

# Loudness perception with pulsatile electrical stimulation: The effect of interpulse intervals<sup>a)</sup>

Colette M. McKay and Hugh J. McDermott

*The University of Melbourne, Department of Otolaryngology, Parkville 3052, Australia*

(Received 23 June 1997; accepted for publication 16 April 1998)

The effect of interpulse intervals on the perception of loudness of biphasic current pulse trains was investigated in eight adult cochlear implantees at three different stimulus levels encompassing the psychophysical dynamic range. Equal-loudness contours and thresholds were obtained for stimuli in which two biphasic pulses were presented in a fixed repetition period (4 and 20 ms), and also for single-pulse/period stimuli with rates varying between 20 and 750 Hz. All stimuli were of 500-ms duration, and the phase durations of each pulse were 100  $\mu$ s or less. The results of these experiments were consistent with predictions of a three-stage model of loudness perception, consisting of a peripheral refractory effect function, a sliding central integration time window, and a central equal-loudness decision device. Application of the model to the data allowed the estimation of neural refractory characteristics of the subjects' remaining peripheral neural population. The average neural spike probability for a 50-Hz stimulus was predicted to be about 0.77, with an associated neural refractory time of 7.3 ms. These predictions did not vary systematically with level, implying that the effect of increasing current level on loudness results more from recruitment of neurons than from any increase in average spike probability. © 1998 Acoustical Society of America. [S0001-4966(98)00708-5]

PACS numbers: 43.66.Mk, 43.66.Cb, 43.66.Ts [JWH]

## INTRODUCTION

For electrical stimulation of the acoustic nerve, as for acoustic stimulation, the main parameter which controls the resultant loudness percept is the stimulus magnitude. This is controlled in biphasic pulsatile stimulation by the current amplitude or phase duration of the pulses, with loudness being a monotonic function of both these parameters. In this paper, we investigate the more subtle effects on loudness of temporal parameters of the signal, in particular the effects of interpulse intervals (the time between onsets of successive biphasic pulses). The effect of varying interpulse interval is difficult to study as an independent factor, because there are associated covariables such as rate, number of stimulus pulses, or stimulus duration, which may also affect loudness. For pulse trains that utilize a single interpulse interval and fixed stimulus duration, the rate of stimulation and number of stimulus pulses covary. In stimuli with a fixed number of stimulus pulses, the stimulus duration covaries with interpulse interval.

Pfingst *et al.* (1996) have studied the effect of pulse separation (the time between the end of the first waveform and the onset of the next waveform) on thresholds of detection in human implantees, when pulse separation was varied in stimuli of fixed 500-ms duration (with rate and total number of stimulus pulses covarying). They showed that the

threshold versus rate function monotonically decreased for shorter pulse durations (500  $\mu$ s per phase), but was bowl shaped for longer pulse durations (2 ms). This result was consistent with previous data obtained with monkeys or humans (Shannon, 1985, 1989; Pfingst and Morris, 1993; Moon *et al.*, 1993), except for a species difference affecting the pulse separation at which the 2-ms pulse duration curve reached the minimum threshold. Pfingst *et al.* also measured thresholds for stimuli with a fixed number of pulses (2 or 10) with stimulus duration covarying. In this case, the threshold versus pulse separation functions had the same shape as for the fixed-duration stimuli, for each pulse duration used (500  $\mu$ s and 2 ms). The main difference between the two-pulse, ten-pulse, and 500-ms stimuli was that stimuli with more pulses had lower thresholds, and produced threshold versus pulse separation curves with steeper slopes. The authors concluded that the pulse separation (or interpulse interval) contributed significantly to the shape of the threshold versus rate curves (as opposed to rate *per se*, or number of stimulus pulses). They stated that their results were consistent with the presence of an integrating mechanism which was most effective within about 5 ms, and that, in the case of very long pulse durations, the non-monotonic threshold versus rate functions were consistent with a neural inhibitory mechanism which operated for short time intervals after long-duration pulses. They suggested that this inhibition may be due to an alteration in the ion channel kinetics of auditory neurons, caused by inactivation of sodium channels for several milliseconds following an action potential, or following a long-duration subthreshold pulse.

The above studies indicate that interpulse interval has a significant effect on detection thresholds. There are several mechanisms which may be involved in temporal interactions

<sup>a)</sup>Some of the data in this paper were presented at the 131st meeting of the Acoustical Society of America, Indianapolis, May 1996 [C. M. McKay and H. J. McDermott (1996) "The effect on pitch and loudness of major interpulse intervals within modulated current pulse trains," *J. Acoust. Soc. Am.* **99**, 2584(A) (1996)], and at the British Society of Audiology short papers meeting, Nottingham, September 1997 [C. M. McKay and H. J. McDermott (1997) "A loudness model for electrical stimulation: implications for physiological differences among cochlear implantees"].

between pulses, both at the level of individual neurons, and at the level of central processing of multiple neural responses. At the level of individual neurons, an inhibition of responsiveness (refractoriness) will follow an action potential. For electrically stimulated auditory neurons, an action potential is followed by an absolute refractory time of about 1 ms, and a relative refractory time of up to 6–10 ms, in which the probability of eliciting a second action potential is zero or reduced below its normal value (Stypulkowski and van den Honert, 1984; Parkins, 1989). On the other hand, an excitatory effect on an individual neuron may occur if residual charge remains on the neural membrane following a subthreshold pulse. Such an excitatory effect has been demonstrated by Butikofer and Lawrence (1979) in a simulation using the Frankenhauser–Huxley model of neural stimulation. Since residual charge dissipates quickly, however, these excitatory effects are unlikely to be significant for pulse separations more than about 400  $\mu$ s.

Mechanisms at a more central level, involving multiple neural responses, are also sensitive to interpulse interval duration. Temporal integration is an example of such a mechanism, and is one of four stages in models which have been commonly applied to account for temporal resolution ability in the auditory system (Viemeister, 1979; Buus and Florentine, 1985; Green and Forrest, 1988; Moore *et al.*, 1988; Plack and Moore, 1990, 1991; Oxenham and Moore, 1994). In these models, the first two stages (auditory filter and compressive nonlinearity) represent the processing which occurs in the periphery and cochlea. The third and fourth stages (sliding temporal integrator with an effective window duration of 3–10 ms and decision device) model more central mechanisms, and are thus likely to be applicable without modification to electrical stimulation. Such a four-stage acoustic model was used in a series of experiments measuring detection of increments and decrements of various durations in sinusoidal stimuli (Moore *et al.*, 1993, 1996; Peters *et al.*, 1995). Those authors found that the model parameters which provided the best fit across frequency and level conditions were an integration time window with equivalent rectangular duration (ERD) of about 7 ms, and a decision device which used a criterion of constant integrator output change (on a dB scale) across conditions to detect changes in intensity (Moore *et al.*, 1996).

Viemeister and Wakefield (1991) proposed an additional feature of central processing to explain the longer term temporal effects such as the decrease in threshold that occurs with increasing signal duration up to about 300 ms. In their model of auditory detection, multiple “looks” are stored in memory and can be used selectively to facilitate decisions about detection and discrimination. Each “look” involves short-term integration, over about 3 ms, of the output of the auditory filters. The time constant of the “look” limits temporal resolution (such as gap detection), but improved information can be obtained by combining “looks” over much longer time scales (up to about 300 ms) for tasks such as threshold detection.

