Pitch discrimination of patterned electric stimulation[a)]

Hongbin Chen,[b)] Yumi Christine Ishihara, and Fan-Gang Zeng[c)]
Hearing and Speech Research Laboratory, Departments of Anatomy and Neurobiology, Biomedical Engineering, Cognitive Sciences and Otolaryngology—Head and Neck Surgery, University of California, Irvine, California 92697

(Received 24 January 2005; revised 21 April 2005; accepted 26 April 2005)

One reason for the poor pitch performance in current cochlear-implant users may be the highly synchronized neural firing in electric hearing that lacks stochastic properties of neural firing in normal acoustic hearing. This study used three different electric stimulation patterns, jittered, probabilistic, and auditory-model-generated pulses, to mimic some aspects of the normal neural firing pattern in acoustic hearing. Pitch discrimination was measured at standard frequencies of 100, 250, 500, and 1000 Hz on three Nucleus-24 cochlear-implant users. To test the utility of the autocorrelation pitch perception model in electric hearing, one, two, and four electrodes were stimulated independently with the same patterned electric stimulation. Results showed no improvement in performance with any experimental pattern compared to the fixed-rate control. Pitch discrimination was actually worsened with the jittered pattern at low frequencies (125 and 250 Hz) than that of the control, suggesting that externally introduced stochastic properties do not improve pitch perception in electric stimulation. The multiple-electrode stimulation did not improve performance but did not degrade performance either. The present results suggest that both "the right time and the right place" may be needed to restore normal pitch perception in cochlear-implant users. [DOI: 10.1121/1.1937228]

PACS number(s): 43.66.Hg, 43.66.Fe, 43.66.Ts [BLM] Pages: 338–345

I. INTRODUCTION

Pitch perception has been studied for more than 150 years but its underlying mechanisms still remain elusive (Ohm, 1843; Helmholtz, 1863). Two theories have been proposed based on temporal neural firing patterns (Wever, 1948; Siebert, 1970; Goldstein and Srulovicz, 1977) and the place of excitation in the cochlea (Zwicker, 1956; Henning, 1967). In the temporal theory, the pitch of a stimulus is determined by the interval between two adjacent neural firings. Strictly speaking, a pure temporal model is independent of the place of excitation in the cochlea. In the place theory, the pitch is determined by the place of excitation in the cochlea, although the exact meaning of the "place" is still debatable with the sharp apical edge, rather than the excitation peak, in the excitation pattern being a more likely code for the place pitch (Chatterjee and Zwislocki, 1997).

To account for pitch of complex stimuli, such as residual pitch and periodicity pitch (Ritsma, 1962; Schouten *et al.*, 1962; Ritsma, 1963), an autocorrelation model has been proposed, in which pitch is interpreted as the peak in the autocorrelation function of the neuron response to the stimuli (Licklider, 1951). In a recently extended model (Meddis and Hewitt, 1991), complex tones were passed through a bank of gammatone filters followed by a hair-cell simulator which converted the mechanical motion of the basilar membrane into spike trains propagated along the auditory nerve fiber. An autocorrelation function on the function (ACF) relating probability of neural firing to time was generated for each auditory nerve fiber. A summary autocorrelation function was generated by summing up the autocorrelation functions from these individual fibers. Pitch was determined by the largest peak in the summary autocorrelation function. The computational model of Meddis *et al.* has successfully predicted several of the classical pitch phenomena, such as the missing fundamental, ambiguous pitch, and inharmonicity.

Variations of the temporal model also exist. For example, several researchers proposed the "first-order" theory (Srulovicz and Goldstein, 1983; Kaernbach and Demany, 1998), suggesting that the auditory system is sensitive only to the first-order intervals, i.e., the interval between two adjacent spikes, between successive spikes in the neural firing. In a stream of neural firing with different intervals, pitch is derived from the longest first-order interval. Carlyon *et al.* (2002) provided evidence for an even stronger first-order model, in which only the interspike intervals contribute to the temporal pitch percept, with the longest first-order intervals receiving more weights than the short first-order intervals.

Based on these temporal pitch theories, the pitch strength ought to be more salient in electric hearing than acoustic hearing, because the neurophysiological studies have found that neural spikes are highly synchronized to electric stimuli (Kiang and Moxon, 1972; Hartmann *et al.*, 1984; Dynes and Delgutte, 1992; Litvak *et al.*, 2001). However, behavioral pitch discrimination data in electric hearing do not support this prediction. Compared to the sharp pitch discrimination in normal-hearing listeners who can detect 1% or less difference for frequencies up to 4000 Hz (Harris, 1952; Moore, 1973; Wier *et al.*, 1977; Nelson *et al.*, 1983),

[a)]Portions of this work were presented at the 2004 American Auditory Society annual meeting, Scottsdale, AZ, 7–9 March 2004.
[b)]Electronic mail: hchen@uci.edu
[c)]Electronic mail: fzeng@uci.edu

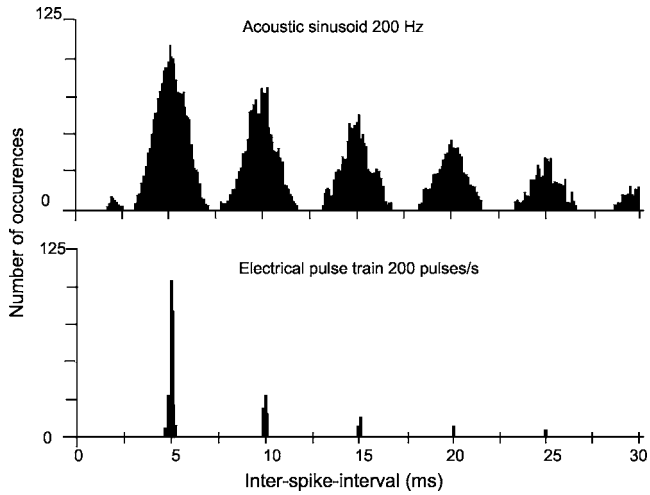

FIG. 1. The histogram of the interspike intervals for a single neuron fiber in response to a pure tone with frequency of 200 Hz (top panel) and a biphasic pulse train with stimulation rate of 200 pulses/s (reproduced from Javel).

cochlear-implant listeners can only detect 10%–25% differences for frequencies up to 500 Hz and typically cannot discriminate any pitch difference for frequencies higher than 500 Hz (Bilger, 1977; Eddington *et al.*, 1978; Shannon, 1983a; Carlyon *et al.*, 2002; Zeng, 2002; Chen and Zeng, 2004). The poor pitch discrimination is likely responsible for the cochlear-implant users' extreme difficulty in speech recognition in noise (Friesen *et al.*, 2001; Garnham *et al.*, 2002), music appreciation (Gfeller and Lansing, 1991; Gfeller *et al.*, 1997; Pijl, 1997; Kong *et al.*, 2004), and tonal language understanding (Wei *et al.*, 2004).

