharge patterns of single auditory nerve fibers (Evans 1978; Javel 1980). The physiological results presented here show that the first effect of pitch shift (de Boer's rule) can be directly explained in terms of an autocorrelation-like population-based predominant interval model for pitch. The more subtle "second effect of pitch shift," a small but systematic deviation from de Boer's rule, was not observed. This effect is thought to be caused by cochlear distortion products, "combination tones" (Smoorenburg 1970), that would be expected to be small given the relatively large component spacings for our stimuli. Although there exist explanations for inharmonic pitch shifts in terms of central mechanisms that perform optimal matching to harmonic spectral templates (Gerson and Goldstein 1978; Goldstein 1973; Terhardt et al. 1982a,b), the physiological data here indicate that these pitch shifts also have direct correlates in the population interspike interval distribution of the auditory nerve.

Pitch ambiguity, pitch circularity, and phase insensitivity

Pitch ambiguity was found to correspond to the appearance of multiple major peaks in pooled interval distributions. Where several discrete pitches would be heard for an inharmonic AM stimulus, multiple independent peaks of comparable height were seen (Fig. 2). Where stimulus parameters are shifted to positions where octave drops in pitch are expected (e.g., $F_c/F_m = n + 0.5$), corresponding patterns of peaks in pooled interspike interval distributions were observed (Fig. 2G). These sorts of patterns have been seen previously in single-fiber responses to inharmonic stimuli: AM tones (Javel 1980), nonuniform click trains (Evans 1978, 1983, 1986), and two-tone complexes (Greenberg and Rhode 1987).

The shifting relative heights of the various peaks in pooled interval distributions roughly parallel the shifting probabilities of particular pitch alternatives. Where the probability of hearing one pitch was greater than that of hearing another alternative pitch, its corresponding interval peak was higher than the other's. Thus the probability that a given pitch is heard among several possibilities appeared to correspond to the height of its associated interval peaks, relative to those associated with the other pitches.

When this probabilistic pitch allocation strategy is applied to the Shepard-Risset stimulus, and assumptions concerning local temporal contiguity of pitch judgments are incorporated, a reasonable account of pitch circularity is given. The pitch allocation principle also suggests possible interval-based explanations for several aspects of musical pitch: the various pitches that can be heard for a chord, the relative strengths of these pitches, and the degree to which they fuse to form a unified percept (Boring 1942; DeWitt and Crowder 1987; Parncutt 1989).

A considerable body of psychophysical work has investigated the role of phase spectrum and waveform envelope in pitch perception. Because changes in phase spectra can result in large changes in waveform envelope without concomitant changes in pitch, invariance of pitch frequency with respect to phase was used to falsify simple temporal models that measured intervals between peaks in the unfiltered waveform (Moore 1989; Wightman 1973a,b). A useful tactic has been to change phase spectra without altering magnitude spectra (Small 1970), thereby allowing spectral pattern cues (e.g., the difference frequencies) to be separated from temporal waveform cues (de Boer 1976; Schouten 1940b). Although low pitches of complex tones can be very sensitive to subtle shifts in magnitude spectra, as in "pitch shift" experiments, they are largely insensitive to changes in the phase spectra. Particular phase configurations can influence salience or distinctness of low pitches, especially for stimuli composed of higher-frequency components (Bilsen 1973; Licklider 1955; Nordmark 1978; Small 1970). However, these phase manipulations generally have little or no effect on the perceived frequencies of low pitches (Bilsen 1973; Meddis and Hewitt 1991b; Patterson 1973; Plomp 1976). These observations notwithstanding, some exceptions, involving inharmonic complex tones and phase-dependent magnitudes of distortion products, have prompted some debate over the issue (Green 1976). Other, high pitches can also be produced by particular phase spectra, but these do not alter the frequencies of the low pitches that are heard (Kubovy and Jordan 1979).

Although their power spectra are identical, waveforms of AM and QFM tones with low-frequency carriers are very different (Fig. 3, A and B). Despite their dissimilar waveforms, pitches produced by AM and QFM tones with low-frequency carriers (or low harmonic numbers) are largely indistinguishable (de Boer 1956; Zwicker 1962). The AM and QFM tones used in this study produce very similar pitches, despite different phase spectra and great dissimilarity in the temporal envelopes of their waveforms. Such stimuli also produced almost indistinguishable pooled interval histograms (Fig. 3). These results are broadly consistent with those obtained from the interval-based model of Meddis and Hewitt (1991a,b). For sets of low-frequency harmonics, the Meddis and Hewitt model produces pitch estimates that are highly independent of phase. The model can also account for some of the subtle differences in pitches and timbres of AM and QFM tones with higher-frequency carriers, higher harmonic numbers, and greater modulation indexes that have been debated in the literature (Ritsma and Engel 1964; Wightman 1973a). Although AM and QFM tones with some of these properties can produce changes in first-order interval distributions for single fibers when $F_c = CF$ (Horst et al. 1992), many of these differences would not be expected to be present in pooled distributions of all-order intervals. In summary, for AM and QFM tones with low-frequency harmonics, the predominant interval appears to be affected little by changes in phase-spectra or the shape of the waveform envelope. This is consistent with the highly phase-invariant character of both simulated population interval distributions and human pitch judgments.