The purpose of the experiments reported below was to study the effects on loudness perception of interpulse intervals in current pulse trains at threshold and suprathreshold

levels. A model of loudness perception will be used to predict the data and infer predicted physiological differences among implantees based on the data. In the first two experiments, equal-loudness relationships were measured for pulse trains containing pairs of biphasic pulses presented at a fixed repetition rate, but with a variable interpulse interval within each pair. In contrast with previous electrical stimulation experiments described above, interpulse interval was the only variable in these stimuli, but there were necessarily two different interpulse intervals (within the pulse pairs, and between the second pulse and the first pulse of the next pair) which covaried over different ranges. With this stimulus pattern, any overall effect on loudness could then be attributed to differences in the effects of interpulse interval over the two covarying interval ranges. In the third experiment, equal-loudness contours were measured for rates of stimulation between 20 and 750 Hz, in order to compare the effect on loudness of interpulse interval alone to the effect for stimuli of differing overall pulse rate.

In this paper, the stimuli used in the three experiments are described first, and a model of loudness perception is proposed which enables the prediction of the effect on loudness of interpulse intervals in electrical stimuli for the three experimental stimulus paradigms. This model incorporates the refractory effects in electrically stimulated peripheral auditory neurons (stage 1), as well as central integration (stage 2), and decision mechanisms (stage 3), the last two stages being similar to those previously postulated from acoustic experiments. The experimental procedures and results are then presented and compared to the model predictions. The model-fitting procedure results in an inferred description of the refractory characteristics of the subjects’ peripheral neurons. In addition, information about the spatial and temporal response patterns of each individual implantee’s population of residual auditory neurons is deduced by studying the individual differences in the data.

## I. STIMULI AND MODEL

### A. Stimuli

The stimuli in these experiments were biphasic current pulse trains, with a total duration of 500 ms, delivered in a bipolar or common-ground stimulation mode. The current amplitude was equal for all pulses in each stimulus. Short duration pulses (50–100  $\mu$ s) were used throughout these experiments, partly because speech processors used with cochlear implants require these short durations to produce rates that can adequately code temporal information extracted from the speech signal. As pulse duration has been shown to affect the refractory and excitatory state of auditory neurons (see above), it is important to note that this factor was held constant in these experiments, and the conclusions drawn later about neural response characteristics will apply only to short-pulse-duration stimuli, such as those used in speech processors.

The test stimuli for experiments 1 and 2 are illustrated in Fig. 1. They consisted of two biphasic pulses which were repeated with a period of 20 ms (experiment 1) or 4 ms (experiment 2). The test interpulse interval was defined as

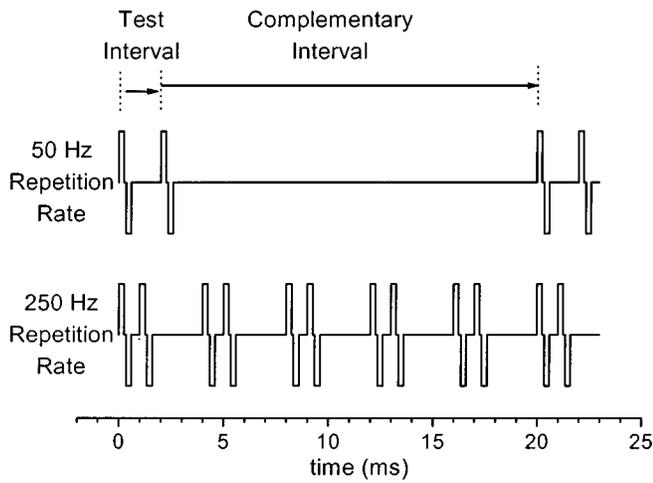


FIG. 1. The test stimuli used in experiments 1 and 2 (23-ms segment). The test interpulse interval was varied between the smallest possible (less than 1 ms) and a maximum of 10 ms for the 50-Hz repetition rate, or 2 ms for the 250-Hz repetition rate. The interphase gap was approximately 43  $\mu$ s.

the time between the start of the leading phases of the two biphasic pulses, and the complementary (covarying) interpulse interval was the interval between the start of the second pulse and the beginning of the first pulse in the next period. The pulse separation [as defined by Pfingst *et al.* (1996) and described above] can be derived from the interpulse interval by subtracting two phase durations (listed in Table I) and the interphase interval of 43  $\mu$ s.

The minimum available pulse separation was limited by the fixed data transmission time of the implant (approximately 445  $\mu$ s). The test interpulse intervals were varied from the minimum (less than 1.0 ms) to a maximum of half the repetition period (10 or 2 ms for experiments 1 and 2, respectively). In addition, pulse trains with a single pulse per period (i.e., a 50-Hz or 250-Hz pulse train) were included. In experiment 3, one-pulse/period pulse trains were also used, with rates between 20 and 750 Hz.

## B. The temporal interaction model of loudness

The first stage of our model describes the effects of refractoriness on peripheral neural excitation elicited by the electrical current pulses within a stimulus. This first stage replaces the first two stages of the acoustic models outlined in the Introduction (auditory filter and compressive nonlinearity). The second stage of our model is a sliding temporal integration window, and the third is a central decision device for comparing loudness percepts.

### 1. Stage 1: Peripheral neural excitation

For the two-pulse stimuli with 20-ms repetition period (experiment 1), we have assumed that the excitation produced by the first of the two biphasic pulses ( $E1$ ) is unaffected by refraction, or other temporal effects, since the no-stimulation period preceding this pulse is always longer than 10 ms. (It follows that the excitation produced by each pulse in the 50-Hz pulse train is also  $E1$ .) The excitation produced by the second biphasic pulse ( $E2$ ) was assumed to depend on the test interval  $t$ , in the following way:

$$E2(t) = E1 \left( 1 - \frac{R}{1 + e^{(t-T)/0.8}} \right). \quad (1)$$

Here  $E2(t)$  is a sigmoid function which was chosen based on the simple assumptions that the activated neurons will have a range of thresholds, leading to a range of spike probabilities and refractory times (since these depend on the signal amplitude relative to the individual thresholds), and that these values are normally distributed.  $E2(t)$  has a value close to  $E1(1 - R)$  for small values of  $t$ , rising to a value of  $E1$  for large values of  $t$ . The parameter  $R$  can range from near zero to 1, and can be thought of as the average spike probability for the first pulse for all neurons with thresholds below the stimulus current (and for the second pulse for all such neurons that did not fire on the first pulse). That is, an  $R$  near zero would represent a low proportion of available neurons firing on the first pulse, resulting in little decrease of excitation for the second pulse no matter how small the test

TABLE I. Details of subjects who took part in the study. The last two columns list the stimulus parameters used in each experiment. The numbers that describe the electrode specify the two rings comprising the bipolar pair used, or in the case of subjects 6 and 7, the single active ring in common ground mode. The rings of an electrode array are numbered 1 to 22 in the basal-to-apical direction. The sixth column lists the electrical dynamic range of the electrode used, measured using 100-Hz pulse trains with pulse durations as listed for experiment 1.

Subject	Age (yr)	Etiology	Implantation date	Length of profound deafness	Electrical dynamic range for 100 Hz (dB current)	Experiment no.	Pulse duration ( $\mu$ s)	Electrode
1	65	Progressive/genetic	May 1986	30 yr	2.2	1, 3	50	(17,19)
						2	50	(17,19)
2	54	Otosclerosis	October 1990	5 yr	4.3	1	100	(15,19)
						2	70	(16,19)
3	55	Meningitis	September 1990	39 yr	3.0	1, 3	100	(18,20)
						2	50	(18,20)
4	35	Sudden onset/unknown	July 1989	5 months	2.2	1	100	(16,18)
						2	50	(16,18)
5	59	Otosclerosis	February 1992	14 yr	2.7	1, 3	100	(10,12)
6	51	Trauma	November 1988	17 yr	3.1	1, 3	70	20
7	48	Progressive/genetic	January 1987	1 yr	5.5	1	100	15
8	41	Otosclerosis	November 1992	6 yr	3.3	1	70	(18,20)

interval. In contrast, a value of  $R$  near 1 would represent a situation where most available neurons fire on the first pulse and are thus in a relative or absolute refractory state for the second pulse, resulting in lower excitation for small test intervals.

