

The representation of periodic sounds in simulated sustained chopper units of the ventral cochlear nucleus

Lutz Wiegrebe^{a)}

Dept. Biologie II, Universität München, Luisenstrasse 14, 80333 München, Germany

Ray Meddis^{b)}

Psychology Department, University of Essex, Wivenhoe Park, Colchester, CO4 3SQ, United Kingdom

(Received 11 July 2003; accepted for publication 1 December 2003)

The nature of the neural processing underlying the extraction of pitch information from harmonic complex sounds is still unclear. Electrophysiological studies in the auditory nerve and many psychophysical and modeling studies suggest that pitch might be extracted successfully by applying a mechanism like autocorrelation to the temporal discharge patterns of auditory-nerve fibers. The current modeling study investigates the possible role of populations of sustained chopper (Chop-S) units located in the mammalian ventral cochlear nucleus (VCN) in this process. First, it is shown that computer simulations can predict responses to periodic and quasiperiodic sounds of individual Chop-S units recorded in the guinea-pig VCN. Second, it is shown that the fundamental period of a periodic or quasiperiodic sound is represented in the first-order, interspike interval statistics of a population of simulated Chop-S units. This is true across a wide range of characteristic frequencies when the chopping rate is equal to the f_0 of the sound. The model was able to simulate the results of psychophysical studies involving the pitch height and pitch strength of iterated ripple noise, the dominance region of pitch, the effect of phase on pitch height and pitch strength, pitch of inharmonic stimuli, and of sinusoidally amplitude modulated noise. Simulation results indicate that changes in the interspike interval statistics of populations of Chop-S units compare well with changes in the pitch perceived by humans. It is proposed that Chop-S units in the ventral cochlear nucleus may play an important role in pitch extraction: They can convert a purely temporal pitch code as observed in the auditory nerve into a temporal *place* code of pitch in populations of cochlear-nucleus, Chop-S with different characteristic frequencies, and chopping rates. Thus, populations of cochlear-nucleus Chop-S units, together with their target units presumably located in the inferior colliculus, may serve to establish a stable rate-place code of pitch at the level of the auditory cortex. © 2004 Acoustical Society of America. [DOI: 10.1121/1.1643359]

PACS numbers: 43.64.Bt, 43.64.Qh, 43.66.Ba [WPS]

Pages: 1207–1218

I. INTRODUCTION

Tracking the pitch of periodic sounds is one of the fundamental tasks of the auditory system. Periodic sounds in nature are normally harmonically structured. The perceived pitch of the stimulus can typically be matched to a sinusoid whose frequency is the same as the fundamental of the harmonic series (F_0). This perception persists even when there is no auditory excitation in the F_0 frequency region, indicating that pitch perception is a synthetic process. The pitch percept is also particularly strong when the harmonic components are well separated relative to the frequency resolving power of the cochlea, suggesting that pitch often involves the combination of information taken across a range of frequency-selective channels. Any model of the physiological basis of pitch processing must take these basic facts into account.

Historically, models of pitch perception have championed either spectral (e.g., Terhardt, 1979; Goldstein, 1973) or temporal (e.g., Schouten, 1970; Licklider, 1951) explanations. Spectral explanations focus on the frequency-resolving

properties of the basilar membrane, while temporal explanations seek to show that the temporal resolution of the auditory-nerve (AN) response is an adequate basis for pitch extraction. Models using an autocorrelation analysis of auditory-nerve response patterns (Licklider, 1959; Lyon, 1984; Meddis and Hewitt, 1991a,b) have suggested that pitch could, in principle, be extracted from the temporal pattern of activity in the peripheral input. More direct evidence in favor of a temporal code for pitch comes from electrophysiological recordings from the cat auditory nerve (Horst *et al.*, 1986; Cariani and Delgutte, 1996a,b): They showed that temporal analyses such as autocorrelation, when applied to the auditory-nerve temporal discharge pattern, results in a reliable estimate of the stimulus pitch for a comprehensive set of test conditions.

The study to be described below takes this analysis as its starting point, but asks how this temporal information is extracted physiologically and made available for later processing. Physiological mechanisms that might serve the process that extracts pitch information from the temporal pattern of activity in the AN have not yet been discovered but may be located at least partly in the cochlear nucleus (CN), the only nucleus to receive direct AN input. de Cheveigne (1998) adopted a temporal approach and proposed that pitch percep-

^{a)}Electronic mail: lutzw@lmu.de

^{b)}Electronic mail: rmeddis@essex.ac.uk

tion may involve “an array of delay lines and inhibitory gating.” Shamma and Klein (2000) have suggested a template approach based on across-frequency coincidence detection. They treat the cochlear nucleus as a second stage containing “a matrix of coincidence detectors that compute the average pairwise instantaneous correlation (or product) between responses from all characteristic frequencies (CFs) across the channels.” Both these approaches are functionally plausible but await direct physiological confirmation. Below it will be suggested that the intrinsic membrane properties of sustained-chopper cells (Chop-S) provide a possible basis for the transformation of an auditory-nerve, temporal pitch code into a temporal *place* code of pitch at the level of the CN. This approach is based on the measured physiological responses of Chop-S units to pitch sounds, and the model has the advantage that it can be tuned to match the physiological responses. Of course, it remains the case that the autocorrelation analysis that has proved successful at the level of the AN would also work at the level of the cochlear nucleus, given that many units preserve the fine-structure information that is present in the AN. However, this would still leave the question as to how this autocorrelation is performed and at what level in the nervous system.

The cochlear nucleus contains different cell types that can be separated both anatomically and in terms of their electrophysiological response characteristics (Blackburn and Sachs, 1989). Shofner (1991, 1999) showed that primarylike units in the chinchilla ventral cochlear nucleus were well able to represent the period of rippled noise and iterated rippled noise in their temporal discharge patterns. Wiegrebe and Winter (2001) and Winter *et al.* (2001) have suggested that Chop-S units in the ventral cochlear nucleus (VCN) may play a role in pitch perception: Chop-S units are characterized by a regular response pattern in response to pure tones at their CF (Young *et al.*, 1988). More recently, it has been shown that the temporal discharge pattern becomes more regular when a Chop-S unit is stimulated with a periodic stimulus whose f_0 is near the unit’s intrinsic chopping rate (CR, Winter *et al.*, 2001; Wiegrebe and Winter, 2001). This selective enhancement of regularity is the focus of the model to be described.

The model consists of a large number of simulated Chop-S units organized according to their CF and CR. An ensemble of units with the same CR, but spread across CFs, is called a “Chop-S population.” Each Chop-S population will synchronize its activity to the f_0 of the input stimulus if the f_0 is close to the CR of the Chop-S population. The important feature of Chop-S units is that this synchronization to f_0 does not only occur in Chop-S units that are spectrally tuned to f_0 but also in Chop-S units that are spectrally tuned to integer multiples (harmonics) of f_0 . Information about the f_0 of the input is therefore distributed across all chopping units in terms of the amount of synchronization present in each, i.e., a synchronization/place code. This synchronization is quantified in the model as its normalized CR-locked rate; its computation is described below. A global estimate of the amount of synchronization to a particular fundamental frequency can be found by averaging the normalized CR-locked rate across a Chop-S population, i.e., across the Chop-S units

with the same CR but different CFs. This is used as the model’s estimate of the strength of a pitch equivalent to the CR. If pitch strength is estimated at a number of values of CR, the CR yielding the highest value can be used to predict the pitch most likely to be selected in a psychophysical pitch-matching experiment.

Clearly, this model can only represent the first stage of a much longer process whereby the synchronization observed in individual Chop-S units is further processed and combined to give a pitch percept. Given that the chopper units are thought to project to the inferior colliculus, it is possible that the next processing step is located there. It may involve coincidence units that are driven when populations of units with similar CF and CR fire in synchrony but not otherwise. Hewitt and Meddis (1994) have already shown how such circuits can convert a change in synchrony into a rate change. However, the present modeling study is only concerned with the processing observed at the level of the cochlear nucleus. The model is intended to demonstrate that the activity of the chopper units provides an adequate basis for later stages of pitch processing. It is already clear that auditory-nerve activity supplies an adequate basis for predicting psychophysical pitch judgments (Horst *et al.*, 1986; Cariani and Delgutte, 1996a,b). The aim of this study is to move forward to the next stage by identifying Chop-S units as the most likely carrier of pitch information by virtue of their ability to synchronize *selectively* to particular pitches. This selectivity converts the temporal code found in the auditory nerve to a temporal *place* code in the cochlear nucleus. As such, it makes a unique and important contribution to pitch processing in the auditory brainstem.