Flanagan-Gutman rate pitch

Click trains of uniform polarity produce strong pitches at their F_0s. For these click trains F_0, click rate, and difference frequency are the same. For click trains of alternating polarity, however, click rate and difference frequency are twice the fundamental. For fundamentals <150 Hz, alternating click trains produce rate pitches at the click rate, whereas higher F_0s produce periodicity pitches at the fundamental (Flanagan and Gutman 1960a,b, 1964) Similar rate pitches

can also be produced by other alternating-polarity stimuli (Pierce 1991; Warren 1982) and by (Seebeck's) click trains with alternating interclick intervals (Evans 1978, 1983, 1986).

The phenomenon of rate pitch produced the only clear contradiction of the predominant interval hypothesis that was found. Observed pooled interval distributions always showed the most frequent interval to be at the fundamental, and consequently no interval correlates of rate pitch were initially seen. This discrepancy was reduced somewhat when correspondences between pooled interval distributions and rate pitches were found for high-pass-filtered click trains. When alternating click trains are passed through a high-pass filter, the rate pitch is strengthened and can be heard at higher fundamentals (Flanagan and Gutman 1964). Because few low-CF fibers would be expected to respond in the high-pass condition, distributions using only higher-CF fibers can be used to estimate the population response. When only intervals from higher-CF fibers are considered, then there are roughly as many intervals at $1/\text{rate}$ as there are at $1/F_0$. Because the first interval maximum is chosen for the pitch, the predominant interval model does correctly predict the rate pitches heard for the high-pass condition.

Another, possibly stronger neural correlate of rate pitch is the presence of periodic patterns of discharge on a population level. For F_0s <150 Hz, pooled PST histograms (Fig. 7) showed clear peaks in population discharge rates for every click in the train, as well as clear troughs during each interclick interval. The time between these peaks corresponds to the interclick period ($1/\text{click rate}$). When fundamentals increase to ≥ 150 Hz, CF-dependent latency differences become a large proportion of the interclick period, and population rate peaks are smeared out. When only high-CF fibers are included in the pooled PST histograms, clear peaks and troughs are present for all F_0s studied (80–160 Hz), consistent with the extension of the existence region of rate pitches by high-pass filtering. High-CF fibers have smaller CF-dependent latency-differences among themselves, and thus they discharge more synchronously as a group than their lower-frequency counterparts. Thus widespread discharge synchrony between fibers can lead to the generation of temporal discharge patterns on a population level. The "roughness" of some sounds (Plomp 1976; Terhardt 1973) might be explicable in such terms (Tramo et al. 1992). Such patterns could strengthen rate pitch relative to periodicity pitch if central analyzers count intervals between fibers (cross-correlation) as well as those within fibers (autocorrelation). Such a process might be consistent with the existence of two pitch mechanisms, one that produces a rough, buzzing, "rattle pitch" or rate pitch (Plomp 1976), and one that produces a smoother periodicity pitch. If both intraneural and interneural intervals are counted, the resulting predominant interval can coincide either with the most common interneural interval (rate pitch) or with the most common intraneural interval (periodicity pitch). Pitch would then be the result of a global correlation operation, possibly implemented via neural temporal cross-correlation architectures, e.g., (Braitenberg 1961; Cariani 1995a; Loeb et al. 1983).

Dominance region for pitch

Some early temporal theories of pitch (Schouten 1940b; Schouten et al. 1962) assumed that the low pitch of a har-

monic complex tone was generated by interaction of harmonics not resolved by peripheral auditory filters, the "residue" of an incomplete aural analysis. Because harmonic spacing remains constant while peripheral filter bandwidths increase with center frequency, residue pitch was thought to be generated mainly by sets of high-frequency components. To test this hypothesis, a number of different psychophysical experiments were carried out to determine the relative importance of low- and high-frequency harmonics for low pitch. When two harmonic complexes with slightly different fundamentals were presented together, human listeners almost invariably heard the pitch associated with the lower-frequency harmonics (Plomp 1967). This pitch is heard presumably because its constituent components carry greater weight in the central neural representation of pitch. Other experiments using different stimuli and psychophysical tasks (Bilsen 1973; Bilsen and Ritsma 1970; Moore and Glasberg 1985; Plomp 1976; Ritsma 1967; Yost 1982) confirmed the presence of a frequency region (the dominance region), roughly between 500 and 1,000 Hz, that disproportionately contributes to low pitch. All of the experiments suggested that resolved, lower-frequency harmonics have a much greater effect on the pitch of complex tones than unresolved, higher-frequency ones, and this led to a resurgence of spectral pattern theories for pitch (de Boer 1976).

The physiological data demonstrate that there are correlates of the dominance of low-frequency stimulus components in the population interspike interval distribution of the auditory nerve. For fundamentals from 160 to 500 Hz, intervals related to the fundamental of the lower-frequency harmonics clearly dominated in pooled distributions. As F_0s and thus component frequencies increase, the saliences of the associated low pitches decline, and this was seen directly in saliences estimated from the physiological data. For fundamentals <125 Hz, intervals related to the fundamental of the higher-frequency harmonics clearly dominated. This puts the crossover frequency for "interval dominance" roughly between 600 and 800 Hz, which is in the vicinity of crossover frequencies determined psychophysically. These results are generally in accord with those obtained by computer simulation (Meddis and Hewitt 1991b).