One apparent reason for poor pitch discrimination in cochlear implant users may be the lack of sharp frequency tuning in electric hearing. Neurophysiological data showed no tuning at all to electric stimulus frequency as long as the stimulus was delivered to the same pair of electrodes (Kiang and Moxon, 1972; Hartmann *et al.*, 1984). The place code can be only crudely reproduced or represented by a limited number of electrodes placed in different sites of the cochlea. More often than not, the number of independent electrodes is further reduced by electrical current filed interaction between electrodes (Shannon, 1983b; Fishman *et al.*, 1997).

The discrepancy between the temporal pitch model prediction and the behavioral data may be due to the significant difference in stochastic neural firing between acoustic and electric stimulation (Rose *et al.*, 1967; Javel *et al.*, 1987; van den Honert and Stypulkowski, 1987). Figure 1 shows the interspike-interval (ISI) histogram for neural responses to acoustic and electric stimuli. The top panel shows the ISI histogram for a single auditory nerve fiber in response to a pure tone at 200 Hz, while the bottom panel shows the ISI histogram in response to a biphasic pulse train at 200 pulses/s. There is clearly greater variability in neural firing in response to acoustic stimulation than electric stimulation. First, at the stimulus period (5 ms) and its multiples, the neural firing has much greater standard deviation in acoustic hearing than in electric hearing. Second, the neural firing occurs randomly at many more modes (the stimulus period and its multiples) in acoustic stimulation than electric stimulation.

Restoring similar stochastic responses in electric stimulation may enhance the pitch extraction process and signal detection at threshold level in cochlear implants (Morse and Evans, 1999; Rubinstein *et al.*, 1999; Zeng *et al.*, 2000). This has been investigated by using high-rate (>2000 Hz) stimulation and adding white noise to a fixed-rate stimulus, while another possible way to introduce stochastic resonance is to temporally modulate pulse trains by a stochastic function. As it is also shown in the auditory model, pitch extraction is an analysis not only "within-channel" but also "between-channel." In the model, if each channel is independent, ACF from different channels will have the same peaks at the period of the stimuli and small peaks at different delay because of the noise. In the summary ACF, the significance of common peaks is the same as that from one channel, but other peaks are smaller because of the process of averaging. Therefore, the model predicts better performance with multiple channels if each channel is independent but has a similar pulse pattern. It is unknown whether this temporal information from different channels could be utilized for high-level pitch extraction in electric hearing.

The main questions addressed in this study were: (1) To what extent do we have to reproduce the normal temporal discharge patterns in electric stimulation to improve cochlear implant pitch discrimination? (2) Can multiple-electrode stimulation improve the performance as predicted by the autocorrelation model? (3) Does the site of stimulation matter in electric pitch discrimination? To answer the first question, we designed three types of novel electric stimulus patterns, including jittered pulses, probabilistic pulses, and auditory-model-generated pulses, which mimic either one or several aspects of the ISI pattern in response to a sinusoid in acoustic stimulation. To answer the second question, we tested pitch discrimination using single- and multiple-electrode stimulation. To answer the third question, we used different spacing between stimulating electrodes in the multiple-electrode condition. To the best of our knowledge, none of the experimental conditions have been reported in the literature.

II. METHODS

A. Subjects

One male and two female adults (S1–S3) who were postlingually deafened and implanted with Nucleus-24 devices took part in this study. The subjects ranged in age from 70 to 79 years with a mean age of 73 years and were all native speakers of American English. All of the subjects had extensive psychophysical test experience and were compensated for their participation in the study. Local IRB approval and informed consent were obtained prior to the experiments. Table I lists the detailed information of the subjects. Vowel stimuli were taken from materials recorded by (Hillenbrand *et al.*, 1995) and consonant stimuli were taken from materials created by (Turner *et al.*, 1992). The stimuli were presented to cochlear implant listeners with custom software (Robert, 1997).

TABLE I. Subject information of the three Nucleus-24 cochlear-implant users who participated in this study.

Subject	Gender	Age (yr)	Cause of deafness	Duration of implant use (yr)	Vowel recognition	Consonant recognition
S1	F	71	Fever	4	51%	54%
S2	M	79	Unknown	2	38%	51%
S3	F	70	Virus	7	64%	70%

B. Stimuli

Three pulse trains with different temporal patterns were used in this study to mimic one or more aspects of the stochastic neural firing in acoustic stimulation. As an example with the 1000 Hz standard frequency, Fig. 2 shows the constructed ISI histograms and the first 20 ms of these three pulse trains.

1. Jittered pulses

In the jittered pulses, the interpulse interval followed a Gaussian distribution with a mean at the standard frequency and a standard deviation of d. In the present experiment, d was set at 0.0, 0.1, 0.2, or 0.3. When $d=0$, the jitters were removed to produce the traditional fixed-rate pulse train. We noted that Dobie and Dillier (1985) conducted jitter discrimination on two Ineraid cochlear-implant users and found a detection threshold of about 10% at the stimulation rate of 1000 Hz. No pitch discrimination of the jittered pulses has been reported. The standard deviation was set at 0.3 in Fig. 2. As can be seen from the ISI histogram in Fig. 2, the jittered pulses produced a random distribution at the first mode (i.e., the period) that was similar to the normally produced ISI pattern in acoustic stimulation (Fig. 1).

2. Probabilistic pulses

In the probabilistic pulses, the occurrence of a pulse in the pulse train was determined by a probability p. When $p=1$, the probabilistic pulses were equivalent to a fixed-rate pulse train. When p was less than 1, the total number of pulses in the probabilistic pulses equaled the number of pulses in the fixed-rate pulse train$\times p$. The interval between two consecutive pulses was multiples of the period in the fixed-rate train. In the present study, the value of p was set at 1, 0.8, 0.5, or 0.3 (the value of 0.3 was used in Fig. 2). The probabilistic pulses produced multiple modes but no jitters in the ISI histogram (middle panel in Fig. 2).