II. SIMULATIONS

A. Simulations of electrophysiological recordings

1. Model structure

The computer model incorporated the Chop-S model of Hewitt *et al.* (1992). Other models (Ghoshal *et al.*, 1992; Lai *et al.*, 1994) could have been used, and it is not expected that they would yield fundamentally different results. The input to the Chop-S model used recently updated models of the auditory periphery. Two different peripheral models were used: one for the guinea pig when modeling physiological observations and one for humans when modeling psychophysical data. The models consisted of (a) a system of band-pass filters designed to simulate the effect of the outer- and middle ear; (b) a dual-resonance, nonlinear filterbank with parameters set to match properties of the guinea pig (Meddis *et al.*, 2001); (c) mechanical to electrical transduction in the IHC; (d) IHC-auditory-nerve (AN) synapse (Sumner *et al.*, 2002); (e) AN spike generation sufficient to give 15 independent fibers with the same CF to each Chop-S unit; (f) the application of a 0.75-ms refractory time; and (g) Chop-S unit dendritic filtering using a first-order low-pass filter with a cutoff frequency of 300 Hz.

The Chop-S units were modeled using the algorithm suggested by MacGregor (1987). Fifteen input fibers were used because recent physiological investigations have revealed that the 60 fibers suggested in Hewitt *et al.* (1992)

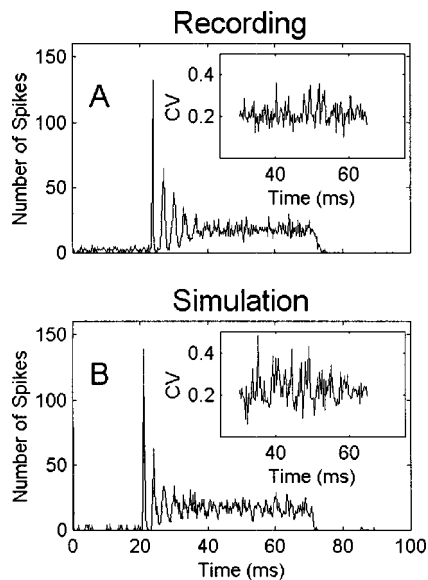


FIG. 1. Comparison of responses of a recorded Chop-S unit with a CF of 3 kHz and a chopping rate (CR) of 250 Hz (A) and its simulation (B) in response to 250 repetitions of 50-ms CF tones at 50 dB above threshold. Peristimulus time histograms are shown in (A) and (B); the coefficients of variation (CV) based on a regularity analysis are shown in the figure insets. Note that the simulation captures the main features of both the PSTH and the regularity analysis.

may be too many. Ferragamo *et al.* (1998), using *in vitro* techniques, estimated that at least five AN fibers converge on the dendrites of VCN stellate cells. We found that five input fibers were too few to generate realistic simulations of our animal data. The CF of the simulated Chop-S unit is determined by the center frequency of the cochlear filter driving its input AN fibers; the CR of the unit was manipulated by changing the potassium recovery time constant (Hewitt *et al.*, 1992).

All model parameters can be obtained from (<http://www.essex.ac.uk/psychology/hearinglab/models>). The model was implemented using publicly available computer coded modules in the Development System for Auditory Modelling (DSAM) library (footnote 1). The DSAM modules were called from MATLAB programs that generated the stimuli and analyzed the response of the modeled Chop-S units.

The electrophysiological recordings shown here are taken from the study by Winter *et al.* (2001).

2. Response to best-frequency pure tones

The parameters of the model were tuned so that the model generated pure-tone responses as similar as possible to those recorded in guinea pig in an earlier study (Winter *et al.*, 2001; Wiegube and Winter, 2001). Figure 1(A) shows the peristimulus time-histogram (PSTH) of a Chop-S unit with a CF of 3 kHz to 250 repetitions of 50-ms CF tones presented at 50 dB above unit threshold. Figure 1(B) shows the PSTH response of the model to the same stimulus. The figure insets show the coefficient of variation (CV) of the responses. The CV is computed by dividing the standard deviation of the interspike intervals by its mean (Blackburn and Sachs, 1989). Both animal and model data have an average CV close to 0.2.

3. Responses to complex tones

Hewitt *et al.* (1992) showed that simulated Chop-S units were able to predict the response characteristics of recorded Chop-S units when stimulated by sinusoidally amplitude modulated (SAM) tones. In our study, the refined Chop-S model was tested with an extended set of stimuli including iterated rippled noise (IRN) and harmonic complexes. The individual tone components of the harmonic complexes were added in either cosine phase (CPH) or random phase (RPH). The IRN stimuli were generated in an add-same configuration (Yost, 1996a). They are designated “IRNS(*d*,*g*,*n*)” where “*d*” is the delay in ms, “*g*” is the linear gain in the delay loop, and “*n*” is the number of iterations. Both RPH complexes and IRN stimuli were refreshed for every presentation for both recording and simulation. The f_0 (equivalent to the IRN delay reciprocal) ranged from 31.25 to 1000 Hz.

Responses were typically obtained using 25 repetitions of 409.6-ms stimuli generated with 20-ms, raised-cosine ramps at a sampling rate of 20 kHz. The presentation level was set so that neural rate responses were always well above threshold. First-order interspike interval histograms (ISIHS) in response to periodic sounds are shown in Fig. 2. Abeles (1982) introduced a method to quantify interspike interval distributions independent of binwidth and overall firing rate. The method is easily implemented by dividing the number of intervals in a bin of an interspike interval histogram by the product of binwidth and the overall number of interspike intervals. All ISIHS presented here are expressed in terms of firing rate as suggested by Abeles (1982). The recorded unit has a CR of 250 Hz and is the same unit as shown in Fig. 1. The figure is arranged with animal ISIHS in the left three columns and the model ISIHS in the right three columns. The three columns represent the response to IRN, RPH, and CPH, respectively. f_0 increases from bottom to top (see the central column). For all stimuli except the CPH with lower f_0 's, the ISIHS of both animal data and model response show a pronounced mode around 4 ms, regardless of the f_0 of the signal. This reflects the intrinsic chopping of the unit. However, this mode becomes taller and thinner when the f_0 of the signal is the same as the CR of the unit (250 Hz; equivalent to a chopping period of 4 ms).

Note that simulated responses to CPH tones show some deviations from the physiological data. Specifically, the ISIHS shows some temporal fine structure which is absent in the recorded data. The lack of fine structure in the recorded data may result from the influence of temporally random inhibition, as suggested by Lai *et al.* (1994).

Overall firing rate responses for the same recordings and simulations are shown in Fig. 3. For IRNS and RPH stimuli, rate responses are largely independent of the f_0 (fine and strong continuous lines). This result agrees with previous observations on modulation tuning in chopper units (Frisina *et al.*, 1990a,b). However, with CPH stimulation alone (dashed lines), both the recordings and simulations show a pronounced increase in the rate response with increasing f_0 . CPH closely resemble click trains. At very low f_0 's, the Chop-S unit responds with a single spike per click. As these become fewer with decreasing f_0 , the rate response decreases.

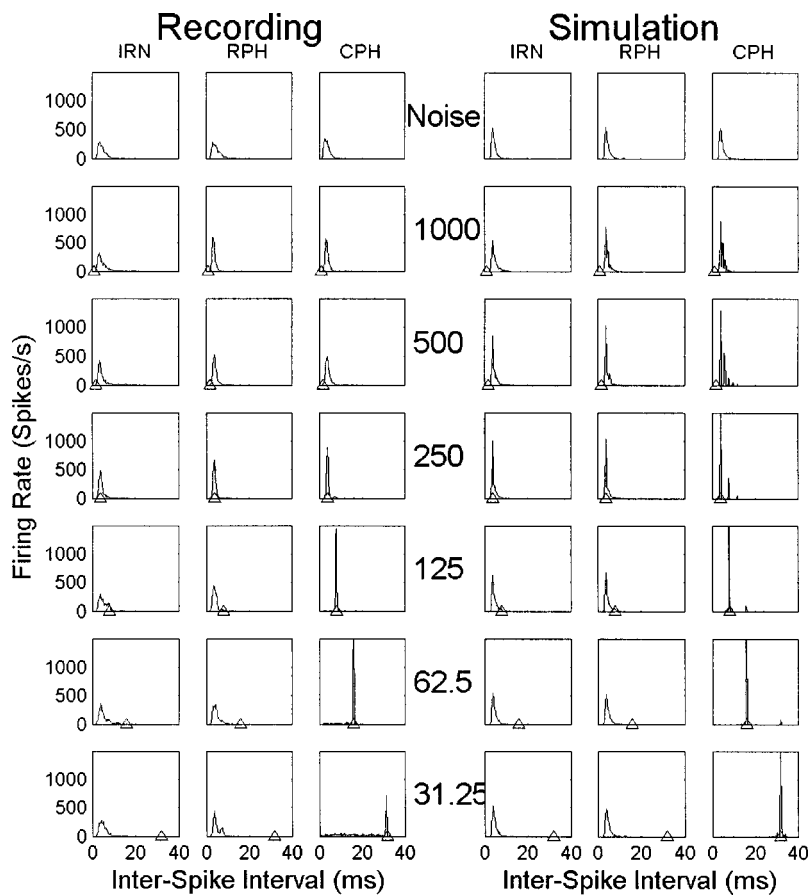


FIG. 2. Interspike interval histograms (ISIHs) of the 3-kHz Chop-S unit and its simulation as shown in Fig. 1 in response to harmonic complex tones and iterated rippled noise. The stimulus F_0 ($=1000/\text{delay}$ for IRNS) is given in the middle column. Stimuli were either IRNS($d,1,16$), random-phase harmonic complexes (RPH), or cosine-phase complexes (CPH). The ordinate of the ISIHs is scaled in terms of firing rate after Abeles (1982) to make the ISIHs independent of binwidth and overall firing rate. Note that when the stimulus F_0 matches the chopping rate (250 Hz), intervals are redistributed in favor of the interval corresponding to F_0 (upward triangles).