Predominant interval models thus provide an alternate, temporal interpretation of the dominance region for pitch. Here, the dominance region is a direct consequence of the differential abilities of auditory nerve fibers to phase-lock to stimulus components of lower and higher frequencies. The relative numbers of intervals produced by a given harmonic complex is thus a function of the aggregate numbers of discharges evoked by the stimulus and the degree to which the discharges phase-lock to stimulus components. Both aggregate numbers of discharges and degree of phase-locking depend on the distribution of fiber CFs and their positions relative to those of the stimulus harmonics present. In contrast to earlier temporal models based on unresolved harmonics, it is the lower-frequency components in the 500- to 1,000-Hz range that produce the most associated intervals. · Thus, from the perspective of a population interval model for pitch, it is immaterial whether the frequency components are psychophysically resolved or not, as long as intervals related to the fundamental are produced (see discussion of Fig. 12 in the companion paper). Interval-based models of

pitch therefore have the advantage of explaining the pitches generated by both resolved and unresolved harmonics in terms of the same mechanism.

Summary of pitch results

Taking the physiological data in its entirety, predominant interval models for pitch yield surprisingly satisfactory, comprehensive, and economical explanations for the pitch of complex tones. Many pitch phenomena are simply explained with the few model assumptions outlined above: the missing fundamental, pitch invariance with respect to level, pitch equivalence of spectrally diverse stimuli, pitch salience, pitch shift, phase insensitivity, and the dominance region for pitch. With the addition of a probabilistic pitch allocation rule, the population interval model can also explain the multiple pitches heard for some complex tones (pitch ambiguities). With assumptions concerning local temporal congruity of pitch judgments, the model can also give a reasonable account of pitch circularity.

Interval-based analysis could also be extended to form models for musical consonance (Tramo et al. 1992). In terms of pooled interval distributions, musical consonance is associated with fewer competing periodicities, whereas dissonance is associated with many competing periodicities (cf. Patterson 1986). Musical intervals with low integer ratios (e.g., 3:2, 4:3, 5:4) produce simpler autocorrelation patterns (fewer competing pitches) than do those with higher integer ratios (e.g., 12:11, 45:32), and these differences are directly reflected in pooled interval distributions.

As it stands, the predominant interval hypothesis has two apparent weaknesses. In the companion paper it was reported that the physiologically estimated salience of the 160-Hz pure tone was lower than expected relative to the estimated saliences of several complex tones. Likewise, in the present paper the dominance of harmonics 3–5 over harmonics 6–12 extended downward only to fundamentals as low as 160 Hz. Although the lower-frequency boundary of the dominance region is less well delineated in the psychophysical literature than the upper one, it was expected that harmonics 3–5 would dominate for fundamentals at least down to 100 Hz. Both of these results could be due to insufficient weight given to very low-CF responses when human population interval distributions are estimated from cat data. The second shortcoming of the predominant interval model is its inability to explain the Flanagan and Gutman click rate pitch. Here much more weight would need to be given to high-CF responses.

Central processing of temporal information for pitch

The strong correspondence between pitch judgments and the population interval distribution of the auditory nerve begs the question of the mechanism by which the central auditory system computes pitch. Perhaps the most convincing evidence for a central temporal analysis comes from electrical stimulation of humans via single-channel cochlear implants, where only very minimal "place" information is thought to be present. Psychophysical studies indicate that temporal periodicities up to several hundred hertz can be distinguished (Eddington et al. 1978; Shannon 1992). How-

ever, the pitches heard are generally weak, and are very poorly defined above 400 or 500 Hz, where just-noticeable frequency differences can be ≥ 100 Hz. These pitches are affected by both temporal pattern and place of stimulation in the cochlea (Eddington et al. 1978). These findings have decidedly mixed implications for temporal models for pitch. On one hand, they constitute an existence proof that purely temporal information is utilized by the central auditory system in forming pitch percepts. On the other, it is known that electrical stimulation produces more precise phase-locking in auditory nerve fibers than does acoustic stimulation (Dynes and Delgutte 1992), so that electrically evoked interspike interval distributions would be expected to produce much finer pitch discriminations than acoustical stimulation, in marked contrast to what is observed. Given the enormous amount of timing information presumed to be available, it can be argued that the purely temporal processing mechanisms normally play only a weak role in pitch perception, participating only in the discrimination of pitches below 300 Hz (Shannon 1992). This strong line of reasoning notwithstanding, it remains to be verified whether peripheral interspike interval distributions in response to complex electrical stimuli are indeed qualitatively similar to their acoustically produced counterparts (Javel et al. 1988; Javel 1988). Recent cochlear compound action potential data from human cochlear implant users (Wilson et al. 1994) suggest that the electrically stimulated auditory nerve may not be capable of conveying interval information for frequencies much above 400 Hz. This might then explain why cochlear implants fail to produce pitch discriminations above that limit.