The value of  $T$  in Eq. (1) specifies the midpoint (in time) of the sigmoid function. It is the average time for which neurons activated on the first pulse remain refractory. Parkins (1989) measured the interspike intervals for 2500-Hz pulse trains in single auditory neurons in squirrel monkeys. He showed that the average time a neuron remains refractory is inversely related to the amount by which the stimulus current exceeds its physiological threshold. The mean interspike times were more than 4 ms for currents within about 1.6 dB of threshold, and reached an asymptote of about 1 ms for currents more than 7 dB above threshold. The standard deviation of refractory times in the population of neurons which responded to the first pulse of our stimulus (which determines the slope of the sigmoid function) was set to 0.8 ms [see Eq. (1)]. Initial analysis of the data showed that goodness-of-fit of the model was insensitive to this parameter within the range 0.5–1.1 ms.

The modeling of relative excitation from the two pulses within the 4-ms repetition period (experiment 2), and for each pulse for the different rates in experiment 3, is more complicated than described above, since, for overall rates exceeding about 100 Hz, every pulse in a stimulus (apart from the first) will be subject to refractory effects. The total effect for any particular pulse will depend on its position within the whole stimulus pulse train. For example, a high-rate pulse train would be expected to produce a large response on the first pulse, followed by a decaying oscillation in response amplitude until a steady-state response occurs [see Javel (1990), Fig. 17.22 for an example]. We will assume for the purpose of our model that the implantee's loudness comparison is based on the steady-state portion of the response. We will also assume that, in the steady part of the response, each stimulus pulse produces the same total excitation which is dependent on the pulse rate (experiment 3), or average pulse rate (experiment 2). Since the range of test interpulse intervals in experiment 2 is small, and the excitation produced by any one pulse will be affected by up to five preceding pulses (occurring within the preceding refractory-time range of about 10 ms), it is expected that the effect of changes to the test interpulse interval on the relative excitation arising from the two pulses will be small, compared to the effect of changing the average interpulse interval (or overall rate).

We will denote the average spike probability for those neurons contributing to the steady-state response to a pulse train of rate  $r$  as  $\theta_r$ . We will assume that the total number of neurons above threshold is dependent only on the current magnitude (and not on the rate). The dependence of total excitation per pulse ( $E_r$ ) on rate at any particular current can then be expressed as

$$E_r = (\theta_r / \theta_{\text{REF}}) E_{\text{REF}}, \quad (2)$$

where  $\theta_{\text{REF}}$  is the average spike probability and  $E_{\text{REF}}$  is the total excitation produced by each pulse in the reference

stimulus. The ratio of the two spike probabilities was used as a fitting parameter in our model for experiments 2 and 3.

## 2. Stage 2: Temporal integration window

The relative excitation per pulse given by the refractory model of stage 1 was used as the input to a temporal integrator (stage 2). We have used a sliding, asymmetric, exponential time window  $W(t)$ , with time constants on each side ( $\tau_1$  and  $\tau_2$ ) such that  $\tau_1$  is 0.63 times  $\tau_2$  [as determined by Oxenham and Moore (1994)], i.e.,

$$\begin{aligned} W(t) &= e^{t/\tau_1}, & t < 0, \\ W(t) &= e^{-t/\tau_2}, & t > 0, \end{aligned} \quad (3)$$

where  $t$  is time (in ms).

The equivalent rectangular duration (ERD), defined as  $\tau_1 + \tau_2$ , was set to 7 ms, based on the parameters found for discrimination of intensity changes in sinusoids by Moore *et al.* (1996). Since this integration most likely occurs more centrally than the acoustic nerve, implantees were assumed to have the same parameters as normally hearing subjects.<sup>1</sup> This assumption may be invalid if the pathology causing deafness in the implantees affected that part of the central auditory pathways which is responsible for the temporal integration. However, our assumption is supported by measurements of temporal resolution in implantees, who have exhibited generally similar performance to that of normally hearing subjects (Shannon, 1993). The suitability of this ERD value for our subjects was supported by initial data fitting procedures in which the ERD was a variable parameter (see Sec. III).

## 3. Stage 3: Equal-loudness decision criterion

The third stage of our model is analogous to the fourth stage of Moore *et al.*'s (1996) model of intensity discrimination, in which a fixed ratio change of output from the integrator was used as a criterion for detection of intensity change. In our model, a decision of equal loudness is based upon a criterion of equal output from the integrator of stage 2 (or a difference less than that used for detection of intensity difference). Since the stimuli used in our experiment had a fixed duration, we assumed that the longer-term temporal integration (which causes the threshold of stimuli to decrease with increasing stimulus duration up to about 300 ms) would not affect our data. That is, we have assumed that if the output of the short-term integration time window is equal for two stimuli, then the total loudness will be equal after applying the processing of the longer integration time window. The validity of this assumption, as assessed by the results of this study, will be discussed below.

The first two stages of our model predict how the output of the integrator changes when interpulse interval changes, given an invariant stimulus current. For experiments 1 and 2, the ratio change of integrator output (expressed in dB) was calculated for the two-pulse/period stimuli versus the one-pulse/period stimulus, as a function of test interpulse interval. Some examples of these model predictions are shown in

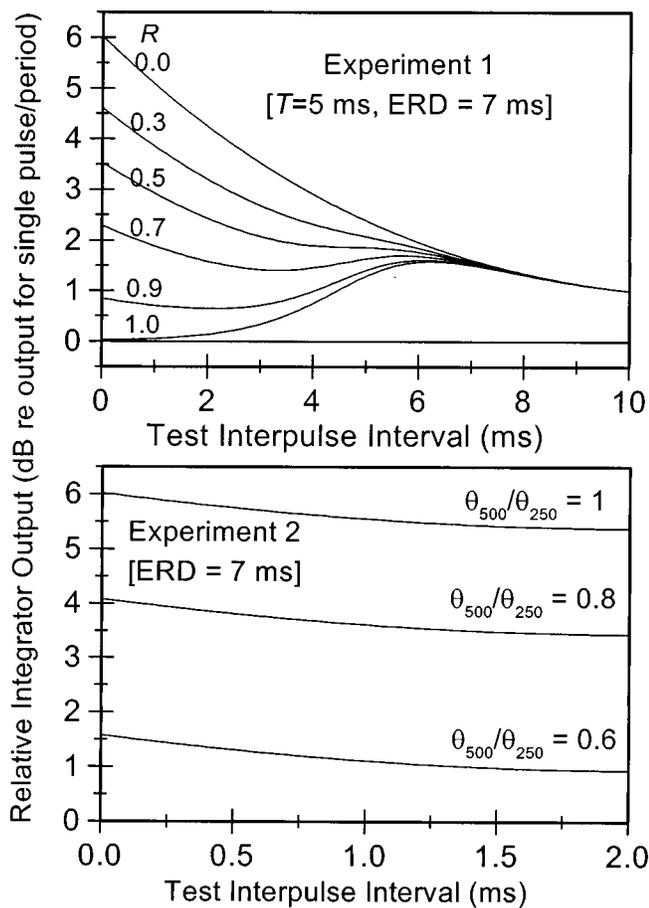


FIG. 2. Some examples of the predicted effect of test interpulse interval on the temporal integration window output for the two-pulse/period stimuli (dB re: the output for single-pulse/period stimuli) for 20-ms repetition period (top panel), and 4-ms period (bottom panel). The separate lines show the effect of changing  $R$  (top panel) and the ratio  $\theta_{500}/\theta_{250}$  (bottom panel), while the other parameters are kept constant as shown.

Fig. 2, where the predicted effects of changing the parameter  $R$  in experiment 1, and the ratio  $\theta_{500}/\theta_{250}$  in experiment 2, are illustrated.