Frisina *et al.* (1990a,b) used amplitude modulated pure tones to demonstrate bandpass modulation tuning functions in VCN chop-S units. This shows that the Chop-S unit synchronized its discharges with the stimulus modulation fre-

quency for only a narrow range of modulation frequencies. A similar phenomenon can be demonstrated with IRN stimuli but a different method of analysis is required to show it, because IRN and RPH stimuli are not amenable to synchronization analysis. This is because their phase spectrum is random and, for IRN stimulation, the phase spectrum is also time variant. Thus, any analysis of regularity or synchrony with this kind of stimulation must be primarily based on interspike interval statistics.

Shofner (1999) quantified changes in the height of the first peak in the neural autocorrelogram (equivalent to an all-order, interspike interval analysis) using the formula

$$\text{Normalized, CR-locked rate} = \frac{R_\tau - R_{\text{ave}}}{R_{\text{ave}}},$$

where R_τ equals the ISIH firing rate at an interspike interval corresponding to the CR of the unit and R_{ave} equals the ISIH firing rate averaged across all interspike intervals. As the ISIH firing rates are independent of binwidth and overall firing rate (Abeles, 1982), also the normalized, CR-locked firing rate is independent of these parameters. We use the same measure of peak height here. Note, however, that, unlike in Shofner (1999), peak heights were calculated based on a first-order ISIH analysis, not on an all-order ISIH analysis. A first-order ISIH analysis was preferred because Winter *et al.* (2001) showed that the bandpass tuning to a specific range of f_0 's is demonstrated only with a first-order ISIH analysis. Binwidths of 250 μs are used throughout the whole body of this paper.

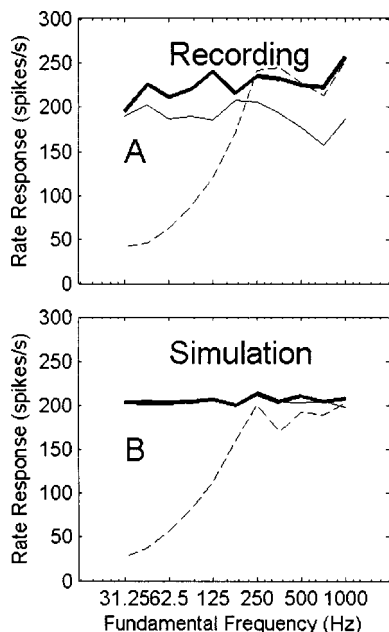


FIG. 3. Overall firing rate responses of the recording shown in Fig. 2. Note that both in the recording (A) and the simulation (B) the rate response is independent of the stimulus F_0 when the stimuli were IRNS or RPH (fine and strong continuous lines). With CPH stimulation (dashed lines), however, the rate response increases with increasing F_0 because of the transient periodic envelope fluctuation of the CPH.

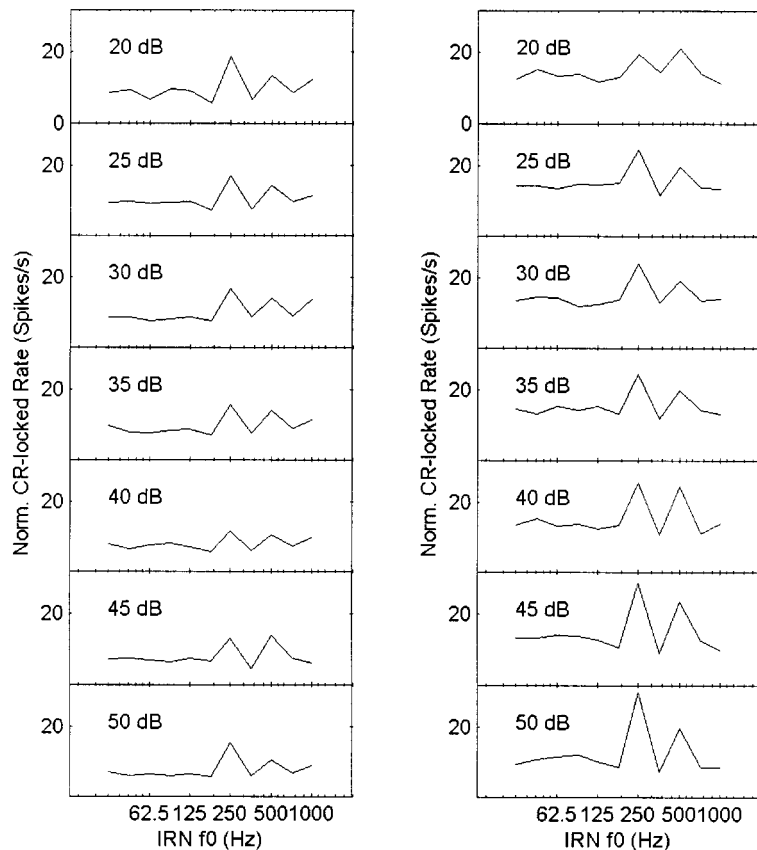


FIG. 4. Normalized CR-locked rate as a function of stimulus f_0 . The stimuli were $\text{IRNS}(d,1,16)$, where d equals the f_0 reciprocal. Recorded responses are shown in the left column; simulated responses are shown in the right column. The attenuation of the sound-pressure level of the stimulus is given in the individual panels. Note that when the IRNS f_0 matches the chopping rate of the unit (250 Hz), there is a clear peak in the normalized CR-locked rate. Moreover, the position of this peak is not affected by the sound-pressure level of stimulation. Thus, this measure is quite stable against changes in sound-pressure level. A secondary peak is observed when the IRNS f_0 is 500 Hz, i.e., twice the chopping rate. This feature of Chop-S units is discussed in Sec. III.

The normalized, CR-locked rate is shown as a function of f_0 in Fig. 4. The stimuli were $\text{IRNS}(d,1,16)$. Again, analyses of the guinea-pig recordings are shown in the left column; analyses of the simulated responses are shown in the right column. The recorded and the simulated unit was a Chop-S unit with a CF of 1000 Hz and a CR of 250 Hz. Different rows represent different stimulus attenuations as shown in the panels. Note that the normalized, CR-locked rate always shows a pronounced peak when the IRNS f_0 matches the CR. Moreover, this peak is stable over the 30-dB range of sound levels tested and it is similar in the recording and the simulation. These data reflect the previous finding that Chop-S units provide a selectivity for f_0 which is independent of sound level when the unit is in saturation (Wiegube and Winter, 2001). A secondary peak is seen at an f_0 of 500 Hz, i.e., when the IRNS f_0 is twice the CR. This secondary peak reflects the fact that a Chop-S unit is capable to lock onto an integer multiple of its input period. This is a possible neurophysiological foundation of octave confusions observed psychophysically. This topic is discussed in Sec. III.

B. Simulations of pitch phenomena in populations of Chop-S units

1. Model structure

Psychophysical studies of pitch perception have employed many different kinds of stimuli. The response of the model will now be evaluated using some of these stimuli. It was necessary to make two major changes to the model in order to make it suitable for evaluation against human, psy-

chophysical data. First, the peripheral model was changed to match that of human listeners. Second, the model was expanded to include populations of Chop-S units with a wide range of CRs (about 100 to 550 Hz). The outer- and middle-ear transfer function was modified to fit the human outer- and middle-ear transfer function (Glasberg and Moore, 2002). Cochlear processing was implemented with a filterbank consisting of dual-resonance, nonlinear, bandpass filters fitted to explain human forward-masking measurements (Lopez-Poveda and Meddis, 2001). The filterbank consisted of 20 nonlinear filters with CFs equally spaced on a logarithmic frequency axis between 200 and 6000 Hz. The simulated Chop-S units form populations of Chop-S units where each population is determined by its common CR. A population consists of 20 Chop-S units (i.e., one per filter) with a common CR and CFs determined by the center frequencies of the auditory filters. In other respects, the model is identical to that used to simulate the physiological recordings.