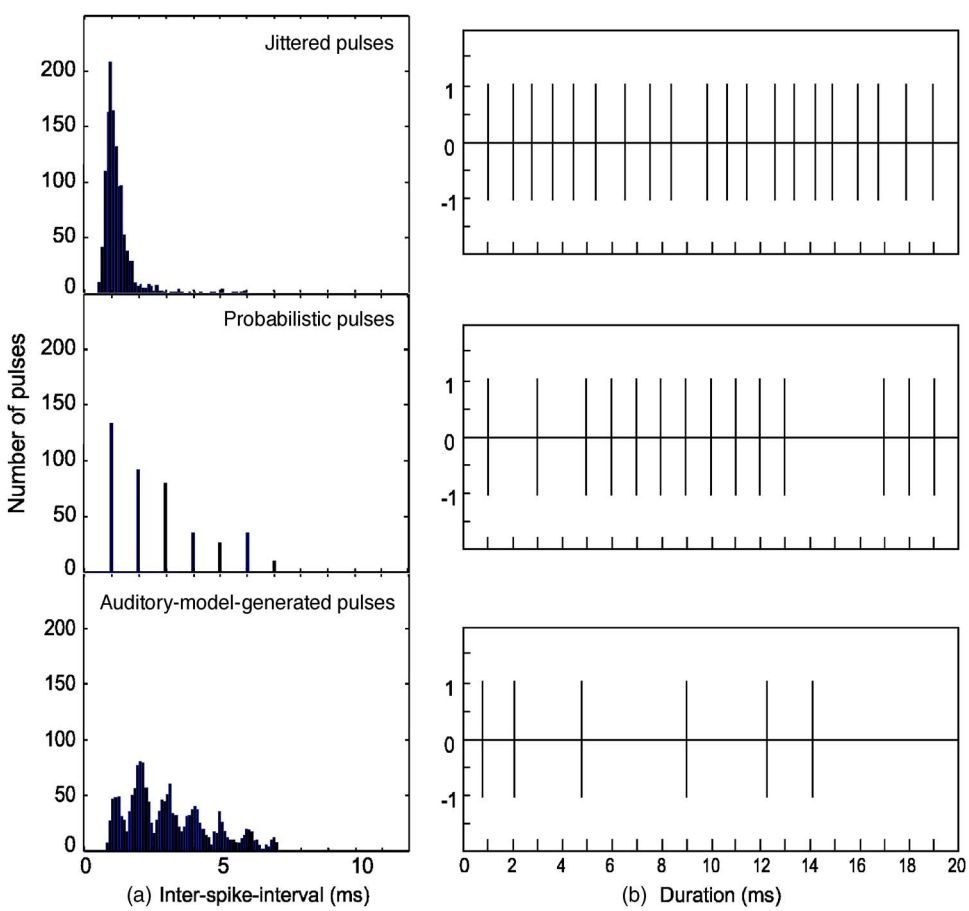

FIG. 2. (a) The interspike-interval histograms for three different types of stimuli: jittered pulses, probabilistic pulses, and auditory-model-generated pulses (from top to bottom). (b) The pulse trains for the three different types of stimuli. Only the first 20 ms of the pulse trains are shown here.

3. Auditory-model-generated pulses

The pulses were generated by Meddis's auditory model, which incorporated a basilar membrane model (DRNL) (Lopez-Poveda and Meddis, 2001), gammatone filters, an inner hair cell model (Meddis, 1986), and an auditory nerve model (Carney, 1993). A C++ implementation of the auditory model was provided by Meddis and his colleagues and implemented on a PC as DSAM libraries. Default values were used for all parameters in the model. Twenty auditory nerve fibers were selected to have characteristic frequencies (CF) ranging from 20 to 10000 Hz. The output of each fiber in response to a sinusoid of a standard frequency was used to define a pulse train in electric hearing. The bottom panel of Fig. 2 shows an auditory nerve fiber's (CF=1000 Hz) ISI response to a 1000 Hz sinusoid.

The present pitch discrimination test used standard frequencies of 100, 250, 500, and 1000 Hz. Stimuli were presented to one, two, or four active electrodes, respectively. When one electrode was presented, the selected electrode had a CF that was based on the Greenwood map (Greenwood, 1990) and was presumably equal to the stimulus's standard frequency. When multiple electrodes were presented, the most apical electrode was the same as in the single-electrode case, while the remaining electrodes were selected based on zero-, one-, or two-electrode separation. For example, the electrode pair selected for single electrode condition at the 1000 Hz standard frequency was [12,14]. In the two-electrode condition, the two electrode pairs were [12,14] and [13,15] with zero separation, [12,14] and [14,16] with one-electrode separation, and [12,14] and [15,17] with two-electrode separation.

Each pulse was converted into a biphasic pulse (negative pulse first followed by a short gap and a positive pulse) in electric stimulation with a total duration of 500 ms, a per-phase duration of 50 μs, and a phase separation of 5 μs. Bipolar (BP+1) configuration mode was always used, resulting in a 1.5 mm spacing between two intracochlear electrodes. In multiple-electrode stimulation, the biphasic pulses were interleaved between electrodes with a delay of 10 μs toward apex. The electric stimuli were delivered to the subject and controlled via a customized research interface[1] (Shannon *et al.*, 1990).