## II. METHODS

### A. Subjects and hardware

Eight adult subjects participated in this study, all users of the Mini System 22 implant, manufactured by Cochlear Ltd. All are postlinguistically profoundly deaf, and had no useful hearing before implantation. Details concerning their etiology, length of profound deafness, and implant experience are given in Table I.

The Mini System 22 cochlear implant consists of an intracochlear array of 22 active electrodes, spaced at intervals of 0.75 mm, which are activated by an implanted receiver-stimulator. The receiver-stimulator receives digitally encoded signals from a speech processor via a transcutaneous inductive link. In these experiments the speech processor was interfaced with an IBM-compatible personal computer which, with specifically designed software, was used to control the parameters of stimuli in each experiment and to record the subject's responses. Stimulation current is

controlled in this implant in discrete steps, and the current produced at each step was determined from the calibration data obtained from the manufacturer for each subject's implant.

The electrodes used, and their dynamic range for 100-Hz pulse trains, are listed in Table I, along with the pulse durations used. The spatial extent (distance between the component rings of the bipolar pair) was chosen for each subject in each experiment based upon the minimum needed to achieve comfortable loudness without excessive currents. Two subjects used common ground mode.<sup>2</sup>

### B. Experimental procedures

#### 1. Experiment 1: 20-ms repetition period

The effect on loudness of varying the interpulse intervals between the pairs of pulses in stimuli with a 50-Hz repetition rate (see Fig. 1) was measured by balancing the loudness of each test stimulus with that of a fixed reference stimulus. The reference stimulus was the two-pulse/period stimulus with the smallest interpulse interval for each subject. This reference was chosen to minimize the perceptual differences (apart from loudness) between the pairs of stimuli being loudness balanced, as any other differences (such as pitch or timbre) could introduce a loudness bias. A 50-Hz pulse train was also balanced with the same reference stimulus. The loudness-balance procedure was conducted at two current levels: first with the reference stimulus at a comfortably loud level, and second with the reference set to half the number of current steps above threshold current compared to the first reference level. Third, the threshold currents for all stimuli were measured using a modified Hughson-Westlake adaptive procedure (Carhart and Jerger, 1959).

The loudness balance procedure was a two-interval forced-choice adaptive procedure, whereby the reference and test stimuli were presented in random order, and the subject was asked whether the first or second stimulus was louder. The current of the test stimulus was initially set so as to produce a percept clearly louder than that of the reference (usually several current steps). The current was then adjusted up or down by one current step (approx. 0.12 dB) whenever the test stimulus was considered softer or louder, respectively, in at least two out of three consecutive trials (see entry 8 in Table I in Levitt, 1971). The procedure continued until 11 turning points were obtained and the last 8 of these were averaged. To overcome potential bias due to the initial loudness or to which stimulus was being adjusted, the procedure was repeated with the two stimuli interchanged (i.e., adjusting the current of the reference stimulus to match the loudness of the test stimulus). In this case, the test stimulus current was set at the previously found equal-loudness current, and the current of the reference stimulus was initially adjusted to produce a loudness greater than that of the test stimulus. Thus two estimates were obtained of the test-stimulus current required for loudness equal to that of the reference stimulus, with the test stimulus starting from above and below the loudness of the reference.

For purposes of comparison with the model predictions, the current reductions (in dB) of the two-pulse/period stimuli

compared to the current of the equally loud one-pulse/period stimulus were calculated by subtracting the average current value for the equally loud one-pulse/period stimulus from each of the current values for the two-pulse/period stimuli.

## 2. Experiment 2: 4-ms repetition period

The adaptive loudness balancing procedure described above was used to equalize the loudness for pairs of biphasic pulses with 4-ms repetition period, and test interpulse intervals of up to 2 ms. In this experiment, the reference stimulus was selected to be the 250-Hz pulse train. This choice of reference stimulus allowed the reduction in current for the two-pulse/period stimuli compared to the equally loud one-pulse/period stimulus to be obtained directly, and was made possible by the smaller pitch or timbre differences among the stimuli. Subjects 1–4 participated in this experiment, and the equal-loudness relationship was measured at the comfortably loud level only.

## 3. Experiment 3: Equal loudness for different rates

The equal-loudness relationship for different rates of stimulation between 20 and 750 Hz were obtained at a comfortably loud level for subjects 1, 3, 5, and 6. Each stimulus was loudness balanced with the 50-Hz stimulus. In this case, the loudness balance was obtained using a method of adjustment. The test and reference (50 Hz) stimuli were presented alternating continuously, separated by 500-ms silent intervals, and the subject altered the current of the test stimulus with an unmarked knob, until the loudness was judged equal to that of the reference. An average result of four balances was calculated (two with the reference fixed, and two with the reference adjusted, with the adjustable stimulus always started at a level perceptibly softer than the fixed). This loudness-balance procedure was chosen because it was more time efficient, and the balanced current differences were expected to be larger than in the previous two experiments, thus not needing the greater measurement precision of the adaptive method. Thresholds for all stimuli were obtained as in experiment 1. The reduction of current for the test stimulus relative to that of the equally loud 50-Hz stimulus (in dB) was then calculated.

## C. Comparison of experimental data with model predictions

### 1. Current-to-excitation transformation

In order to quantitatively relate the experimental results to the model output, it is necessary to adopt a transformation between the change in stimulus current and the change in neural excitation. However, the precise nature of the relationship between current and neural excitation is unknown, and it may also be subject, electrode, or level dependent. As a first approximation, we have assumed that the excitation produced by a current pulse can be described by a simple power function of current, with exponent  $S$ , at least over the small current range used to balance the loudness in each experiment. In other words, the change in integrator output (in dB) resulting from a current change  $\Delta i$  (in dB) was assumed to be  $S \Delta i$ . This scaling factor was assumed to be constant over

small changes of current, and to be stimulus independent, but was allowed to be subject and level dependent.

The model predictions, when divided by  $S$ , therefore predict the current increase which would have elicited the same increase in excitation as the change in stimulus type (from one- to two-pulse/period, or 50 Hz to a different rate) while keeping current constant. This is equal to the *reduction* in current (relative to the reference stimulus current) required to reduce the excitation of the test stimulus to a value equal to that of the reference stimulus (as measured in the experiments).

## 2. Model-fitting procedure

For experiment 1, the prediction for the current increase which would produce the same increase in excitation as the change from one- to two-pulse/period stimuli was fitted to the experimentally measured reduction in current for the two-pulse/period stimuli (relative to the reference stimulus current), to make it equal in loudness to the one-pulse/period stimulus (see above).

The fitting procedure was an iterative  $\chi^2$  minimization procedure based on the Levenberg–Marquardt algorithm (Marquardt, 1963). Three fitting parameters were used:  $R$ ,  $T$ , and  $S$ . As mentioned above, the integration time window was assumed to have an ERD of 7 ms. Initially, however, fitting was attempted with ERD as a free fourth parameter (but constrained to be equal across the three stimulus levels for each subject) to investigate whether the data were consistent with this 7-ms value. In five out of the eight subjects, the dependence of ERD on the value of at least one other parameter was too great to obtain a valid estimate of ERD. The results for the other three subjects gave estimates for ERD which were generally consistent with the 7-ms value. The ERDs (in ms) and their standard errors were: for  $S_4$ , ERD=7.5(1.3); for  $S_6$ , ERD=5.3(0.7); for  $S_7$ , ERD=6.8(0.5). Based on these results, and their consistency with our initial assumptions about the integration time window, the ERD was fixed at 7 ms for all subjects and stimuli.

In experiment 2, the current-to-excitation exponent,  $S$ , which was obtained in the fitting procedure for experiment 1 for each subject at comfortable loudness, was reused. Thus there was a single fitting parameter,  $\theta_{500}/\theta_{250}$ , used to fit the model predictions to the data.