2. Simulation paradigm

Unless otherwise stated, simulated responses are based on 25 repetitions of 409.6-ms stimuli presented at 40 dB above rate threshold. Stimuli are always windowed with 20-ms raised-cosine gates. Again, data are analyzed in terms of the Chop-S first-order ISIH scaled as suggested by Abeles (1982). Binwidth was always 250 μs . The dependent variable is the normalized CR-locked rate produced by a stimulus. The model implies that the normalized CR-locked rate reflects the perceived pitch strength, while the pitch itself is

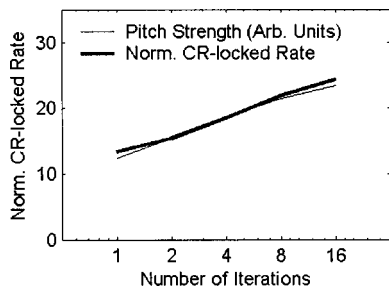


FIG. 5. The change of pitch strength (fine line; after Yost, 1996b) and normalized CR-locked rate (strong line) as a function of the number of iterations of IRNS(5,1, n). The normalized CR-locked rate was averaged over an array of 20 Chop-S units with CFs between 200 and 6000 Hz and a CR of 250 Hz. The data indicate the linear relationship between normalized CR-locked rate and the perceived pitch strength.

reflected in a distributed manner in terms of which population of Chop-S units shows the strongest normalized CR-locked rate.

3. Simulation 1: The representation of the pitch strength of IRN

Yost (1996b) showed that the pitch strength of IRN can be predicted using the height of the first major peak in the autocorrelation function of the stimulus waveform ($h1$). Specifically, the perceived pitch strength is proportional to 10 raised to the power of $h1$. The Chop-S model is evaluated by measuring the normalized CR-locked rate to IRN stimuli with different pitch strengths. The growth of the normalized CR-locked rate (averaged over a population of Chop-S units) is plotted against the number of iterations for both Yost's function and the model (Fig. 5). For this demonstration, the Chop-S units are spread over a wide range of CFs (200 to 6000 Hz) but CRs were fixed at 200 Hz, equal to the IRN delay reciprocal. Thus, the stimuli were IRNS(5,1, n), where n , the number of iterations, was 1, 2, 4, 8, and 16. The fine line in Fig. 5 represents Yost's estimate of pitch strength of the stimulus, based on $h1$, as a function of the number of iterations. The strong line show the normalized CR-locked rate averaged across an array of simulated Chop-S units. It can be seen that the normalized CR-locked rate for a population of simulated Chop-S units grows in proportion with Yost's estimate of perceived pitch strength. This is an important feature of the normalized CR-locked rate because it implies that the measure can be used additively when it is investigated how different units contribute to overall pitch strength. The feature is exploited in the subsequent simulations.

4. Simulation 2: The dominance region of harmonic sounds

For simple harmonic complexes, Ritsma (1967) showed that it is the low harmonics that contribute most to pitch strength. He asked listeners to match the pitch of a reference stimulus to that of a test stimulus. The test stimulus was a harmonic complex manipulated so that harmonics 1 to n were shifted down to be based on a lower f_0 ($F_{0_low} = 0.9 \cdot F_0$) and harmonics $n+1$ to 12 were shifted up to be based on a higher f_0 ($F_{0_high} = 1.1 \cdot F_0$). The reference

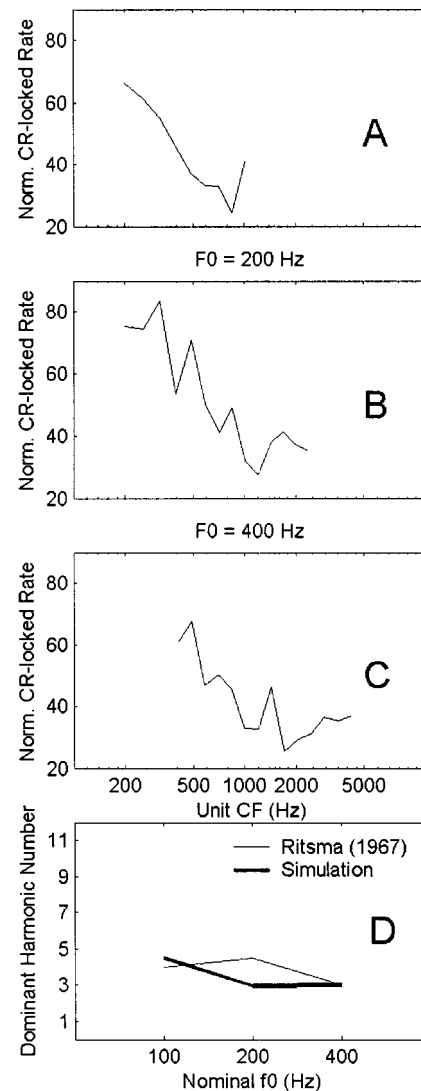


FIG. 6. Simulated neural correlates of the dominance region of pitch: Normalized CR-locked rate of a Chop-S population tuned to the stimulus F_0 is shown as a function of unit CF in (A) to (C) for F_0 's of 100, 200, and 400 Hz, respectively. Note that the low harmonics always produce the strongest normalized CR-locked rate. From these data, the center of dominance is calculated as described in the text. The predicted dominant harmonic number is shown along with the data of Ritsma (1967) in (D).

stimulus was a harmonic complex consisting of all 12 unmodified harmonics. Listeners were asked to judge whether the test stimulus was higher or lower in pitch compared with the reference stimulus. The center of dominance corresponds to that harmonic, n , which, when assigned either to the lower F_0 or to the higher F_0 in the test stimulus, results in a change in the listener's decision. However, this does not mean that n is the strongest harmonic; it means that the aggregated pitch strength of harmonics 1 to n is equal to the aggregated pitch strength of harmonics $n+1$ to 12 (Plomp, 1976).

The same stimuli were used in an evaluation of the current model. Normalized CR-locked rate in a Chop-S population with a CR equal to F_0 is shown as a function of unit CF for 12 harmonics of F_0 's of 100, 200, and 400 Hz in Figs. 6(A) to (C), respectively. The center of dominance is determined by that specific CF where the integrated, normalized

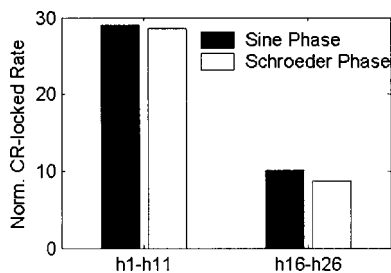


FIG. 7. Normalized CR-locked rate averaged across a population of Chop-S units with 20 CFs between 200 and 6000 Hz and a CR of 200 Hz in response to harmonic complexes with an F_0 of 200 Hz. Harmonics are added either in sine phase (solid bars) or in Schroeder phase (open bars). The left two bars show the normalized CR-locked rate averaged across a series of low harmonics ($h1-h11$); the right two bars show normalized CR-locked rate averaged across high harmonics ($h16-h26$). In agreement with the perception of pitch strength, the normalized CR-locked rate is strong and phase independent for the resolved harmonic complexes ($h1-h11$), and it is generally weaker for high complexes ($h16-h26$). Moreover, as in the perception, the normalized CR-locked rate for high Schroeder-phase complexes is slightly weaker than for high sine-phase complexes.

CR-locked rate below this CF equals the integrated, normalized CR-locked rate above this CF. The resulting CF is converted to harmonic number and plotted in Fig. 6(D) as a function of the stimulus F_0 (strong line). The data show that the simulated center of dominance matches the data by Ritsma (1967) reasonably well.

5. Simulation 3: The influence of phase on pitch strength

Houtsma and Smurzynski (1990) investigated the extent to which the pitch strength of harmonic complexes depends on the phase relations between the harmonics. They measured frequency difference limens to obtain a quantitative measure of pitch strength. The influence of phase was measured for a set of low harmonics (1–11) as well as for a set of high harmonics (16–26). Houtsma and Smurzynski (1990) found that for a set of low harmonics pitch strength was high, and the phase relation between the harmonics did not influence pitch strength. For a set of high harmonics, pitch strength was generally much lower, and the pitch strength of harmonics added in sine phase was slightly stronger than when the harmonics were added in Schroeder phase. Adding harmonics in Schroeder phase minimizes the amplitude modulation of the waveform envelope, whereas sine-phase harmonics result in a “peaky” waveform. When simulating their experiment, harmonic complexes consisted of either 11 low harmonics ($h1$ to $h11$) or 11 high harmonics ($h16$ to $h26$). F_0 was set to 200 Hz as used by Houtsma and Smurzynski (1990). Simulated neural output was obtained from 25 repetitions of 409.6-ms stimuli. The simulated Chop-S population had a CR of 200 Hz (corresponding to the F_0 of 200 Hz).