C. Procedures

The dynamic range for each selected electrode was measured individually using jittered pulses with $d=0.2$, probabilistic pulses with $p=0.8$, and auditory-model-generated pulses, respectively, at all standard frequencies. The threshold (THR) was the level that the stimuli were just perceptible for subjects. The most comfortable level (MCL), defined as 65%–70% of the maximum loudness level that subjects could tolerate, was employed as the stimulus level. For multiple-electrode stimulation, the dynamic range of each electrode was measured first and the MCL of the stimulation with all electrodes was obtained by proportionally increasing the amplitude of each electrode based on its dynamic range. For example, in the condition of two electrodes, the dynamic ranges for electrode 20 (THR=45 dB) and electrode 18 (THR=50 dB) were 10 and 5 dB, respectively. To measure the MCL of the stimulation with these two electrodes, subjects first received threshold stimulation on the both electrodes, 18 and 20. And then the amplitude of each electrode was increased by a certain percentage (10%) of the dynamic range (2 dB on electrode 20 and 1 dB on electrode 18) until the MCL was reached. A loudness balance procedure was used to balance all stimuli before the test. To further discourage the use of the loudness cue, the amplitude of each interval was roved by a value that was uniformly distributed between −1 and 0 dB (see Chen and Zeng, 2004) so that the levels of all three stimuli in each trial were randomized 1 dB lower than their MCLs. Finally, the stochastic nature of these pulses produced additional uncertainties in terms of the total number of the pulses and the interpulse duration, making the use of the loudness cue highly unlikely. The only exception was for the auditory-model-generated pulses, whose number would decrease when the signal frequency was increased to be greater than the electrode's CF. This was because the neural activities gradually shifted to the next adjacent channel in the model when signal frequency approached the center frequency of the next channel. If this were the case, subjects would recognize the signal by picking the relatively softer sound. However, because of the amplitude roving, the loudness might be a cue only if the frequency of the signal was much higher than the standard frequency and it would not necessarily affect the result of the experiment.

Pitch discrimination was measured using a three-interval, forced-choice, adaptive procedure. In each trial, a subject heard 3 sounds, including two sounds with the standard frequency and a signal with higher frequency. The order of presentation was randomized. The subject was asked to identify the interval with the highest pitch by pressing a button on the computer monitor. Graphical feedback was given after each trial. A two-down, one-up decision rule was employed to track the 70.7% correct point on the psychometric function. To complete each run, the subject had to incur either 13 reversals or 60 trials with at least 8 reversals. The step size was about 25% of the standard frequency for the first 4 reversals and reduced to 3%–5% after that. All subjects completed 3 successful runs with an average standard deviation of about 50% of the mean value.

III. RESULTS

A. Jittered pulses

Figure 3 shows the individual data (panels) in pitch discrimination of jittered pulses in terms of the difference limen (Hz) as a function of the standard deviation at four standard frequencies (100, 250, 500, and 1000 Hz). Each data point is the mean of three trials and the error bar represents the standard deviation of the mean. Had the jitters helped pitch discrimination performance, a negative-sloping curve would be expected in the data. Except for the 250 Hz condition in S3 where the difference limen decreased with the standard deviation, there was no evidence supporting the idea that adding jitters improved pitch discrimination. The jitter degraded the performance at the 100 Hz standard frequency while producing no effects on performance at high frequencies

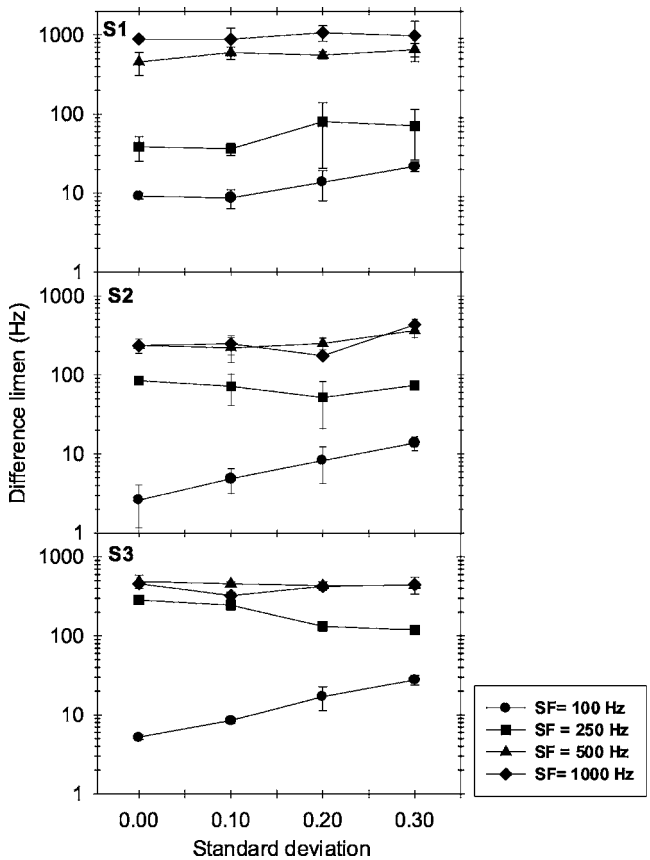

FIG. 3. Individual data (panels) in pitch discrimination of jittered pulses in terms of the difference limen (Hz) as a function of the standard deviation at four standard frequencies (100, 250, 500 and 1000 Hz). Each data point is the mean of three trials and the error bars show the standard deviation of these trials.

(>250 Hz). A within-subjects ANOVA supported these observations by revealing a significant effect of the standard frequency $[F(3,6)=5.9, p<0.05]$, with a significant effect of the standard deviation at only the 100 Hz standard frequency $[F(3,6)=5.9, p<0.05]$.

Figure 4 shows the average pitch discrimination data for jittered pulses as a function of the number of active electrodes at four standard frequencies (panels). Since the electrode spacing produced no significant effect in either the two-electrode condition $[F(2,4)=3.3, p>0.05]$ or the four-electrode condition $[F(1,2)=6.9, p>0.05]$, pitch discrimination data were averaged across the different electrode spacing conditions for presentation and analysis. Different symbols represent pitch discrimination with the standard deviation of 0.0, 0.1, 0.2, and 0.3, respectively. With a 0.0 standard deviation (open circles), the jittered pulses were the same as the traditionally used fixed-rate pulses. The one-electrode data were the same as plotted in Fig. 3. Three important points can be noted in this figure. First, there was no effect of the number of electrodes on pitch discrimination because the overall pattern of the data was essentially flat. Second, pitch discrimination in terms of the difference limen increased with standard frequencies, noting the ascending trend in difference limens from left to right panels. Third, consistent with the single-electrode data, jitters degraded the performance only at the 100 Hz frequency. A three-way (standard frequency $\times$ electrode number $\times$ standard deviation) repeated-measures ANOVA showed a significant effect of the standard frequency $[F(3,6)=167.1, p<0.01]$. Although the standard deviation was not a significant factor in the three-way ANOVA due to the overwhelmingly large variability caused by the standard frequency, a two-way ANOVA showed a significant effect of standard deviation only at the 100 Hz standard frequency $[F(3,6)=23.9, p<0.01]$.