Is enough timing information available at central auditory stations to support a temporally based mechanism for pitch? There are several ways by which interspike interval information might be conveyed and utilized by central auditory stations. One possibility is that temporal discharge patterns in the auditory periphery are transformed to discharge rate codes more centrally (Eggermont 1993; Epping and Eggermont 1986; Langner 1992; Schreiner and Langner 1988). Central auditory units whose discharge rate and/or stimulus-locked synchronization of discharges is maximal for particular stimulus periodicities have been found in many central auditory stations (Epping and Eggermont 1986; Langner 1992; Langner and Schreiner 1988; Schreiner et al. 1983). However, central representations for pitch based on these "modulation detectors" have several major difficulties: *1*) modulation tuning is coarse and degrades at high stimulus levels (Rees and Møller 1987) and in noise (Rees and Palmer 1989), *2*) the range of periodicities represented decreases dramatically at more central auditory stations, such that the range of best F_ms at the cortex is far below the range of periodicity pitches heard (Eggermont 1993, 1994), *3*) sharply tuned cortical "periodicity detectors" have yet to be found (Lyon and Shamma 1995; Schwartz and Tomlinson 1990; Sheich 1991), and *4*) analysis of envelope modulation is not how the central processor computes periodicity pitch, as was ruled out by experiments demonstrating the pitch shift of inharmonic AM tones (de Boer 1956; Schouten et al. 1962). Whatever the central mechanism for pitch, it must operate on either the fine structure of the waveform or the harmonic structure of its spectrum.

Some peripheral interspike interval information may be

preserved in higher centers. In the cochlear nucleus, pitch-related interspike intervals (Boerger 1974; Cariani 1995b,c; Kim and Leonard 1988; Kim et al. 1990; Rhode 1995; Shofner 1991) and discharge periodicities (Blackburn and Sachs 1990; Frisina et al. 1990; Glattke 1969; Greenberg and Rhode 1987; Palmer et al. 1986; Wang and Sachs 1994) are observed in many types of neurons. In the auditory midbrain and cortex, discharge periodicities up to 1 kHz have been observed in single- and multiple-unit responses to periodic stimuli (Langner 1992; Licklider 1959; de Ribaupierre et al. 1972; Rouiller et al. 1981), although periodicities above a few hundred hertz are rare in cortical units. Even evoked potentials, which detect only those temporal discharge patterns that are widely synchronized over neural populations, show periodicities up to 2 kHz at the level of the midbrain (Greenberg 1980) and up to several hundred hertz at the level of the cortex (Goldstein and Kiang 1958; Kiang and Goldstein 1959; Mäkelä et al. 1990; Steinschneider et al. 1980). A central interval code for periodicity pitch is thus not completely out of the question.

Finally, synchronous temporal codes in the periphery may be transformed into asynchronous temporal codes more centrally. These could include patterns of long intervals at fundamental subharmonics, interspike intervals jittered with respect to the stimulus and each other, intervals that are interleaved with other spike patterns, and interval patterns that are gated in time. These kinds of patterns might be expected from the high degree of divergence and convergence found in the auditory pathway, but none of the methods commonly utilized in the central auditory system to analyze spike trains would detect these transformations. Period histograms of their derived measures (synchronization and modulation indexes), for example, detect only stimulus-locked time patterns, and will miss more subtle, asynchronous ones. When more powerful analytical methods have been used to study neural coding in other parts of the brain, complex temporal patterns have been found (Abeles et al. 1993; Chung et al. 1970; Emmers 1981; Lestienne and Strehler 1987; Mountcastle 1993; Richmond et al. 1989; Villa 1992; Wasserman 1992). Until better methods are used to study responses to complex stimuli in more central auditory stations, complex temporal pattern codes cannot be ruled out.

If temporal information is preserved and/or transformed in the ascending auditory pathway, how this information might be utilized by central auditory processors to discriminate pitch remains a key question. Neural architectures that compute temporal autocorrelations (Chung et al. 1970; Licklider 1951, 1959; Reichardt 1961) and temporal cross-correlations (Braitenberg 1961, 1967; Jeffress 1948) have been proposed in the past, and such architectures may provide starting points for new neurally based temporal processing strategies for pitch. Other possibilities lie in the development of artificial, adaptive time-delay neural networks (Cariani 1995a; Chappelier and Grumbach 1994; Hopfield 1995; MacKay 1962; Mozer 1993; Pratt 1990; Tank and Hopfield 1987; Torras i Genis 1985), which in turn may inspire entirely new classes of physiologically based models for the perception of auditory forms.

REFERENCES

ABELES, M., BERGMAN, H., MARGALIT, E., AND VAADIA, E. Spatiotemporal firing patterns in the frontal cortex of behaving monkeys. *J. Neurophysiol.* 70: 1629–1638, 1993.

BILSEN, F. A. On the influence of the number and phase of harmonics on the perceptibility of the pitch of complex signals. *Acustica* 28: 60–65, 1973.

BILSEN, R. A. AND RITSMA, R. J. Some parameters influencing the perceptibility of pitch. *J. Acoust. Soc. Am.* 47: 469–475, 1970.

BLACKBURN, C. C. AND SACHS, M. B. The representation of the steady-state vowel sound /e/ in the discharge patterns of cat anteroventral cochlear nucleus neurons. *J. Neurophysiol.* 63: 1191–1212, 1990.

DE BOER, E. On the "Residue" in Hearing (PhD thesis). Amsterdam: Univ. of Amsterdam, 1956.

DE BOER, E. On the "residue" and auditory pitch perception. In: *Handbook of Sensory Physiology,* edited by W. D. Keidel and W. D. Neff. Berlin: Springer-Verlag, 1976, p. 479–583.