The normalized CR-locked rate averaged across unit CFs is shown for low harmonics ($h1-h11$) and high harmonics ($h16-h26$) added in either sine phase (filled bars) or Schroeder phase (open bars) in Fig. 7. As expected from the previous simulation, the normalized CR-locked rate caused by low harmonics is considerably larger than the normalized

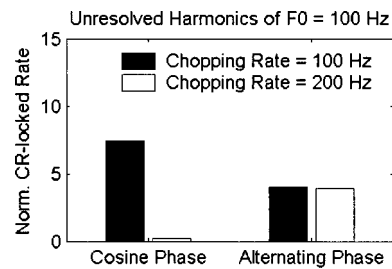


FIG. 8. Normalized CR-locked rate averaged across two Chop-S populations (one with a CR of 100 Hz, solid bars, and one with a CR of 200 Hz, open bars) in response to spectrally unresolved harmonics 16 to 26 of an F_0 of 100 Hz. Harmonics were added in either cosine phase or alternating phase. The unresolved cosine-phase harmonics produce a strong, normalized CR-locked rate in the Chop-S population tuned to 100 Hz and almost no normalized CR-locked rate in the Chop-S population tuned to 200 Hz. The unresolved alternating-phase harmonics produce normalized CR-locked rate of similar amount in both the 100-Hz Chop-S population and in the 200-Hz population.

CR-locked rate caused by high harmonics. This finding is in agreement with Houtsma and Smurzynski (1990), who showed that F_0 difference limens (a psychophysical measure of pitch strength) are better for low harmonics than for high harmonics.

Moreover, for high harmonics, the simulations show a smaller normalized CR-locked rate for Schroeder-phase complexes than for sine-phase complexes. Several repetitions of the simulations showed that this difference is highly significant. Again, this is in agreement with Houtsma and Smurzynski (1990) because they showed that pitch strength (quantified as F_0 difference limens) was slightly stronger when high harmonics were added in sine phase than when high harmonics were added in Schroeder phase.

6. Simulation 4: The influence of phase on the pitch of harmonic complexes

Patterson (1987) and Shackleton and Carlyon (1994) investigated the pitch of harmonic complexes where the harmonics were added either in cosine phase or in alternating phase (odd-order harmonics have a phase of 0° ; even-order harmonics have a phase of 90°) as a function of whether their listeners could resolve the individual spectral components. They found that when harmonics were added in cosine phase, the perceived pitch corresponds to F_0 regardless of whether the harmonics were spectrally resolved or unresolved. However, when the harmonics were added in alternating phase, the perceived pitch corresponds to F_0 only for spectrally resolved harmonics. Spectrally unresolved harmonics added in alternating phase produce a pitch that matches to $2F_0$.

For the simulation, stimuli consisted of harmonics 16 to 26 of a harmonic complex with an F_0 of 100 Hz and added in either cosine phase or alternating phase. Neural responses were obtained for two Chop-S populations, one with a CR equal to F_0 and one with a CR equal to $2F_0$.

Normalized CR-locked rate averaged across each Chop-S population is shown in Fig. 8. In the Chop-S population tuned to $F_0 = 100$ Hz (solid bars), the normalized CR-locked rate is stronger for the unresolved cosine-phase harmonics than for the unresolved alternating-phase harmonics.

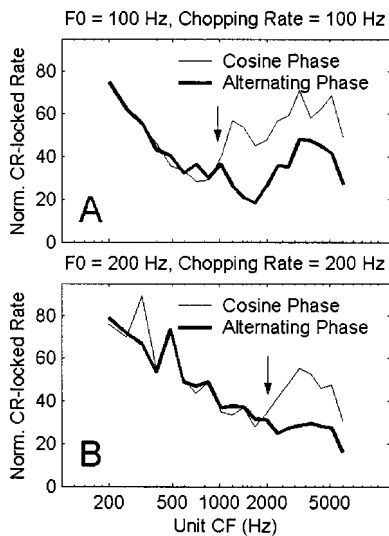


FIG. 9. Normalized CR-locked rate as a function of unit CF for harmonic complexes added in either cosine phase (fine lines) or alternating phase (strong lines). The stimulus F_0 and the population CR were either 100 Hz (A) or 200 Hz (B). Note that the normalized CR-locked rate to spectrally resolved harmonics is phase independent. The CF at which neural responses start to become phase sensitive (arrows) is at a fixed harmonic number (about 10), and not at a fixed frequency.

In the Chop-S population tuned to $F_0 = 200$ Hz (open bars), the unresolved cosine-phase harmonics generate no normalized CR-locked rate, whereas the alternating-phase stimuli generate a normalized CR-locked rate that almost equals that for the Chop-S population tuned to 100 Hz.

Shackleton and Carlyon (1994) also showed that the perceived pitch does not depend on the absolute frequency region (low vs high frequency) but on the harmonic number (low vs high harmonics). To verify, if this was the case for the Chop-S model we calculated the normalized CR-locked rate for cosine-phase complexes and alternating-phase complexes with F_0 's of 100 and 200 Hz with both resolved and unresolved harmonics. The normalized CR-locked rate of a Chop-S population tuned to 100 Hz in response to a 100-Hz harmonic complex as a function of unit CF is shown in Fig. 9(A). Chop-S responses are phase independent up to about 1000 Hz (arrow). The same plot is shown for an F_0 of 200 Hz and a Chop-S population tuned to 200 Hz in Fig. 9(B). With this F_0 , the phase sensitivity starts at about 2000 Hz, i.e., at about the same harmonic number (10), not at the same frequency.

These simulations show that the Chop-S model is insensitive to the phase of spectrally resolved harmonics and it is sensitive to the phase of unresolved harmonics. Thus, it produces predictions of pitch in qualitative agreement with the psychophysical data of Shackleton and Carlyon (1994).

7. Simulation 5: Inharmonic pitch shifts and pitch ambiguity

The pitch shifts and pitch ambiguity resulting from inharmonic complexes (consisting of equally spaced but inharmonic components) have strongly influenced modern theories of pitch perception. Chop-S units have classically been regarded as sensitive to the waveform envelope (modulation

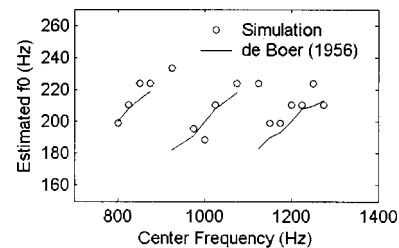


FIG. 10. Simulation of pitch shifts produced by inharmonically shifted complexes. The spacing of the five harmonics was always 200 Hz. The lines show the pitches matched after de Boer (1956) as a function of the frequency of the central component (the carrier). The open circles indicate the CR of that Chop-S population which produces the strongest, normalized CR-locked rate in response to the complex.

sensitivity: Frisina *et al.* 1990a,b). However, the envelope of the waveform of a harmonic complex is not affected by inharmonic shifts. Thus, the pitches evoked by inharmonic complexes appear to represent a special challenge for a model of pitch based exclusively on Chop-S units.

The first simulation is an attempt to reproduce the data obtained by de Boer (1956). He asked listeners to match the pitch of a carrier tone amplitude modulated with the sum of a 200- and 400-Hz tone. This results in a harmonic complex with five harmonics and a harmonic spacing of 200 Hz. The overall position of the five harmonics depends exclusively on the frequency of the carrier tone. When the frequency is an integer multiple of 200 Hz, the resulting complex is harmonic; otherwise, it is inharmonically shifted by the signed difference of the carrier frequency and the nearest multiple of 200 Hz. The matching results by de Boer (1956) indicated that when a shift of -80 Hz was applied, listeners matched a pitch of about 173 Hz. With increasing shift, the pitch match increased and reached a value of about 214 Hz for an inharmonic shift of $+80$ Hz. Moreover, the pitch matches were circular in that a further increase of pitch shift to $+100$ Hz and beyond resulted in a jump of the perceived pitch from about 230 Hz back down to 170 Hz. The magnitude of the pitch shifts tends to decrease with increasing harmonic number of the carrier frequency (see also Moore and Moore, 2003).