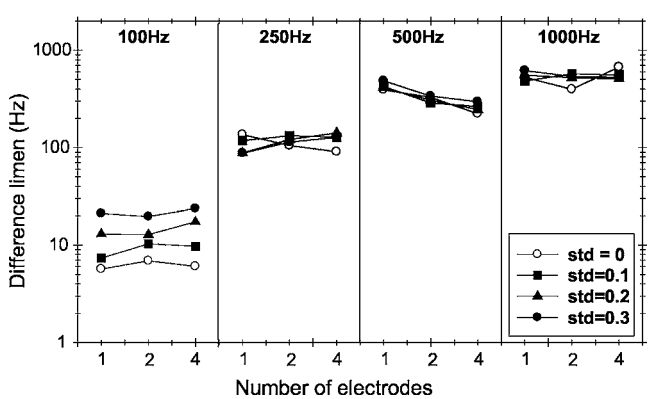

FIG. 4. The average pitch discrimination data for jittered pulses as a function of the number of active electrodes at four standard frequencies (panels). In each panel, different symbols represent pitch discrimination in difference limens with standard deviation of 0, 0.1, 0.2, and 0.3.

B. Probabilistic pulses

Figure 5 shows individual pitch discrimination data for probabilistic pulses as a function of pulse probability. Different symbols represent the data at standard frequencies of 100, 250, 500, and 1000 Hz. No improvement in pitch discrimination was observed in any of the three subjects. The general trend was similar to the data obtained with the fixed-rate stimuli. Difference limen increased with the standard frequency and was nearly independent of the probability except for the 100 Hz standard frequency. A within-subjects ANOVA revealed that the standard frequency produced a significant effect $[F(3,6)=12.0, p<0.01]$, and that the probabilistic pulses significantly increased the difference limen only at the 100 Hz standard frequency $[F(3,6)=12.5, p<0.01]$.

Figure 6 shows the average pitch discrimination data for probabilistic pulses as a function of the number of active electrodes at four standard frequencies (panels). Since the electrode spacing produced no significant effect in either the two-electrode condition $[F(2,4)=4.4, p>0.05]$ or the four-electrode condition $[F(1,2)=0.01, p>0.05]$, pitch discrimination data were averaged across the different electrode spacing conditions for presentation and analysis. Open circles denote the pitch discrimination for $p=1$, at which the probabilistic-pulse train was the same as the fixed-rate pulse train. Similar to the three trends in the jittered pulse experiment, the difference limen for the probabilistic pulses was independent of the number of electrodes $[F(2,4)=2.4, p>0.05]$, increased with the standard frequency $[F(3,6)=19.1, p<0.05]$, and decreased with the probability only at the 100 Hz standard frequency $[F(3,6)=32.8, p<0.01]$.

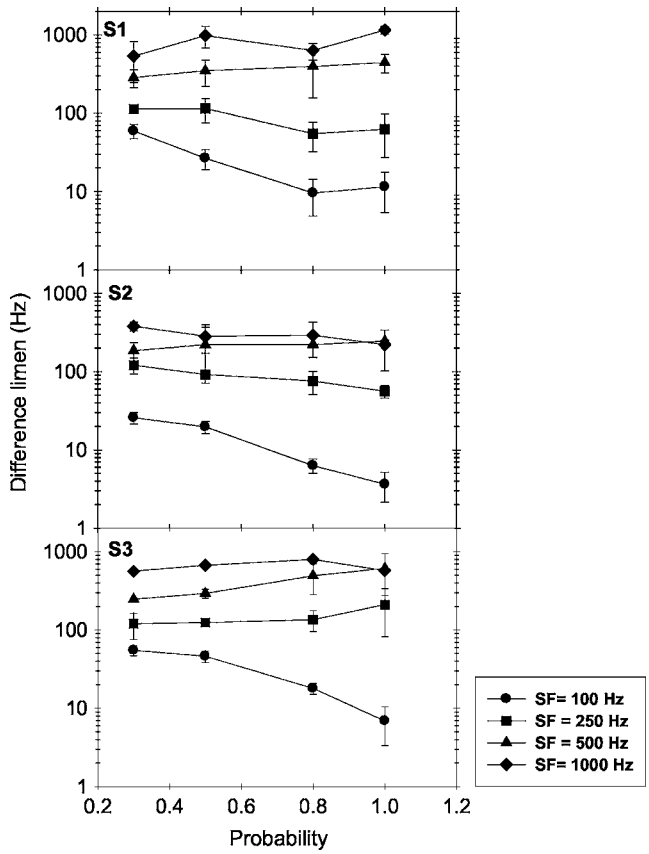

FIG. 5. Individual data (panels) in pitch discrimination of probabilistic pulses in terms of the difference limen (Hz) as a function of the pulse probability at four standard frequencies (100, 250, 500, and 1000 Hz). Each data point is the mean of three trials and the error bars show the standard deviation of these trials.

C. Auditory-model-generated pulses

Figure 7 shows the average pitch discrimination data for the auditory-model-generated pulses (closed symbols connected by the solid line) and the fixed-rate pulses (open circles connected by the dashed line) as a function of the number of active electrodes at four standard frequencies (panels). Since the electrode spacing produced no significant effect in either the two-electrode condition $[F(2,4)=0.001, p>0.05]$ or the four-electrode condition $[F(1,2)=11.3, p>0.05]$, pitch discrimination data were averaged across the different electrode spacing conditions for presentation and analysis. Similar to the trends in the jittered and probabilistic pulse experiments, the difference limen for the auditory-model-generated was independent of the number of electrodes $[F(2,4)=0.38, p>0.05]$ and increased with the standard frequency $[F(3,6)=18.5, p<0.01]$. In addition, the auditory-model-generated pulses produced significantly poorer performance than the fixed-rate pulses across the number of electrodes and standard frequencies $[F(1,2)=32.1, p<0.01]$.