In experiment 3,  $S$  was also constrained to be equal to the values for comfortably loud and threshold stimuli derived by the fitting procedure for experiment 1. The fitting parameter in this case was the ratio  $\theta_r/\theta_{50}$ .

## III. RESULTS

### A. Experiment 1: 20-ms repetition period

The results of experiment 1 for the eight subjects are shown in Fig. 3, where the reductions in current for the two-pulse/period stimuli (in dB relative to that of the equally loud one-pulse/period stimulus) are plotted against the test interpulse interval for comfortably loud, mid-intensity, and threshold stimuli. It can be seen from the figure that the repeated loudness balancing results at comfortably loud and

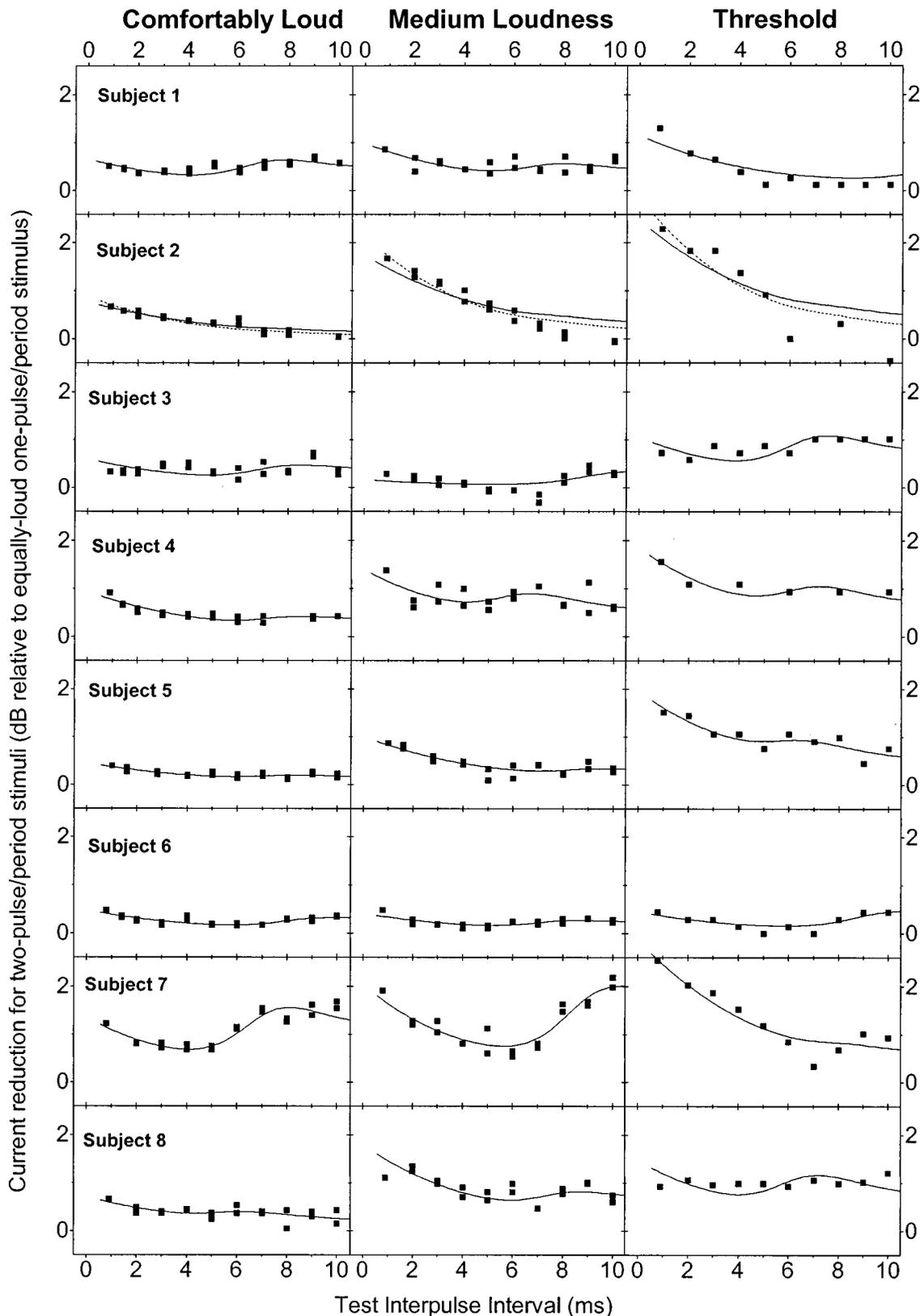


FIG. 3. The threshold and loudness balancing results for different test interpulse intervals in experiment 1, with the best-fit model predictions. Each row contains the data for one subject, and each column the data for each stimulus level. The vertical axes show the current reduction required to make each two-pulse/period stimulus equal in loudness to the one-pulse/period stimulus. The horizontal axes represent the test interpulse interval (see Fig. 1). The pair of symbols at each interpulse interval in the comfortably loud and medium loudness panels show the two loudness balance results (see text). The solid lines represent the best-fit model predictions, using an ERD of 7 ms and the parameter values shown in Table II. The dotted lines for subject 2 represent predictions with a smaller ERD (5.5 ms).

TABLE II. The values of the fitted parameters  $S$ ,  $R$ , and  $T$  which gave the best predictions of the data, for each of the eight subjects at the three levels tested. All fitted functions were with a time-window ERD of 7 ms, with the exception of the second set of values for subject 2, where 5.5 ms was used. Asterisks denote the values which were fixed to avoid high dependency among parameters (see text).

Subject	Comfortably loud level fitted parameters (standard errors)	Mid-dynamic range level fitted parameters (standard errors)	Threshold fitted parameters (standard errors)
1	$S=1.85$ (0.17) $R=0.85$ (0.2) $T=6.7$ (0.3) ms	$S=2.05$ (0.3) $R=0.74$ (0.4) $T=7.0$ (0.5) ms	$S=1.41$ (0.18) $*R=0.795$ $T=12$ (5) ms
2	$S=6.1$ (0.9) $*R=0.3$ $*T=7.3$ ms	$S=2.6$ (0.2) $*R=0.3$ $*T=7.3$ ms	$S=1.9$ (0.15) $*R=0.3$ $*T=7.3$ ms
2 ERD=5.5 ms	$S=5.3$ (0.7) $*R=0.3$ $*T=7.3$ ms	$S=2.4$ (0.2) $*R=0.3$ $*T=7.3$ ms	$S=1.7$ (0.1) $*R=0.3$ $*T=7.3$ ms
3	$S=2.3$ (0.4) $R=0.83$ (0.04) $T=7.3$ (0.7) ms	$S=2.6$ (1.0) $R=0.95$ (0.05) $T=9.2$ (1.7) ms	$S=1.15$ (0.12) $R=0.85$ (0.2) $T=8.4$ (0.4) ms
4	$S=2.5$ (0.6) $R=0.69$ (0.08) $T=7.6$ (1.0) ms	$S=1.59$ (0.12) $R=0.71$ (0.03) $T=5.9$ (0.4) ms	$S=1.23$ (0.17) $R=0.70$ (0.05) $T=6.4$ (0.6) ms
5	$S=5.3$ (1.8) $R=0.68$ (0.11) $T=7.5$ (1.3) ms	$S=2.8$ (0.8) $R=0.62$ (0.11) $T=8.5$ (1.1) ms	$S=1.59$ (0.12) $R=0.70$ (0.05) $T=6.4$ (0.6) ms
6	$S=2.9$ (0.5) $R=0.83$ (0.2) $T=8.4$ (0.7) ms	$S=3.8$ (0.6) $R=0.80$ (0.03) $T=7.5$ (0.5) ms	$S=1.85$ (0.5) $R=0.90$ (0.03) $T=5.9$ (0.7) ms
7	$S=0.73$ (0.04) $R=0.88$ (0.01) $T=6.9$ (0.2) ms	$S=0.45$ (0.04) $R=0.89$ (0.01) $T=8.7$ (0.3) ms	$S=1.4$ (0.3) $R=0.40$ (0.15) $*T=7.8$ ms
8	$S=3.85$ (0.7) $R=0.63$ (0.10) $T=5.6$ (1.1) ms	$S=1.27$ (0.15) $R=0.70$ (0.04) $T=7.7$ (0.5) ms	$S=1.12$ (0.12) $R=0.79$ (0.03) $T=6.2$ (0.4) ms

mid-intensity levels were generally within 0.1 dB. Also shown in this figure are the best-fit predictions of the model to these data.