The simulation paradigm investigates the representation of such an inharmonic, five-component complex in 11 discrete Chop-S populations with CRs between 170 and 235 Hz. Inharmonic shifts matched those produced by de Boer (1956). The F_0 was 200 Hz. The normalized CR-locked rate was calculated for each Chop-S population in response to an inharmonic complex with each of the inharmonic shifts. Figure 10 shows the CR of that Chop-S population which produced the strongest normalized CR-locked rate for each inharmonic shift. Simulation results (open circles) are plotted together with the results of de Boer (1956).

Yost (1997) investigated the pitch of IRN when the IRN gain is negative (corresponding to a delay-and-subtract algorithm instead of delay-and-add). A negative gain results in a shift of all spectral peaks in the IRN stimulus by a frequency corresponding to half the delay reciprocal, $0.5/d$. This results in a stimulus spectrum similar to an inharmonic complex with a frequency shift equal to half the f_0 . Yost (1997) showed that the pitch of IRN with negative gain depends not

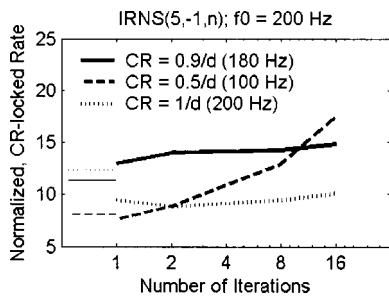


FIG. 11. Simulated neural correlate of inharmonic pitch shifts with IRN stimuli: The stimulus is IRNS(5, -1, n), with n ranging from 1 to 16 iterations in doublings. Normalized CR-locked rate averaged over a population of Chop-S units is shown with a population CR of $0.9/d$ (180 Hz), $0.5/d$ (100 Hz), and $1/d$ (200 Hz). IRNS with negative gain produces a pitch of $0.9/d$ or $1.1/d$ for a low number of iterations and a pitch equal to $0.5/d$ for a high number of iterations. In qualitative agreement with this perception, normalized CR-locked rate for a Chop-S population tuned to $0.9/d$ is strongest for a low number of iterations but, with increasing number of iterations, the Chop-S population tuned to $0.5/d$ starts to produce the strongest, normalized CR-locked rate. Note, however, that the change in perceived pitch occurs typically between 2 and 4 iterations, whereas the crossover of the rate functions occurs between 8 and 16 iterations. The reason for the deviation between the experimental data and the simulations probably lies in the different values of the normalized CR-locked rate in response to white-noise stimuli (horizontal lines on the left). Compensation of these different offsets would result in an improved fit to the experimental data.

only on the IRN gain but also on the number of iterations. With one or two iterations, listeners matched pitches corresponding to $1.1/d$ or $0.9/d$, whereas with a high number of iterations, listeners tended to match a pitch equal to $0.5/d$. Simulations were performed using IRNS(5, -1, n) where n , the number of iterations, is variable. These stimuli would produce a pitch corresponding to 200 Hz if a positive gain were used. Populations of Chop-S units with three different CRs were used. These CRs were equal to $0.9/d$, $0.5/d$, and $1/d$. Figure 11 shows the normalized CR-locked rate averaged across each population for the three populations as a function of the number of iterations.

For a low number of iterations, the Chop-S population with a CR at $0.9/d$ shows the strongest normalized CR-locked rate (solid line). With increasing number of iterations, the normalized CR-locked rate of the Chop-S population with a CR of $0.5/d$ grows fastest. For 16 iterations, the Chop-S population with a CR of $0.5/d$ shows the strongest normalized CR-locked rate. The Chop-S population with a CR of $1/d$ shows only a weak, normalized CR-locked rate. This latter population produces the strongest normalized CR-locked rate when the gain is positive (cf. Fig. 5). Note that for this population, the normalized CR-locked rate is below that obtained for white-noise (aperiodic) stimulation (short horizontal lines on the left).

8. Simulation 6: The segregation of concurrent pitches

Assmann and Paschall (1998) investigated the extent to which listeners can match the pitch of each of two vowel-shaped harmonic complexes as a function of the separation of the f_0 's of the two complexes. They found that when the f_0 difference of the vowels was below four semitones (24%), listeners matched a single pitch to the complex con-

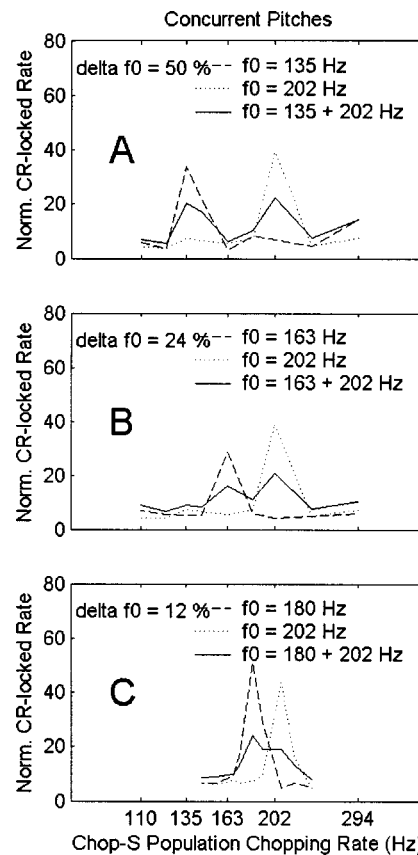


FIG. 12. Segregation of concurrent pitches in Chop-S populations: Normalized CR-locked rate responses are shown as a function of the population CR when stimulated with either of two harmonic complexes or their sum. The dashed lines show responses to the lower of the two f_0 's; dotted lines show responses to the higher of the two f_0 's. Solid lines show the responses to the sum of the two complexes. Note that when the f_0 deviation is 50% (A) or 24% (B) the two pitches are well resolved (solid lines). With a 12% f_0 deviation (C) a bimodal distribution of the estimated pitches is not found, although the spacing of the population CRs was halved compared to (A) and (B).

sisting of two f_0 's. When the f_0 difference was four semitones, the listeners provided bimodal pitch matches, i.e., sometimes they matched the lower pitch and sometimes they matched the higher pitch. This performance indicates the listeners' ability to "hear out" each of the two f_0 's. The following simulation investigates the extent to which the Chop-S model is capable to segregate two simultaneously presented harmonic complexes. The stimuli consisted of CPH with either a lower f_0 or a higher f_0 or the sum of the two CPHs. Averaged, normalized CR-locked rates were obtained for populations of Chop-S units with closely spaced CRs in the range surrounding the two f_0 's. Simulation results are shown in Fig. 12. For an f_0 difference of 50% or 24% (A) and (B), the simulations reveal a clear peak in response to each single complex and a bimodal distribution in response to the sum of the two CPHs. With an f_0 difference of 12% [Fig. 12(C)], there is no clear bimodal distribution despite the fact that the spacing of the CRs was halved compared to Figs. 12(A) and (B).

III. DISCUSSION

The simulations demonstrate that the normalized CR-locked rate measure applied to a model of VCN Chop-S units

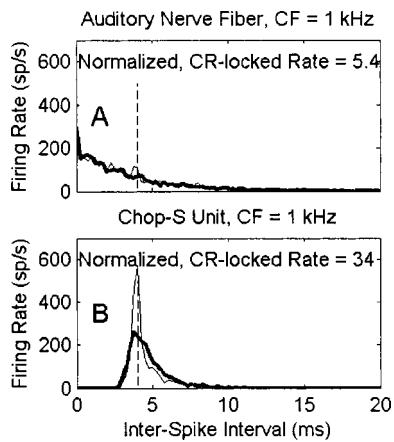


FIG. 13. Comparison of the interspike interval histograms of a simulated auditory-nerve fiber (A) and a Chop-S unit (B). Stimuli were either white noise (strong lines) or a random-phase harmonic complex (RPH) with an F_0 of 250 Hz (fine lines). Note that both in the simulated auditory-nerve fiber and in the simulated Chop-S unit, the RPH leads to a redistribution of interspike intervals in favor of the intervals equal to the stimulus period (vertical dashed lines), compared to white-noise stimulation. However, the degree of this redistribution, quantified as normalized CR-locked rate, is much higher in the Chop-S unit compared to the auditory-nerve fiber.

can reflect the strength of pitch percepts observed in a wide range of psychophysical studies. The demonstrations present a *prima facie* case that Chop-S units may play a role in transmitting pitch-related information to higher stations in the nervous system. The model can simulate responses to a variety of pitch phenomena. It is able to predict changes of the IRN pitch strength with increasing number of iterations, the dominance region of pitch, changes of pitch strength which depend on the phase relations between the harmonics, changes of the pitch when harmonics are added in cosine phase compared to alternating phase, the pitches of inharmonic complexes and the pitches of IRN with negative gain (which also produces inharmonic shifts), and the pitches of concurrent harmonic complexes, and the pitch and pitch strength of SAM noise.