BOERGER, G. Coding of repetition noise in the cochlear nucleus of the cat. In: *Facts and Models in Hearing,* edited by E. Zwicker and E. Terhardt. New York: Springer-Verlag, 1974, p. 206–215.

BORING, E. G. *Sensation and Perception in the History of Experimental Psychology.* New York: Appleton-Century-Crofts, 1942.

BRAITENBERG, V. Functional interpretation of cerebellar histology. *Nature Lond.* 190: 539–540, 1961.

BRAITENBERG, V. Is the cerebellar cortex a biological clock in the millisecond range? *Prog. Brain Res.* 25: 334–346, 1967.

CARIANI, P. As if time really mattered: temporal strategies for neural coding of sensory information. *Commun. Cognit. Artif. Intell.* 12: 161–229, 1995a.

CARIANI, P. Physiological correlates of periodicity pitch in the cochlear nucleus. *Assoc. Res. Otolaryn. Abstr.* 128, 1995b.

CARIANI, P. A. Temporal coding of periodicity pitch in interspike interval distributions of cochlear nucleus populations (Abstract). *Ann. Biomed. Eng.* 23: S81, 1995c.

CHAPPELIER, J. C. AND GRUMBACH, A. Time in neural networks. *SIGART Bull.* 5: 3–11, 1994.

CHUNG, S. H., RAYMOND, S. A., AND LETTVIN, J. Y. Multiple meaning in single visual units. *Brain Behav. Evol.* 3: 72–101, 1970.

DEWITT, L. A. AND CROWDER, R. G. Tonal fusion of consonant musical intervals: the oomph in Stumpf. *Percept. Psychophys.* 41: 73–84, 1987.

DYNES, S. B. C. AND DELGUTTE, B. Phase-locking of auditory-nerve discharges to sinusoidal stimulation of the cochlea. *Hear. Res.* 58: 19–90, 1992.

EDDINGTON, D. K., DOBELLE, W. H., BRACKMAN, D. E., MLADEJOVSKY, M. G., AND PARKIN, J. Place and periodicity pitch by stimulation of multiple scala tympani electrodes in deaf volunteers. *Trans. Am. Soc. Artif. Intern. Organs* 24: 1–5, 1978.

EGGERMONT, J. J. Functional aspects of synchrony and correlation in the auditory nervous system. *Concepts Neurosci.* 4: 105–129, 1993.

EGGERMONT, J. J. Temporal modulation transfer functions for AM and FM stimuli in cat auditory cortex. Effects of carrier type, modulating waveform and intensity. *Hear. Res.* 74: 51–66, 1994.

EMMERS, R. *Pain: A Spike-Interval Coded Message in the Brain.* New York: Raven, 1981.

EPPING, W. J. M. AND EGGERMONT, J. J. Sensitivity of neurons in the auditory midbrain of the grassfrog to temporal characteristics of sound. I. Stimulation with acoustic clicks. *Hear. Res.* 24: 37–54, 1986.

EVANS, E. F. Place and time coding of frequency in the peripheral auditory system: some physiological pros and cons. *Audiology* 17: 369–420, 1978.

EVANS, E. F. Pitch and cochlear nerve fibre temporal discharge patterns. In: *Hearing: Physiological Bases and Psychophysics,* edited by R. Klinke and R. Hartmann. Berlin: Springer-Verlag, 1983, p. 140–146.

EVANS, E. F. Cochlear nerve fibre temporal discharge patterns, cochlear frequency selectivity and the dominant region for pitch. In: *Auditory Frequency Selectivity,* edited by B. C. J. Moore and R. D. Patterson. New York: Plenum, 1986, p. 253–264.

FLANAGAN, J. L. AND GUTMAN, N. On the pitch of periodic pulses. *J. Acoust. Soc. Am.* 32: 1308–1319, 1960a.

FLANAGAN, J. L. AND GUTMAN, N. Pitch of periodic pulses without fundamental component. *J. Acoust. Soc. Am.* 32: 1319–1328, 1960b.

FLANAGAN, J. L. AND GUTMAN, N. Pitch of high-pass-filtered pulse trains. *J. Acoust. Soc. Am.* 36: 757–765, 1964.

FRISINA, R. D., SMITH, R. L., AND CHAMBERLAIN, S. C. Encoding of amplitude modulation in the gerbil cochlear nucleus. I. A hierarchy of enhancement. *Hear. Res.* 44: 99–122, 1990.

GERSON, A. AND GOLDSTEIN, J. L. Evidence for a general template in central optimal processing for pitch of complex tones. *J. Acoust. Soc. Am.* 63: 498–510, 1978.

GLATTKE, T. J. Unit responses of the cat cochlear nucleus to amplitude-modulated stimuli. *J. Acoust. Soc. Am.* 45: 419–425, 1969.

GOLDSTEIN, J. L. An optimum processor theory for the central formation of the pitch of complex tones. *J. Acoust. Soc. Am.* 54: 1496–1516, 1973.

GOLDSTEIN, M. H., JR. AND KIANG, N. Y. S. Synchrony of neural activity in electric responses evoked by transient acoustic stimuli. *J. Acoust. Soc. Am.* 30: 107–114, 1958.

GREEN, D. M. *An Introduction to Hearing.* Hillsdale, NJ: Erlbaum, 1976.

GREENBERG, S. *Neural Temporal Coding of Pitch and Vowel Quality: Human Frequency-Following Response Studies of Complex Signals.* Los Angeles, CA: UCLA Working Papers in Phonetics #52, 1980.