Table II contains the values of the parameters and their confidence limits for the model predictions plotted in Fig. 3. Using the three free parameters  $R$ ,  $T$ , and  $S$ , 19 out of the 24 sets of data were successfully fitted with the fitting algorithm without excessive parameter dependence leading to uncertainty about parameter values. Threshold data for subjects 1 and 7, and all three sets of data for subject 2, required at least one of the three parameters to be fixed in order to reduce the parameter dependence. These fixed parameter values were chosen based on the analysis of the remaining data described below, and are marked with an asterisk in Table II.

The 19 sets of data which were successfully fitted with three free parameters were used to investigate the relationships between the parameter values and intensity level. One-way analysis of variance showed that level did not significantly affect  $R$  for the five subjects where  $R$  was derived for all three levels ( $p=0.91$ ). In addition, paired  $t$  tests showed no effect of level on  $R$  between comfortably loud and mid-level stimuli for seven subjects ( $p=0.93$ ), and between mid-level and threshold ( $p=0.89$ ) or comfortably loud and

threshold stimuli ( $p=0.53$ ) for five subjects. The mean (and standard deviation) of the 19 estimates of  $R$  was 0.77 (0.11).

One-way analysis of variance showed that level also did not significantly affect  $T$  for the five subjects where  $T$  was derived for all three levels ( $p=0.47$ ). Similarly, paired  $t$  tests showed no effect of level on  $T$  between comfortably loud and mid-level stimuli for seven subjects ( $p=0.29$ ), and between mid-level and threshold ( $p=0.33$ ) or comfortably loud and threshold stimuli ( $p=0.37$ ) for five subjects. The mean (and standard deviation) of the 19 estimates of  $T$  was 7.3 (1.1) ms.

One-way analysis of variance showed that level did significantly affect  $S$  for the five subjects where  $S$  was derived for all three levels ( $p=0.02$ ). Paired  $t$  tests for these subjects showed a significant difference in  $S$  between comfortably loud and threshold stimuli ( $p=0.02$ ) and between mid-level and threshold stimuli ( $p=0.04$ ), although the difference between comfortably loud and mid-level stimuli did not reach significance for a paired  $t$  test with seven subjects ( $p=0.24$ ). The means and standard deviations for  $S$  for the five subjects where all three levels were analyzed were: at threshold 1.4 (0.1); at mid-level 2.4 (1.0); and at comfortably loud level 3.3 (1.5).

The above analysis was used to fix various parameters in the five remaining sets of data, in order to obtain the best fit that was consistent with the other data sets. In the case of the threshold data for subjects 1 and 7, initial fitting with all three parameters produced a situation where  $S$  was highly dependent on  $R$  (and *vice versa*). Based on the insignificant effect of level on  $R$  and  $T$  values seen in the other subjects, best fits were obtained first with the  $R$  value set to the average of the other two  $R$  values at the higher levels, and then with the  $T$  value similarly fixed. The best outcome of these two fitting procedures is shown in Fig. 3, and the parameters used are in Table II. For subject 1, the best-fitting option was with  $R$  fixed, and for subject 7 with  $T$  fixed.

Subject 2's data were more difficult to fit, as the initial fitting with all three parameters resulted in  $S$  being highly dependent on  $R$  for all three stimulus levels. This was due to the monotonic exponential shape of the data. It can be seen from Fig. 2 that this shape is consistent with a low value of  $R$ , where refractory effects are likely to have only a small influence. All values of  $R$  below about 0.3 produced an equally good fit (with  $S$  covarying with the  $R$  value). Given that the  $R$  values for other subjects were all greater than 0.6 (with one exception), subject 2's  $R$  value was set to 0.3, which was the maximum value that produced a reasonable fit. Since  $T$  had a very insignificant effect on the model prediction for this  $R$  value, it was set to the average (7.3 ms) found for the 19 well-fitted sets of data. It can be seen from Fig. 3 that the model predictions for these values (solid lines) fail to predict the data shape for long interpulse intervals. To improve the model predictions at longer interpulse intervals, it was necessary to narrow the integration time window. The fitting procedure described above for subject 2 was repeated at successively smaller ERDs and the  $\chi^2$  values compared. It was found that ERD values within 0.5 ms of 5.5 ms provided the best model predictions. The dotted lines in Fig. 3 show the predictions for an ERD of 5.5 ms for subject 2.

### B. Experiment 2: 4-ms repetition period

The results of experiment 2 are shown for subjects 1–4 in Fig. 4, where the reductions in current for the two-pulse/period stimuli (in dB relative to that of the equally loud one-pulse/period stimulus) are plotted against the test interpulse interval. Also shown in this figure are the best-fit predictions of the model for these data. The values of  $S$  which were derived from experiment 1 at comfortable loudness were used to make the model predictions. The only fitting parameter was the ratio  $\theta_{500}/\theta_{250}$ . It should be noted that this parameter influences the vertical position of the predicted curve and not its slope, the latter being determined by  $S$  and the ERD of the time window. The two fitted lines for subject 2 correspond to the two ERD values used for this subject in experiment 1. It can be seen in Fig. 4 that the slopes of the predicted curves match the slopes of the data fairly well, thus lending support to the  $S$  values obtained from experiment 1. The smaller ERD value of 5.5 ms for subject 2 improved the fit of the slope for these data, as it did in experiment 1. The values of the ratio  $\theta_{500}/\theta_{250}$  (and the standard errors) for subjects 1–4 respectively were 0.671 (0.003); 0.69 (0.1); 0.568 (0.003); and 0.655 (0.006). The use

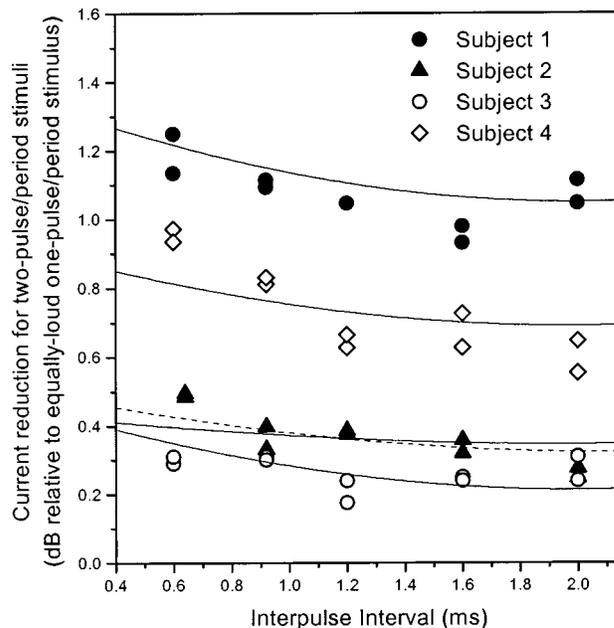


FIG. 4. The current reduction required to make the two-pulse/period stimuli equal in loudness to the one-pulse/period stimulus in experiment 2, along with the best-fit model predictions. As in Fig. 3, the two results for each interpulse interval are represented separately, and the dotted line for subject 2 represents the prediction for ERD=5.5 ms.

of a smaller ERD for subject 2 did not produce a significantly different ratio [0.682 (0.006)]. Thus, for all four subjects, the model predicted that doubling the overall rate from 250 to 500 Hz would lead to a drop in the average spike probability by a factor of approximately  $\frac{2}{3}$ .