Chop-S units have traditionally been characterized as followers of signal envelope. This modeling study suggests that they may have an important role to play in transmitting information concerning stimulus fine structure as well. Most physiological experiments have used high-frequency (CF) carrier stimuli subject to low-frequency modulation. Few if any studies have looked at synchronization to unmodulated pure tones other than CF tones. However, psychophysical pitch studies have shown that it is the resolved (and therefore only weakly modulated or unmodulated) sinusoidal components of harmonic complexes that contribute most to the strength of the pitch percept. The model predicts that a strong normalized CR-locked rate is observed in response to resolved harmonics when the frequency of the harmonic is equal to the intrinsic CR or equal to an integer multiple of the CR of the unit.

The latter is a crucial feature of Chop-S units: Imagine a Chop-S unit with a CF of 750 Hz and a CR of 250 Hz being stimulated by a harmonic complex with an f_0 of 250 Hz. Spectrally, this unit is tuned to the third harmonic of the complex. As this harmonic is spectrally well resolved, the

auditory-nerve representation of this third harmonic will show very little amplitude modulation corresponding to f_0 . Instead, the Chop-S unit will receive an auditory-nerve input with a dominant period (summed across all auditory-nerve fibers converging on the Chop-S unit) of 1.33 ms, the reciprocal of 750 Hz. The important function of the Chop-S unit is that it is capable to skip input spikes and lock onto an integer multiple of the 1.33-ms period. It will choose a period which is most suitable for its CR. Thus, the Chop-S unit will fire with a highly regular interspike interval of 4 ms, which means firing on every third input maximum. Thus, a Chop-S unit is capable to “interpret” the 750-Hz not as a 750-Hz pure tone but as being the third harmonic of an f_0 of 250 Hz even if the unit has no indication that the 250-Hz component is actually present.

The simulations described above indicate that assemblies of Chop-S units may supply a useful neural representation of pitch information. However, Chop-S units are just one of a variety of cell types found in the CN. Primarylike units have been shown to transmit the highest temporal precision of neural discharge. Their firing characteristics are very similar to those of auditory-nerve fibers, which have been shown to represent a reliable pitch code in their all-order interspike intervals (Horst *et al.*, 1986; Cariani and Delgutte, 1996a,b). The special advantage of Chop-S units is that they are narrowly tuned in terms of which fundamental frequencies produce a strong, normalized CR-locked rate. Other unit types are typically broadly tuned to f_0 , if at all (Winter *et al.*, 2001). This is illustrated in Fig. 13, where ISI histograms and normalized CR-locked rate of a simulated Chop-S unit (B) is compared to its auditory-nerve input (A). As outlined above, this input consists of the aggregated firing activity of 15 auditory-nerve fibers. Responses are shown to white noise (strong lines) and an RPH with $F_0 = 250$ Hz (fine lines). Note the redistribution of ISI in favor of the 4-ms period (vertical dashed line) in both unit types. However, as a consequence of its bandpass temporal tuning, the Chop-S unit provides a much higher normalized CR-locked rate compared to the auditory nerve. Thus, the temporal tuning of a Chop-S unit to a narrow range of ISIs results in a twofold benefit for the system: first, it increases its sensitivity to changes in the ISI distribution when the ISI is in this tuning range. Second, the temporal tuning of Chop-S units may serve as a physiological basis for the identification of common harmonicity across the many components of a harmonic complex, even though they are spread across a range of frequencies.

Chop-S units show less modulation gain in response to envelope-modulated stimuli than other unit types like onset units (Kim *et al.*, 1990). However, considering the fact that the low, weakly modulated harmonics of a harmonic complex contribute most to pitch strength, it is questionable whether sensitivity to envelope modulation is the most important physiological parameter for neural pitch extraction. The current study shows that Chop-S units reveal a good sensitivity to the period of the low, weakly modulated harmonics. As such, they may represent the first stage of sensitivity to pitch.

Shortcomings of the model

The current model depends on the assumption that there is a population of Chop-S units for each f_0 encountered. The existence region of the f_0 related pitch of harmonic complexes spans from about 30 Hz to about 1400 Hz. However, Chop-S units are not frequently seen with CRs greater than 500 Hz or smaller than 100 Hz. Thus, at present, the Chop-S model cannot account for these pitches. Concerning higher f_0 's, it is possible that Chop-S units show synchronization to a subharmonic of f_0 when this subharmonic fits the CR of the unit. This is reflected in Fig. 2, where an increased normalized CR-locked rate at the natural CR will be seen in response to these f_0 multiples. This model behavior might be interpreted as a neural correlate of octave confusions, as they are often encountered in pitch judgments.

This study has not considered pitches based upon the many different types of dichotic signals that generate a pitch sensation (Culling, 2000; Culling *et al.*, 1998a,b). At the level of the CN, the neural representation of these stimuli will be no different than for an aperiodic stimulus because of the monaural nature of the CN. Thus, the pitch sensation evoked by dichotic-pitch stimuli cannot be mediated by the processing of Chop-S units. However, at the next prospective stage of the model under development, the Chop-S units project to coincidence detector units in the inferior colliculus. At this stage, there is pronounced binaural interaction and it is possible that binaural properties of coincidence detector units may be involved in the perception of dichotic pitches. The behavior of a model fully developed to the level of the inferior colliculus remains to be tested.

Another potential problem for the model concerns the firing rates of sustained chopper units at very low signal levels. The model assumes that chop-S units have low thresholds and low saturation thresholds. Above saturation levels, the firing rate is roughly constant (regardless of level) and will synchronize to acoustic events at that frequency. At low signal levels, the rate/constancy assumption clearly does not hold and the pitch percept should be affected by reductions in rate. Unfortunately, it is difficult to estimate from the animal data the levels at which this would become an issue. We also know relatively little about how low pitch percepts behave close to absolute threshold. Nevertheless, this issue is potentially problematic and will need to be kept in view in the future.

May *et al.* (1998) described responses of chopper units to vowel sounds where the position of the vowel formants was varied systematically. They showed that a variation in formant position resulted in a change of the units' rate responses regardless of the overall sound-pressure level of the presented vowel. This finding represents an interesting challenge for the current suggestion that temporal response characteristics of Chop-S units play an important role in pitch perception.

First, however, it must be pointed out that the suggested role in pitch perception applies to relatively low pitches with fundamental frequencies below about 500 Hz. The role of Chop-S units in the encoding of spectral envelope features applies to formants that are typically located at frequencies of 1 kHz and above. Thus, the encoding of the low harmon-

ics numbers (which are the most important for pitch perception) is not likely to be affected by the unit's sensitivity to spectral envelope features. For higher harmonic numbers, however, it must be assumed that there will be interference between a Chop-S unit's rate-response changes in response to changes in the spectral envelope and the unit's sensitivity to the fundamental frequency of these high harmonics. Such interference, however, is also observed perceptually: It is difficult for listeners to estimate the low pitch of higher-order harmonics independently of spectral envelope and overall sound level. One example is the study by Walliser (1969), where it was shown that the pitch of a 300-Hz harmonic complex octave band filtered around 4000 Hz depends significantly on sound-pressure level. Another example is the occurrence of edge pitches.

An extensive study of Chop-S units with iterated rippled noise, where the spectral envelope in the spectrally unresolved frequency region is flat, has shown that (for these "nonvowel" sounds) Chop-S units reach their rate saturation at relatively low sound-pressure levels and they preserve their sensitivity to periodic sounds in rate saturation (Wiegube and Winter, 2001).

In summary, the temporal tuning of a Chop-S unit to a narrow range of ISIs results in a twofold benefit for the system: first, it increases its sensitivity to changes in the ISI distribution when the ISI is in this tuning range. Second, the temporal tuning of Chop-S units may serve as a physiological basis for the identification of common harmonicity across the many components of a harmonic complex, even though they are spread across a range of frequencies.

ACKNOWLEDGMENTS

We would like to thank Ian Winter, Alain de Cheveigne, and an unknown reviewer for very helpful comments on earlier versions of this manuscript. We would also like to thank Roy Patterson and Ian Winter for many fruitful discussions on pitch processing and the cochlear-nucleus physiology.