GREENBERG, S. AND RHODE, W. S. Periodicity coding in cochlear nerve and ventral cochlear nucleus. In: *Auditory Processing of Complex Sounds,* edited by W. A. Yost and C. S. Watson. Hillsdale, NJ: Erlbaum, 1987, p. 225–236.

HOPFIELD, J. J. Pattern recognition computation using action potential timing for stimulus representation. *Nature Wash. DC* 376: 33–36, 1995.

HORST, J. W., JAVEL, E., AND FARLEY, G. R. Coding of fundamental frequency in auditory nerve fibers: effects of signal level and phase spectrum. In: *The Processing of Speech,* edited by M. E. H. Schouten. Berlin: Mouton-DeGruyer, 1992, p. 29–36.

JAVEL, E. Acoustic and electrical encoding of temporal information. In: *Cochlear Implants: Models of the Electrically Stimulated Ear,* edited by J. M. Miller and F. A. Spelman. New York: Springer-Verlag, 1988, p. 246–295.

JAVEL, E. Coding of AM tones in the chinchilla auditory nerve: implications for the pitch of complex tones. *J. Acoust. Soc. Am.* 68: 133–146, 1980.

JAVEL, E., MCGEE, J. A., HORST, W., AND FARLEY, G. R. Temporal mechanisms in auditory stimulus coding. In: *Auditory Function: Neurobiological Bases of Hearing,* edited by G. M. Edelman, W. E. Gall, and W. M. Cowan. New York: Wiley, 1988, p. 515–558.

JEFFRESS, L. A. A place theory of sound localization. *J. Comp. Physiol. Psychol.* 41: 35–39, 1948.

KIANG, N. Y. S. AND GOLDSTEIN, M. H., JR. Tonotopic organization of the cat auditory cortex for some complex stimuli. *J. Acoust. Soc. Am.* 31: 786–790, 1959.

KIANG, N. Y. S., WATANABE, T., THOMAS, E. C., AND CLARK, L. F. *Discharge Patterns of Single Fibers in the Cat's Auditory Nerve.* Cambridge: MIT Press, 1965.

KIM, D. O. AND LEONARD, G. Pitch-period following response of cat cochlear nucleus neurons to speech sounds. In: *Basic Issues in Hearing,* edited by H. Duifhuis, J. W. Horst, and H. P. Wit. London: Academic, 1988, p. 252–260.

KIM, D. O., SIRIANNI, J. G., AND CHANG, S. O. Responses of DCN-PVCN neurons and auditory nerve fibers in unanesthetized decerebrate cats to AM and pure tones: analysis with autocorrelation/power-spectrum. *Hear. Res.* 45: 95–113, 1990.

KUBOVY, M. AND JORDAN, R. Tone-segregation by phase: on the phase sensitivity of the single ear. *J. Acoust. Soc. Am.* 66: 100–106, 1979.

LANGNER, G. Periodicity coding in the auditory system. *Hear. Res.* 60: 115–142, 1992.

LANGNER, G. AND SCHREINER, C. E. Periodicity coding in the inferior colliculus of the cat. I. Neuronal mechanisms. *J. Neurophysiol.* 60: 1799–1822, 1988.

LESTIENNE, R. AND STREHLER, B. L. Time structure and stimulus dependence of precise replicating patterns present in monkey cortical neuronal spike trains. *Brain Res.* 43: 214–238, 1987.

LICKLIDER, J. C. R. A duplex theory of pitch perception. *Experientia* VII: 128–134, 1951.

LICKLIDER, J. C. R. Influence of phase coherence upon the pitch of complex sounds (Abstract). *J. Acoust. Soc. Am.* 27: 996, 1955.

LICKLIDER, J. C. R. Three auditory theories. In: *Psychology: A Study of a Science. Study I. Conceptual and Systematic,* edited by S. Koch. New York: McGraw-Hill, 1959, p. 41–144.

LOEB, G. E., WHITE, M. W., AND MERZENICH, M. M. Spatial cross-correlation. A proposed model for acoustic pitch perception. *Biol. Cybern.* 47: 149–163, 1983.

LYON, R. AND SHAMMA, S. Auditory representations of timbre and pitch. In: *Auditory Computation,* edited by H. Hawkins, T. McMullin, A. N. Popper, and R. R. Fay. New York: Springer-Verlag, 1995, p. 517.

MACKAY, D. M. Self-organization in the time domain. In: *Self-Organizing Systems 1962,* edited by M. C. Yovitts, G. T. Jacobi, and G. D. Goldstein. Washington, DC: Spartan, 1962, p. 37–48.

MÄKELÄ, J. P., KARMOS, G., MOLNAR, M., CSEPE, V., AND WINKLER, I. Steady-state responses from cat auditory cortex. *Hear. Res.* 45: 41–50, 1990.

MEDDIS, R. AND HEWITT, M. J. Virtual pitch and phase sensitivity of a computer model of the auditory periphery. I. Pitch identification. *J. Acoust. Soc. Am.* 89: 2866–2882, 1991a.

MEDDIS, R. AND HEWITT, M. J. Virtual pitch and phase sensitivity of a computer model of the auditory periphery. II. Phase sensitivity. *J. Acoust. Soc. Am.* 89: 2883–2894, 1991b.