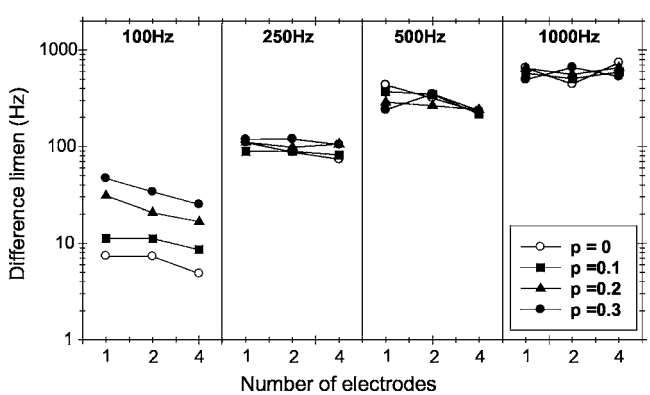

FIG. 6. The average pitch discrimination data for probabilistic pulses as a function of the number of active electrodes at four standard frequencies (panels). In each panel, different symbols represent pitch discrimination in difference limens with probability of 1, 0.8, 0.5, and 0.3.

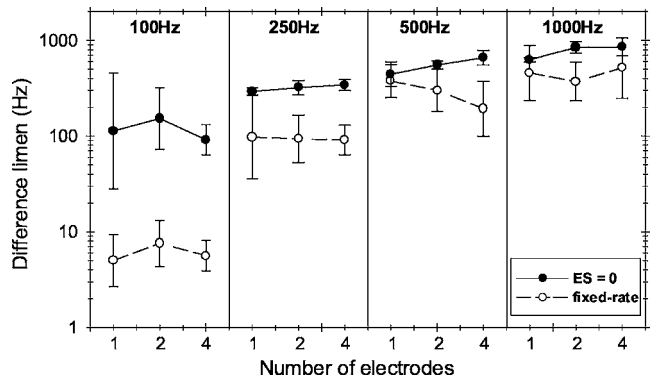

FIG. 7. The average pitch discrimination data auditory-model-generated pulses (closed symbols connected by the solid line) and the fixed-rate pulses (open circles connected by the dashed line) as a function of the number of active electrodes at four standard frequencies (panels). Data were averaged across three subjects and different electrode spacing. The error bars show the standard deviation of the three subjects.

IV. DISCUSSION

A. Temporal patterns

The most important finding in the present study was that none of the three types of stochastic pulses produced better performance in pitch discrimination than the traditional fixed-rate pulses. To the extent that the ISI histogram reflects the normal temporal firing properties, the auditory-model-generated pulses best represent the normal neural temporal response. However, the present result shows worst performance with the auditory-model-generated pulses, especially at low standard frequencies. As mentioned in Sec. II B, if loudness were an effective cue in the present experiment, we would expect better performance with the auditory-model-generated pulses. The present result was not consistent with the potential use of loudness cue in pitch discrimination.

It is difficult to explain why the auditory-model-general pulses produced the worst performance. In acoustic stimulation, not only is the spike activity stochastic in single auditory nerve fibers but also the activities across different fibers are stochastic and independent (Johnson and Kiang, 1976). The brain may need to compare statistically independent temporal firing patterns between fibers to produce optimal performance. These statistically independent across-fiber spike activities are difficult to achieve with current cochlear implant technology.

The randomness introduced by the present methods did not affect performance at high standard frequencies (>250 Hz) but significantly degraded performance at the

100 Hz standard frequency range. The present result indicates that an accurate temporal cue is essential for pitch discrimination at low frequencies in electric hearing with current cochlear implant technology. Disruption of this temporal cue by introducing randomness to either the period in jittered pulses or the first-order interval in probabilistic pulses will degrade pitch discrimination at low frequencies. At high frequencies, the neuronal membrane and central circuitry will likely produce stochastic response, thus adding randomness in the stimulus will less likely affect pitch discrimination at these high frequencies.

B. Autocorrelation models

According to the autocorrelation theory, pitch information extracted from multiple channels should be more salient than that from a single channel if each channel is independent but contains the same information. Furthermore, pitch perception for multiple channels should be robust in the presence of jitters, because these jitters would likely be averaged out in the process of the summarizing between channels. However, the present result showed that neither multiple-electrode stimulation nor electrode spacing produced any significant difference in pitch discrimination compared with single-electrode stimulation. This result was inconsistent with the autocorrelation theory's prediction.

There are three possible reasons that could explain the failure of improvement of pitch in cochlear implant users using auditory models. The first reason is that current cochlear implants do not provide the critical number of independent channels for this autocorrelation model to aggregate and then extract adequate pitch information. The second possible reason is that pitch requires that not only the proper temporal information be present in each channel, but also this temporal information come from the proper place (Miller and Sachs, 1984; Shamma, 1985b; Oxenham *et al.* 2004). While it is relatively difficult to differentiate these models in acoustic hearing because of the tightly coupled temporal-place information, modern cochlear implants provide ample opportunities to explore and test these working hypotheses. The third plausible class of explanation comes from the lack of consistent timing differences between channels, as proposed by Loeb *et al.* (1983) and by Shamma (1985a). For the simplicity of the implementation, the present study used 10 μs fixed delay between channels disregarding the electrode and electrode spacing. The future implementation of the model should take this into account and a variable time delay depending on the location of the stimulation could be used.

C. Practical considerations

Current speech strategies in most cochlear implants deliver electric stimulation by temporally amplitude-modulating a fixed-rate pulse train, disregarding the fine structure information. Pitch information, in these strategies, is coded by either the modulation frequency in the time domain or the location of the stimulation in the "place" domain. Unfortunately, neither delivers a salient pitch percept as evidenced by the poor pitch discrimination via stimulation rate only (Bilger, 1977; Eddington *et al.* 1978; Shannon, 1983a; Carlyon *et al.* 2002; Zeng *et al.* 2002; Chen and Zeng, 2004) or electrode position only, due to electrode interaction (Shannon, 1983b; Fishman *et al.* 1997) and frequency-to-electrode mismatch (Townshend *et al.*, 1987). In a normal auditory system, pitch information is encoded by both place and temporal cues. The failure to improve pitch perception by stochastic pulses suggests that to restore normal pitch perception in cochlear implants, future processing strategies may need to take both place and temporal cues into account.

V. CONCLUSIONS

Three stochastic temporal patterns were used to frequency modulate a fixed-rate pulse train in an attempt to improve pitch perception in cochlear-implant users. The three temporal patterns simulated one or several aspects of the stochastic temporal firing pattern observed in a normal auditory nerve fiber in response to a pure-tone stimulus in acoustic hearing. Perceptual results showed that, compared with the traditional fixed-rate pulse train stimulation, the three stochastic temporal patterns did not improve pitch discrimination in electric hearing but actually degraded performance at low frequencies ($<$250 Hz). Neither multiple-electrode stimulation nor electrode spacing significantly affected pitch discrimination in cochlear implants, suggesting that a strict version of the autocorrelation model for pitch perception needs to be required, and additionally, that the absolute place information may need to be taken into account to restore normal pitch perception in cochlear implant users.