- Abeles, M. (1982). "Quantification, smoothing, and confidence limits for single-units' histograms," *J. Neurosci. Methods* **5**, 317–325.
- Assmann, P. F., and Paschall, D. D. (1998). "Pitches of concurrent vowels," *J. Acoust. Soc. Am.* **103**, 1150–1160.
- Blackburn, C. C., and Sachs, M. B. (1989). "Classification of unit types in the anteroventral cochlear nucleus: PST histograms and regularity analysis," *J. Neurophysiol.* **62**, 1303–1329.
- Burns, E. M., and Viemeister, N. F. (1981). "Played-again Sam: Further observation on the pitch of amplitude-modulated noise," *J. Acoust. Soc. Am.* **70**, 1655–1660.
- Cariani, P. A., and Delgutte, B. (1996a). "Neural correlates of the pitch of complex tones. I. Pitch and pitch salience," *J. Neurophysiol.* **76**, 1698–1716.
- Cariani, P. A., and Delgutte, B. (1996b). "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," *J. Neurophysiol.* **76**, 1717–1734.
- Culling, J. F. (2000). "Dichotic pitches as illusions of binaural unmasking. III. The existence region of the Fourcin pitch," *J. Acoust. Soc. Am.* **107**, 2201–2208.
- Culling, J. F., Marshall, D. H., and Summerfield, A. Q. (1998a). "Dichotic pitches as illusions of binaural unmasking. II. The Fourcin pitch and the dichotic repetition pitch," *J. Acoust. Soc. Am.* **103**, 3527–3539.
- Culling, J. F., Summerfield, A. Q., and Marshall, D. H. (1998b). "Dichotic pitches as illusions of binaural unmasking. I. Huggins' pitch and the binaural edge pitch," *J. Acoust. Soc. Am.* **103**, 3509–3526.

- de Boer, E. (1956). "On the 'residue' in hearing," Doctoral dissertation, University of Amsterdam.
- de Cheveigne, A. (1998). "Cancellation model of pitch perception," *J. Acoust. Soc. Am.* **103**, 1261–1271.
- Ferragamo, M. J., Golding, N. L., and Oertel, D. (1998). "Synaptic inputs to stellate cells in the ventral cochlear nucleus," *J. Neurophysiol.* **79**, 51–63.
- Frisina, R. D., Smith, R. L., and Chamberlain, S. C. (1990a). "Encoding of amplitude modulation in the gerbil cochlear nucleus. I. A hierarchy of enhancement," *Hear. Res.* **44**, 99–122.
- Frisina, R. D., Smith, R. L., and Chamberlain, S. C. (1990b). "Encoding of amplitude modulation in the gerbil cochlear nucleus. II. Possible neural mechanisms," *Hear. Res.* **44**, 123–141.
- Ghoshal, S., Kim, D. O., and Northrop, R. B. (1992). "Amplitude-modulated tone encoding behavior of cochlear nucleus neurons: Modeling study," *Hear. Res.* **58**, 153–165.
- Glasberg, B. R., and Moore, B. C. J. (2002). "A model of loudness applicable to time-varying sounds," *J. Audio Eng. Soc.* **50**, 331–342.
- Goldstein, J. L. (1973). "An optimum processor theory for the central formation of the pitch of complex tones," *J. Acoust. Soc. Am.* **54**, 1496–1516.
- Hewitt, M. J., and Meddis, R. (1994). "A computer model of amplitude-modulation sensitivity of single units in the inferior colliculus," *J. Acoust. Soc. Am.* **95**, 2145–2159.
- Hewitt, M. J., Meddis, R., and Shackleton, T. M. (1992). "A computer model of a cochlear-nucleus stellate cell: Responses to amplitude-modulated and pure-tone stimuli," *J. Acoust. Soc. Am.* **91**, 2096–2109.
- Horst, J. W., Javel, E., and Farley, G. R. (1986). "Coding of spectral fine structure in the auditory nerve. I. Fourier analysis of period and interspike interval histograms," *J. Acoust. Soc. Am.* **79**, 398–416.
- Houtsma, A. J. M., and Smurzynski, J. (1990). "Pitch identification and discrimination for complex tones with many harmonics," *J. Acoust. Soc. Am.* **87**, 304–310.
- Kim, D. O., Sirianni, J. G., and Chang, S. O. (1990). "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.
- Lai, Y. C., Winslow, R. L., and Sachs, M. B. (1994). "A model of selective processing of auditory-nerve inputs by stellate cells of the antero-ventral cochlear nucleus," *J. Comput. Neurosci.* **1**, 167–194.
- Licklider, J. C. R. (1951). "A duplex theory of pitch perception," *Experientia* **7**, 128–133.
- Licklider, J. C. R. (1959). "Three auditory theories," in *Psychology, A Study of A Science*, edited by S. Koch (McGraw-Hill, New York).
- Lopez-Poveda, E. A., and Meddis, R. (2001). "A human nonlinear cochlear filterbank," *J. Acoust. Soc. Am.* **110**, 3107–3118.
- Lyon, R. F. (1984). "Computational models of neural auditory processing," in *Proc IEEE ICASSP* 84.3.
- MacGregor, R. J. (1987). *Neural and Brain Modeling* (Academic, San Diego).
- May, B. J., Prell, G. S., and Sachs, M. B. (1998). "Vowel representations in the ventral cochlear nucleus of the cat: Effects of level, background noise, and behavioral state," *J. Neurophysiol.* **79**, 1755–1767.
- Meddis, R., and Hewitt, M. J. (1991a). "Virtual pitch and phase sensitivity of a computer model of the auditory periphery. I. Pitch identification," *J. Acoust. Soc. Am.* **89**, 2866–2882.
- Meddis, R., and Hewitt, M. J. (1991b). "Virtual pitch and phase sensitivity of a computer model of the auditory periphery. II. Phase sensitivity," *J. Acoust. Soc. Am.* **89**, 2883–2894.
- Meddis, R., O'Mard, L. P., and Lopez-Poveda, E. A. (2001). "A computational algorithm for computing nonlinear auditory frequency selectivity," *J. Acoust. Soc. Am.* **109**, 2852–2861.
- Moore, G. A., and Moore, B. C. (2003). "Perception of the low pitch of frequency-shifted complexes," *J. Acoust. Soc. Am.* **113**, 977–985.
- Patterson, R. D. (1987). "A pulse ribbon model of monaural phase perception," *J. Acoust. Soc. Am.* **82**, 1560–1586.
- Plomp, R. (1976). *Aspects of Tone Sensation: A Psychophysical Study* (Academic, London, New York, San Francisco).
- Ritsma, R. J. (1967). "Frequencies dominant in the perception of the pitch of complex sounds," *J. Acoust. Soc. Am.* **42**, 191–198.
- Schouten, J. F. (1970). "The residue revisited," in *Frequency Analysis and Periodicity Detection in Hearing*, edited by R. Plomp and G. F. Smoorenburg (Sijthoff, Leiden), pp. 39–58.
- Shackleton, T. M., and Carlyon, R. P. (1994). "The role of resolved and unresolved harmonics in pitch perception and frequency modulation discrimination," *J. Acoust. Soc. Am.* **95**, 3529–3540.
- Shamma, S., and Klein, D. (2000). "The case of the missing pitch templates: How harmonic templates emerge in the early auditory system," *J. Acoust. Soc. Am.* **107**, 2631–2644.
- Shofner, W. P. (1991). "Temporal representation of rippled noise in the anteroventral cochlear nucleus of the chinchilla," *J. Acoust. Soc. Am.* **90**, 2450–2466.
- Shofner, W. P. (1999). "Responses of cochlear nucleus units in the chinchilla to iterated rippled noises: Analysis of neural autocorrelograms," *J. Neurophysiol.* **81**, 2662–2674.
- Sumner, C. J., Lopez-Poveda, E. A., O'Mard, L. P., and Meddis, R. (2002). "A revised model of the inner-hair cell and auditory-nerve complex," *J. Acoust. Soc. Am.* **111**, 2178–2188.
- Terhardt, E. (1979). "Calculating virtual pitch," *Hear. Res.* **1**, 155–182.
- Walliser, K. (1969). "Zusammenhänge zwischen dem Schallreiz und der Periodentonhöhe," *Acustica* **21**, 319–329.
- Wiegrefe, L., and Winter, I. M. (2001). "Temporal representation of iterated rippled noise as a function of delay and sound level in the ventral cochlear nucleus," *J. Neurophysiol.* **85**, 1206–1219.
- Winter, I. M., Wiegrefe, L., and Patterson, R. D. (2001). "The temporal representation of the delay of iterated rippled noise in the ventral cochlear nucleus of the guinea-pig," *J. Physiol. (London)* **537**, 553–566.
- Yost, W. A. (1996a). "Pitch of iterated rippled noise," *J. Acoust. Soc. Am.* **100**, 511–518.
- Yost, W. A. (1996b). "Pitch strength of iterated rippled noise," *J. Acoust. Soc. Am.* **100**, 3329–3335.
- Yost, W. A. (1997). "Pitch strength of iterated rippled noise when the pitch is ambiguous," *J. Acoust. Soc. Am.* **101**, 1644–1648.
- Young, E. D., Robert, J. M., and Shofner, W. P. (1988). "Regularity and latency of units in ventral cochlear nucleus: Implications for unit classification and generation of response properties," *J. Neurophysiol.* **60**, 1–29.