MOORE, B. C. J. *Introduction to the Psychology of Hearing* (3rd ed.). London: Academic, 1989.

MOORE, C. J. AND GLASBERG, B. R. Relative dominance of individual partials in determining the pitch of complex tones. *J. Acoust. Soc. Am.* 77: 1853–1860, 1985.

MOUNTCASTLE, V. Temporal order determinants in a somatosthetic frequency discrimination: sequential order coding. *Ann. NY Acad. Sci.* 682: 151–170, 1993.

MOZER, M. C. Neural net architectures for temporal sequence processing. In: *Predicting the Future and Understanding the Past,* edited by A. Weigend and N. Gershenfeld. Redwood City, CA: Addison-Wesley, 1993, p. 243–264.

VAN NOORDEN, L. Two channel pitch perception. In: *Music, Mind and Brain,* edited by M. Clynes. New York: Plenum, 1982, p. 251–269.

NORDMARK, J. O. Frequency and periodicity analysis. In: *Handbook of Perception,* edited by E. C. Carterette and M. P. Friedman. New York: Academic, 1978, p. 243–282.

PALMER, A. R., WINTER, I. M., AND DARWIN, C. J. The representation of steady-state vowel sounds in the temporal discharge patterns of the guinea pig cochlear nerve and primarylike cochlear nucleus neurons. *J. Acoust. Soc. Am.* 79: 100–113, 1986.

PARNCUTT, R. *Harmony: A Psychoacoustical Approach.* Berlin: Springer-Verlag, 1989.

PATTERSON, R. D. The effects of relative phase and the number of components on residue pitch. *J. Acoust. Soc. Am.* 53: 1565–1572, 1973.

PATTERSON, R. D. Spiral detection of periodicity and the spiral form of musical scales. *Psychol. Music* 14: 44–61, 1986.

PENROSE, L. S. AND PENROSE, R. Impossible objects: a special type of visual illusion. *Br. J. Psychol.* 49: 31–33, 1958.

PIERCE, J. R. Periodicity and pitch perception. *J. Acoust. Soc. Am.* 90: 1889–1893, 1991.

PLOMP, R. Pitch of complex tones. *J. Acoust. Soc. Am.* 41: 1526–1533, 1967.

PLOMP, R. *Aspects of Tone Sensation.* London: Academic, 1976.

PRATT, G. *Pulse Computation* (PhD thesis). Cambridge, MA: MIT, 1990.

REES, A. AND MØLLER, A. R. Stimulus properties influencing the responses of inferior colliculus neurons to amplitude-modulated sounds. *Hear. Res.* 27: 129–143, 1987.

REES, A. AND PALMER, A. R. Neuronal responses to amplitude-modulated and pure-tone stimuli in the guinea pig inferior colliculus, and their modification by broadband noise. *J. Acoust. Soc. Am.* 85: 1978–1994, 1989.

REICHARDT, W. Autocorrelation, a principle for the evaluation of sensory information by the central nervous system. In: *Sensory Communication,* edited by W. A. Rosenblith. New York: MIT, 1961, p. 303–317.

RHODE, W. S. Interspike intervals as correlates of periodicity pitch in cat cochlear nucleus. *J. Acoust. Soc. Am.* 97: 2414–2429, 1995.

DE RIBAUPIERRE, F., GOLDSTEIN JR., M. H., AND YENI-KOMSHIAN, G. Cortical coding of repetitive acoustic pulses. *Brain Res.* 48: 205–225, 1972.

RICHMOND, B. J., OPTICAN, L. M., AND GAWNE, T. J. Neurons use multiple messages encoded in temporally modulated spike trains to represent pictures. In: *Seeing Contour and Colour,* edited by J. J. Kulikowski and C. M. Dickenson. New York: Pergamon, 1989, p. 705–713.

RISSET, J. C. Paradoxes de hauteur: le concept de hauteur sonore n'est pas les meme pour tout le monde. In: *7th International Congress on Acoustics.* Budapest, 1971.

RITSMA, R. J. Frequencies dominant in the perception of the pitch of complex sounds. *J. Acoust. Soc. Am.* 42: 191–198, 1967.

RITSMA, R. J. AND ENGEL, F. L. Pitch of frequency-modulated signals. *J. Acoust. Soc. Am.* 36: 1637–1644, 1964.

ROUILLER, E., DE RIBAUPIERRE, Y., TOROS-MOREL, A., AND DE RIBAUPIERRE, F. Neural coding of repetitive sounds in the medial geniculate body of cat. *Hear. Res.* 5: 181–100, 1981.

SCHOUTEN, J. F. The perception of pitch. *Phillips Tech. Rev.* 5: 286–294, 1940a.

SCHOUTEN, J. F. The residue, a new concept in subjective sound. *Proc. K. Ned. Akad. Wet.* 43: 356–365, 1940b.

SCHOUTEN, J. F., RITSMA, R. J., AND CARDOZO, B. L. Pitch of the residue. *J. Acoust. Soc. Am.* 34: 1418–1424, 1962.

SCHREINER, C., URBAS, J. V., AND MEHRGARDT, S. Temporal resolution of amplitude modulation and complex signals in the auditory cortex of the cat. In: *Hearing—Physiological Bases and Psychophysics,* edited by R. Klinke and R. Hartmann. Berlin: Springer-Verlag, 1983, p. 169–175.