ACKNOWLEDGMENTS

We thank our cochlear-implant subjects for their time and dedication. We also thank Dr. Ray Meddis and Dr. Lowel O'Mard for providing the code of DSAM and technical assistance during the implementation of the auditory-model in the HEINRI system. Thanks also to the associate editor Brenda Lonsbury-Martin for her very helpful comments on the manuscript. This work was partially supported by a NIH Grant No. 2R01DC02267.

[1]J. Wygonski and M. E. Robert, "HEI Nucleus research interface specification," House Ear Institute (2001).

Bilger, R. C. (**1977**). "Psychoacoustic evaluation of present prostheses," Ann. Otol. Rhinol. Laryngol. Suppl. **86**, 92–140.

Carlyon, R. P., van Wieringen, A., Long, C. J., Deeks, J. M., and Wouters, J. (**2002**). "Temporal pitch mechanisms in acoustic and electric hearing," J. Acoust. Soc. Am. **112**, 621–633.

Carney, L. H. (**1993**). "A model for the responses of low-frequency auditory-nerve fibers in cat," J. Acoust. Soc. Am. **93**, 401–417.

Chatterjee, M. and Zwislocki, J. J. (**1997**). "Cochlear mechanisms of frequency and intensity coding. I. The place code for pitch," Hear. Res. **111**, 65–75.

Chen, H. and Zeng, F.-G. (**2004**). "Frequency modulation detection in cochlear implant subjects," J. Acoust. Soc. Am. **116**, 2269–2277.

Dobie, R. A. and Dillier, N. (**1985**). "Some aspects of temporal coding for single-channel electrical stimulation of the cochlea," Hear. Res. **18**, 41–55.

Dynes, S. B. and Delgutte, B. (**1992**). "Phase-locking of auditory-nerve discharges to sinusoidal electric stimulation of the cochlea," Hear. Res. **58**, 79–90.

Eddington, D. K., Dobelle, W. H., Brackmann, D. E., Mladejovsky, M. G., and Parkin, J. L. (**1978**). "Auditory prostheses research with multiple channel intracochlear stimulation in man," Ann. Otol. Rhinol. Laryngol. **87**, 1–39.

Fishman, K. E., Shannon, R. V., and Slattery, W. H. (**1997**). "Speech recognition as a function of the number of electrodes used in the SPEAK cochlear implant speech processor," J. Speech Lang. Hear. Res. **40**, 1201–1215.

Friesen, L. M., Shannon, R. V., Baskent, D., and Wang, X. (**2001**). "Speech recognition in noise as a function of the number of spectral channels: Comparison of acoustic hearing and cochlear implants," J. Acoust. Soc. Am. **110**, 1150–1163.

Garnham, C., O'Driscoll, M., Ramsden R., and Saeed, S. (**2002**). "Speech understanding in noise with a Med-El COMBI 40+ cochlear implant using reduced channel sets," Ear Hear. **23**, 540–552.

Gfeller, K. and Lansing, C. R. (**1991**). "Melodic, rhythmic, and timbral perception of adult cochlear implant users," J. Speech Hear. Res. **34**, 916–920.

Gfeller, K., Woodworth, G., Robin, D. A., Witt, S., and Knutson, J. F. (**1997**). "Perception of rhythmic and sequential pitch patterns by normally hearing adults and adult cochlear implant users," Ear Hear. **18**, 252–260.

Goldstein, J. L. and Srulovicz, P. (**1977**). "Auditory-nerve spike intervals as an adequate basis for aural frequency measurement," in *Psychophysics and Physiology of Hearing*, edited by E. F. Evans and J. P. Wilson (Academic, London), pp. 337–347.

Greenwood, D. D. (**1990**). "A cochlear frequency-position function for several species—29 years later," J. Acoust. Soc. Am. **87**, 2592–2605.

Harris, J. D. (**1952**). "Pitch discrimination," J. Acoust. Soc. Am. **24**, 750–755.

Hartmann, R., Topp, G., and Klinke, R. (**1984**). "Discharge patterns of cat primary auditory fibers with electrical stimulation of the cochlea," Hear. Res. **13**, 47–62.

Helmholtz, H. L. F. (**1863**). "Die lehre von den tonempfindungen als physiologische grundlage fur die theorie der musik," in *On the Sensations of Tone as a Physiological Basis for the Theory of Music* (Dover, New York).

Henning, G. B. (**1967**). "A model for auditory discrimination and detection," J. Acoust. Soc. Am. **42**, 1325–1334.

Hillenbrand, J., Getty, L. A., Clark, M. J., and Wheeler, K. (**1995**). "Acoustic characteristics of American English vowels," J. Acoust. Soc. Am. **97**, 3099–3111.

Javel, E., Tong, Y. C., Shepherd, R. K., and Clark, G. M. (**1987**). "Responses of cat auditory nerve fibers to biphasic electrical current pulses," Ann. Otol. Rhinol. Laryngol. Suppl. **128**, 26–30.

Johnson, D. H. and Kiang, N. Y. (**1976**). "Analysis of discharges recorded simultaneously from pairs of auditory nerve fibers," Biophys. J. **16**, 719–734.

Kaernbach, C. and Demany, L. (**1998**). "Psychophysical evidence against the autocorrelation theory of auditory temporal processing," J. Acoust. Soc. Am. **104**, 2298–306.

Kiang, N. Y. and Moxon, E. C. (**1972**). "Physiological considerations in artificial stimulation of the inner ear," Ann. Otol. Rhinol. Laryngol. **81**, 714–730.

Kong, Y. Y., Cruz, R., Jones, J. A., and Zeng, F. G. (**2004**). "Music perception with temporal cues in acoustic and electric hearing," Ear Hear. **25**, 173–185.

Licklider, J. C. (**1951**). "A duplex theory of pitch perception," Experientia **7**, 128–134.

Litvak, L., Delgutte, B., and Eddington, D. (**2001**). "Auditory nerve fiber responses to electric stimulation: Modulated and unmodulated pulse trains," J. Acoust. Soc. Am. **110**, 368–379.