### C. Experiment 3: Equal loudness for different rates

The results of experiment 3 are shown in Fig. 5 for subjects 1, 3, 5, and 6, where the current reduction relative to the current for the equally loud 50-Hz stimulus is shown for rates from 20 to 750 Hz. It can be seen that, as rate increased, there was a monotonic increase in the current reduction required to maintain equal loudness. The threshold data are in accordance with previous published results for short duration pulses as described in the Introduction. The comfortably loud data show a similar increase in current reduction for increasing rate, but the magnitude of the effect is smaller than that for threshold, leading to a steadily increasing dynamic range.

Unlike the stimuli in the first two experiments, the total number of stimulus pulses in experiment 3 covaried with interpulse interval for the different rates. The model predicts that there will be two opposing influences on integrator output as rate increases: there will be an increasing number of pulses within the integration window; and there will be a decrease in excitation per pulse caused by a falling spike probability (as seen in experiment 2).

The effect of the increasing number of pulses in the integration time window was first calculated (i.e., keeping spike probability constant). The lines in Fig. 5 show the predicted current reduction due to this effect alone, using an ERD of 7 ms and the values of  $S$  derived in experiment 1 for comfortably loud and threshold stimuli. The model predicts

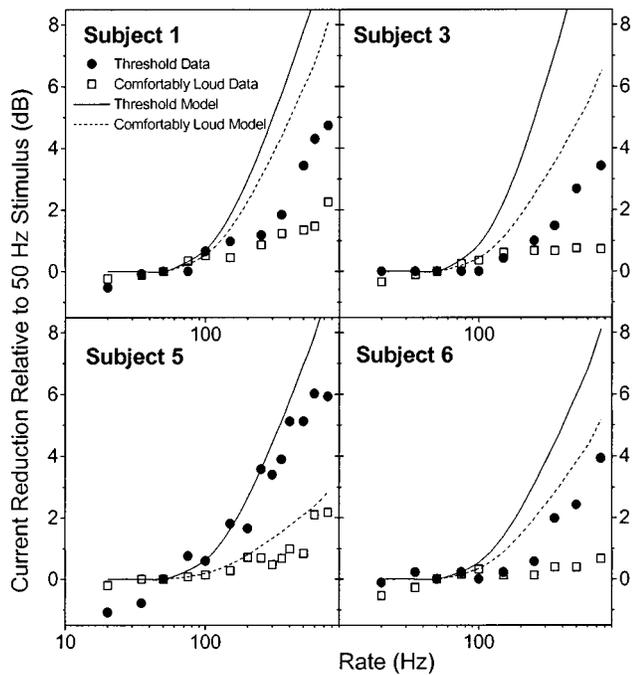


FIG. 5. The loudness balancing results at comfortable loudness and thresholds for stimuli with different rates in Experiment 3. The current reduction relative to the threshold or reference current for the equally loud 50-Hz stimulus is shown, along with the model predictions for the effect of the integration time-window without refractory effects.

that the actual reduction in current should be less than that given by these lines by an amount given by  $(20/S)\log(\theta_r/\theta_{50})$ .

The ratios  $\theta_r/\theta_{50}$  were calculated from the vertical distance between each data point and the corresponding curve in Fig. 5. These were then multiplied by the subject's  $R$  values (equal to  $\theta_{50}$ ), determined from experiment 1, to create the derived functions of  $\theta_r$  seen in Fig. 6.

Since  $\theta_r$  was an arbitrary rate-dependent fitting parameter, which would allow a perfect fit to any data of current adjustment versus rate, the plausibility of our model then rests with the appropriateness of the derived functions of  $\theta_r$ , as descriptions of the averaged spike probabilities across auditory neurons. Javel (1990) measured the spike probability versus intensity for a single auditory neuron in a cat at different stimulation rates. The spike probability versus rate data, interpolated from his Fig. 17.21 for four fixed intensities within about 3 dB of the neuron's threshold of activation, are plotted in Fig. 6, in the same panel as subject 3's data. The behavior of the predicted average spike probability is consistent with these data, given that it represents an average over many neurons all at different levels in their dynamic range. The proportion of neurons that are activated at saturation probability would, in general, make the slope of the average data less steep than for a single neuron near threshold.

There are two further points of interest in the derived average spike probability versus rate functions. First, they predict that spike probability is essentially unaffected by stimulus rates less than 100 Hz, which is consistent with the absence of refractory changes for time intervals longer than about 10 ms. (In some cases, there is a slight drop in derived

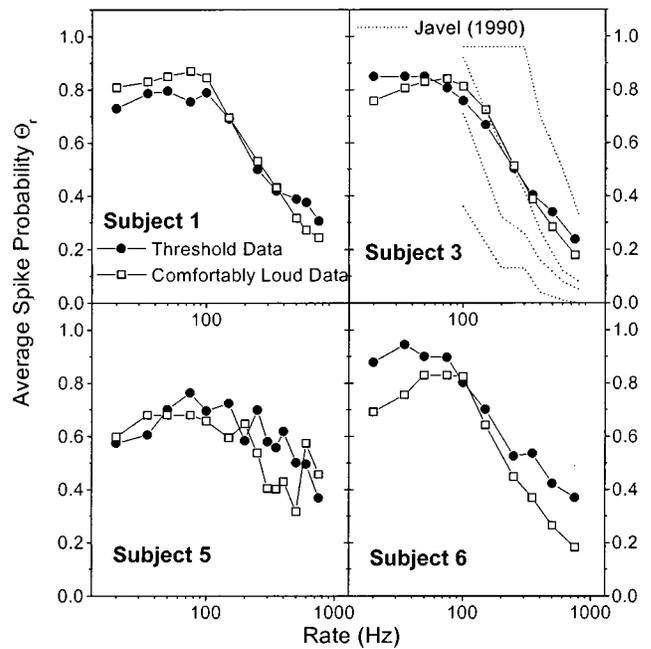


FIG. 6. The predicted effect of rate on  $\theta_r$  derived from Fig. 5, assuming that the vertical distances between the data and the corresponding line in that figure are due to a reduction in average spike probability, and that  $\theta_{50}$  is equal to the  $R$  value (at threshold or comfortable loudness) determined in experiment 1 (Table II). The four dotted lines in the second panel show the spike probability versus rate functions extrapolated from Fig. 17.21 in Javel (1990), for a single neuron at four different current levels: 51, 51.5, 52, and 54 dB (*re*: 1  $\mu$ A).

spike probability below 50 Hz, which will be considered below.) Second, the curves for threshold and comfortably loud levels are similar, implying that the effect of rate on average spike probability is largely independent of stimulation level.

#### IV. DISCUSSION

The results of experiment 1 have shown that the effect of changing the test interpulse interval within pairs of pulses having a 20-ms repetition period could be successfully accounted for by a model which considered both the effects of refractoriness on the second of the two pulses,<sup>3</sup> and a sliding central integration window with ERD of 7 ms. The application of the model to the experimental results allowed the individual specification of the "average" refractory characteristics ( $R$  and  $T$ ), and the transformation between current change and neural excitation change (given by the exponent  $S$ ). In experiments 2 and 3, the effects of refractoriness were modeled as a reduction of spike probability with increasing rate. The application of the model, using the values of  $S$  which were derived in experiment 1, provided good predictions for the results of experiments 2 and 3, and led to a plausible prediction of how spike probability changes with rate. Thus, the model was able to consistently describe the effect of interpulse intervals on loudness for three different stimulus structures.