SCHREINER, C. E. AND LANGNER, G. Coding of temporal patterns in the central auditory system. In: *Auditory Function: Neurobiological Bases of Hearing,* edited by G. M. Edelman, W. E. Gall, and W. M. Cowan. New York: Wiley, 1988, p. 337–362.

SCHROEDER, M. R. S. Auditory paradox based on fractal waveform. *J. Acoust. Soc. Am.* 79: 186–189, 1986.

SCHWARTZ, D. W. F. AND TOMLINSON, R. W. W. Spectral response patterns of auditory cortex neurons to harmonic complex tones in alert monkey (*Macaca mulatta*). *J. Neurophysiol.* 64: 282–298, 1990.

SHANNON, R. V. Temporal modulation transfer functions in patients with cochlear implants. *J. Acoust. Soc. Am.* 91: 2156–2164, 1992.

SHEICH, H. Auditory cortex: comparative aspects of maps and plasticity. *Curr. Opin. Neurobiol.* 1: 236–247, 1991.

SHEPARD, R. N. Circularity in judgments of relative pitch. *J. Acoust. Soc. Am.* 36: 2346–2353, 1964.

SHOFNER, W. P. Temporal representation of rippled noise in the anteroventral cochlear nucleus of the chinchilla. *J. Acoust. Soc. Am.* 90: 2450–2466, 1991.

SLANEY, M. AND LYON, R. F. On the importance of time—a temporal representation of sound. In: *Visual Representations of Speech Signals,* edited by M. Cooke, S. Beet, and M. Crawford. New York: Wiley, 1993, p. 95–118.

SMALL, A. M. Periodicity pitch. In: *Foundations of Modern Auditory Theory,* edited by J. V. Tobias. New York: Academic, 1970, p. 3–54.

SMOORENBURG, G. F. Pitch perception of two-frequency stimuli. *J. Acoust. Soc. Am.* 48: 924–942, 1970.

STEINSCHNEIDER, M., AREZZO, J., AND VAUGHAN, H. G., JR. Phase locked cortical responses to a human speech sound and low frequency tones in the monkey. *Brain Res.* 198: 75–84, 1980.

TANK, D. W. AND HOPFIELD, J. J. Neural computation by concentrating information in time. *Proc. Natl. Acad. Sci. USA* 84: 1896–1900, 1987.

TERHARDT, E. Pitch, consonance, and harmony. *J. Acoust. Soc. Am.* 55: 1061–1069, 1973.

TERHARDT, E., STOLL, G., AND SEEWANN, M. Pitch of complex signals according to virtual-pitch theory: test, examples, and predictions. *J. Acoust. Soc. Am.* 71: 671–678, 1982a.

TERHARDT, E., STOLL, G., AND SEEWANN, M. Algorithm for extraction of pitch and pitch salience from complex tonal signals. *J. Acoust. Soc. Am.* 71: 679–688, 1982b.

TORRAS I GENIS, C. *Temporal-Pattern Learning in Neural Models.* Berlin: Springer-Verlag, 1985.

TRAMO, M. J., CARIANI, P., AND DELGUTTE, B. Representation of tonal consonance and dissonance in the temporal firing patterns of auditory nerve fibers. *Soc. Neurosci. Abstr.* 18: 382, 1992.

VILLA, A. E. P. Temporal aspects of information processing in the central nervous system. *Annales du groupe CARNAC (Swiss Federal Institute of Technology EPFL)* 5: 15–42, 1992.

WANG, X. AND SACHS, M. B. Neural encoding of single-formant stimuli in the cat. II. Responses of anteroventral cochlear nucleus units. *J. Neurophysiol.* 71: 59–78, 1994.

WARREN, R. M. *Auditory Perception: A New Synthesis.* New York: Pergamon, 1982.

WASSERMAN, G. S. Isomorphism, task dependence, and the multiple meaning theory of neural coding. *Biol. Signals* 1: 117–142, 1992.

WIERSINGA-POST, E. AND DUIFHUIS, H. Probability distributions of the pitch of harmonic and inharmonic three tone complexes: Symmetric mistuning. In: *Advances in Hearing Research. Proceedings of the 10th International Symposium on Hearing, June 26–July 1, 1994, Irsee, Bavaria,* edited by G. A. Manley, G. M. Klump, C. Köppl, H. Fastl, and H. Oeckinghaus. Singapore: World Scientific, 1995, p. 488–497.

WIGHTMAN, F. L. Pitch and stimulus fine structure. *J. Acoust. Soc. Am.* 54: 397–406, 1973a.

WIGHTMAN, F. L. The pattern-transformation model of pitch. *J. Acoust. Soc. Am.* 54: 407–416, 1973b.

WILSON, B. S., FINLEY, C. C., ZERBI, M., AND LAWSON, D. T. *Speech Processors for Auditory Prostheses.* Quarterly Progress Report No. 7. Center for Auditory Prosthesis Research, Research Triangle Institute, 1994.

YOST, W. A. The dominance region and ripple noise pitch: a test of the peripheral weighting model. *J. Acoust. Soc. Am.* 72: 416–425, 1982.

ZWICKER, E. Direct comparisons between the sensations produced by frequency modulation and amplitude modulation. *J. Acoust. Soc. Am.* 34: 1425–1430, 1962.