Loeb, G. E., White, M. W., and Jenkins, W. M. (**1983**). "Biophysical considerations in electrical stimulation of the auditory nervous system," Ann. N.Y. Acad. Sci. **405**, 123–136.

Lopez-Poveda, E. A. and Meddis, R. (**2001**). "A human nonlinear cochlear filterbank," J. Acoust. Soc. Am. **110**, 3107–3118.

Meddis, R. (**1986**). "Simulation of mechanical to neural transduction in the auditory receptor," J. Acoust. Soc. Am. **79**, 702–711.

Meddis, R. and Hewitt, M. (**1991**). "Virtual pitch and phase sensitivity studied using a computer model of the auditory periphery: Pitch identification," J. Acoust. Soc. Am. **89**, 2866–2882.

Miller, M. I. and Sachs, M. B. (**1984**). "Representation of voice pitch in discharge patterns of auditory-nerve fibers," Hear. Res. **14**, 257–279.

Moore, B. C. (**1973**). "Frequency difference limens for short-duration tones," J. Acoust. Soc. Am. **54**, 610–619.

Morse, R. P. and Evans, E. F. (**1999**). "Additive noise can enhance temporal coding in a computational model of analogue cochlear implant stimulation," Hear. Res. **133**, 107–119.

Nelson, D. A., Stanton, M. E., and Freyman, R. L. (**1983**). "A general equation describing frequency discrimination as a function of frequency and sensation level," J. Acoust. Soc. Am. **73**, 2117–2123.

Ohm, G. S. (**1843**). "Uber die definition des tones, nebst daran geknufter theorie der sirene and ahnlicher tonbildender vorichtungen," Ann. Phys. Chem. **59**, 513–565.

Oxenham, A. J., Bernstein, J. G. W., and Penagos, H. (**2004**). "Correct tonotopic representation is necessary for complex pitch perception," Proc. Natl. Acad. Sci. U.S.A. **101**, 1421–1425.

Pijl, S. (**1997**). "Labeling of musical interval size by cochlear implant patients and normally hearing subjects," Ear Hear. **18**, 364–372.

Ritsma, R. J. (**1962**). "Existence region of the tonal residue. I," J. Acoust. Soc. Am. **34**, 1224–1229.

Ritsma, R. J. (**1963**). "Existence region of the tonal residue. II," J. Acoust. Soc. Am. **35**, 1241–1245.

Robert, M. E. (**1997**). "AIPSS-ID—Phoneme identification software," House Ear Institute, Los Angeles.

Rose, J. E., Brugge, J. F., Anderson, D. J., and Hind, J. E. (**1967**). "Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey," J. Neurophysiol. **30**, 769–793.

Rubinstein, J. T., Wilson, B. S., Finley, C. C., and Abbas, P. J. (**1999**). "Pseudospontaneous activity: Stochastic independence of auditory nerve fibers with electrical stimulation," Hear. Res. **127**, 108–118.

Schouten, J. F., Ritsma, R. J., and Cardozo, B. L. (**1962**). "Pitch of the residue," J. Acoust. Soc. Am. **34**, 1418–1424.

Shamma, S. A. (**1985a**). "Speech processing in the auditory system. I. The representation of speech sounds in the responses of the auditory nerve," J. Acoust. Soc. Am. **78**, 1612–1621.

Shamma, S. A. (**1985b**). "Speech processing in the auditory system. II. Lateral inhibition and the central processing of speech evoked activity in the auditory nerve," J. Acoust. Soc. Am. **78**, 1622–1632.

Shannon, R. V. (**1983a**). "Multichannel electrical stimulation of the auditory nerve in man. I. Basic psychophysics," Hear. Res. **11**, 157–189.

Shannon, R. V. (**1983b**). "Multichannel electrical stimulation of the auditory nerve in man. II. Channel interaction," Hear. Res. **12**, 1–16.

Shannon, R. V., Adams, D. D., Ferrel, R. L., Palumbo, R. L., and Grandgenett, M. (**1990**). "A computer interface for psychophysical and speech research with the Nucleus cochlear implant," J. Acoust. Soc. Am. **87**, 905–907.

Siebert, W. M. (**1970**). "Frequency discrimination in the auditory system: Place or periodicity mechanisms," Proc. IEEE **58**, 723–730.

Srulovicz, P. and Goldstein, J. L. (**1983**). "A central spectrum model: A synthesis of auditory-nerve timing and place cues in monaural communication of frequency spectrum," J. Acoust. Soc. Am. **73**, 1266–1276.

Townshend, B., Cotter, N., Van Compernolle, D., and White, R. L. (**1987**). "Pitch perception by cochlear implant subjects," J. Acoust. Soc. Am. **82**, 106–115.

Turner, C. W., Fabry, D. A., Barrett, S., and Horwitz, A. R. (**1992**). "Detection and recognition of stop consonants by normal-hearing and hearing-impaired listeners," J. Speech Hear. Res. **35**, 942–949.

van den Honert, C. and Stypulkowski, P. H. (**1987**). "Temporal response patterns of single auditory nerve fibers elicited by periodic electrical stimuli," Hear. Res. **29**, 207–222.

Wei, C. G., Cao, K., and Zeng, F. G. (**2004**). "Mandarin tone recognition in cochlear-implant subjects," Hear. Res. **197**, 87–95.

Wever, E. G. (1948). "Theory of Hearing," in (Wiley, New York).

Wier, C. C., Jesteadt, W., and Green, D. M. (**1977**). "Frequency discrimination as a function of frequency and sensation level," J. Acoust. Soc. Am. **61**, 178–184.

Zeng, F.-G. (**2002**). "Temporal pitch in electric hearing," Hear. Res. **174**, 101–106.

Zeng, F.-G., Fu, Q. J., and Morse, R. (**2000**). "Human hearing enhanced by noise," Brain Res. **869**, 251–255.

Zeng, F.-G., Grant, G., Niparko, J., Galvin, J. J., Shannon, R. V., Opie, J., and Segel, P. (**2002**). "Speech dynamic range and its effects on cochlear implant performance," J. Acoust. Soc. Am. **111**, 377–386.

Zwicker, E. (**1956**). "Die elementaren grundlagen zur bestimmung der informationskapazitat des gehors," Acustica **6**, 356–381.