As mentioned in the Introduction, refractoriness is only one of several factors that can affect the amount of excitation from a pulse which is preceded by another pulse. For stimuli which use very long pulse durations (over 1 ms), there ap-

pears to be an additional inhibitory mechanism which operates after subthreshold pulses, and, for pulses separated by less than 400  $\mu$ s, there is possibly an excitatory mechanism whereby charge can be integrated on the neural membrane across more than one pulse to achieve a response spike. If either of these stimulus conditions are present (on a single electrode), then these effects would have to be incorporated into the function which describes the excitation for the second pulse compared to that for the first pulse [Eq. (1)]. For the stimuli used in these experiments, and for stimuli typically produced by existing speech processors, these conditions are not expected to arise, and a model based on refractoriness alone is sufficient to describe the peripheral temporal effects.

### A. Average peripheral refractory characteristics

If the model assumptions about peripheral neural refractory behavior and the subsequent process of neural spike integration are valid, then it is possible to make inferences from the data about the refractory behavior of the activated neural population of individual subjects who took part in the experiment. Single-neuron spike probabilities are determined by the magnitude of the stimulus current relative to the neural threshold until saturation occurs. Similarly, the value of  $R$  is determined by the magnitude of the stimulus current compared to the distribution of neural thresholds that are lower than that current. The  $R$  values from experiment 1 (for 19 sets of data) had a mean value of 0.77, and ranged from 0.4–0.9 (although subject 2 may have had values lower than this range). The fact that, on average,  $R$  remained constant across the subjects' dynamic ranges implies that, as current increased, and neurons activated at the lower current moved towards a higher or saturated spike probability, the average spike probability remained relatively constant due to other neurons, with higher thresholds, being activated at lower spike probabilities. Thus, for the "average" subject, the increase in total excitation per pulse arising from an increase in current is due mostly to an increase in the number of neurons activated, rather than any increase in the average spike probability across those neurons.

The  $T$  values (the mean time that neurons which fired on the first pulse remained refractory) depend in a similar fashion on the relative magnitude of the stimulus current compared to the distribution of neural thresholds that are lower than that current. It is not surprising, then, that the mean value of  $T$  (7.3 ms) was also not significantly affected by stimulus level. The interspike interval versus "neural sensation level" data for single neurons measured by Parkins (1989) indicate that a refractory time of 7.3 ms would be expected for a neuron activated less than 1 dB above its threshold. Thus, the average spike probability and interspike interval data obtained in these experiments support the hypothesis that, as current increases, significant numbers of additional neurons are recruited, leading to large proportions of neurons being activated close to their thresholds, regardless of the stimulus current.

The rate of recruitment of neurons with level is likely to be different for each subject, depending on the local density and type of neural fibers surrounding the stimulation site.

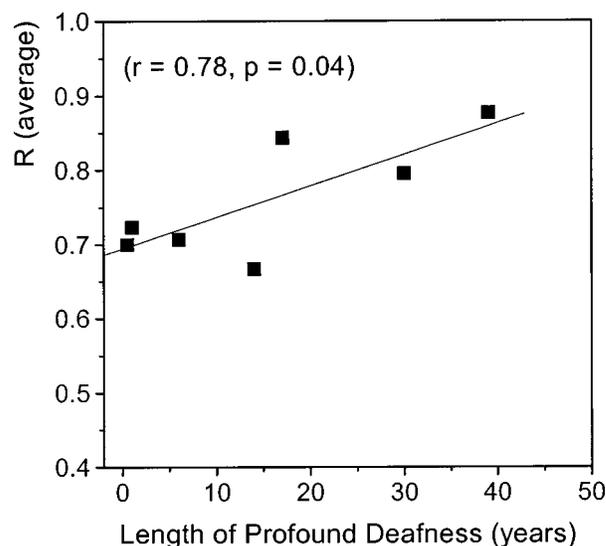


FIG. 7. The relationship of  $R$  determined in experiment 1 (averaged over the three levels) to the length of profound deafness prior to implantation, for the seven subjects where  $R$  values were uniquely determined. The correlation and statistical significance ( $r$  and  $p$ ) are shown in the panel.

This would lead to individual differences in the way  $R$  and  $T$  vary with level. For example, subject 7 had a large increase in  $R$  (0.4 to 0.9) between threshold and mid-dynamic-range (see Table II). This could be due to a lower density of neural fibers at sites more distant from the stimulation site, so that at higher currents there were not enough neurons being recruited to maintain a low average spike probability.

It is possible that the different average spike probabilities among subjects may reflect the overall density of surviving neural fibers. A subject with a very sparse population of nerve fibers may have a slow rate of recruitment of additional fibers as the current is increased, and this would lead to a higher average spike probability than a subject with a dense population of available nerve fibers. One factor which may affect the density of surviving nerve fibers is the length of profound deafness. For our subjects there was a significant relationship between the  $R$  values (averaged across the three levels) and the length of profound deafness before implantation. Figure 7 shows the average spike probabilities and lengths of profound deafness for seven subjects along with the results of linear regression analysis ( $r=0.78$ ,  $p=0.04$ ). (Subject 2 was omitted, as no definite  $R$  predictions were obtained for him.) This result supports the proposition that there is a loss of neural density over time in the absence of auditory stimulation, and that this leads to higher average spike probability among the stimulated fibers. Another factor which may influence the density of surviving nerve fibers is the etiology of deafness. Three of the subjects (2, 5, and 8) had a partially conductive deafness (due to otosclerosis), and so may have more surviving nerve fibers than those whose deafness was wholly sensorineural. These three subjects had three out of the four lowest average  $R$  values (assuming that subject 2 had a value below 0.65).

In summary, then, the results of the experiments suggest that average spike probability does not change much with level of stimulation, but that there are differences in spike

probabilities among the subjects which are consistent with varying degrees of neural survival.

## B. The current-to-excitation transformation

To fit the data from experiment 1 with the model, the scaling factor applied to the current adjustment to transform it into neural excitation units needed to be significantly larger for higher levels of stimulation than that needed at threshold. Given the assumptions of the model, this result suggests that a single simple power function is not appropriate to describe the current-to-excitation transformation over the entire dynamic range. In contrast, the data imply that a ratio current change will cause a larger ratio increase in neural excitation at higher stimulus levels compared to lower levels.

The same conclusion was reached by Nelson *et al.* (1996), who studied the effect of electrical stimulus level on intensity discrimination. They found that, for most subjects, Weber fractions decreased as a power function of intensity ( $I$ ) relative to absolute threshold. When the electrical intensity was normalized between subjects by using “percent of the dynamic range” on a dB scale (%DR), the slope of the Weber function (averaged across seven subjects) was  $-0.08$  dB/%DR, i.e.,

$$Wf_{dB} = 10 \log (\Delta I/I) = -0.08\{\% DR\} + 10 \log (\beta). \quad (4)$$

The  $10 \log (\beta)$  term in the above equation represents the subject’s overall sensitivity to intensity change. They interpreted their data as suggesting that the effects of current on neural excitation increased with increasing level. This interpretation was based on the assumption that the central decision device for intensity increment was similar to that in acoustic hearing, and involved a level-invariant criterion based on a set neural excitation ratio increment. It is possible to derive how our scaling factors (of current to neural excitation on log scales) would have changed with level for their seven subjects to produce the relationship in Eq. (4). The derived scaling factors for different points in the dynamic range (as a ratio of the threshold factor) are shown in Fig. 8, along with the ratios from our experiment. The loudness that our subjects called “comfortably loud” would probably have been considerably lower than the “maximum acceptable loudness” measured by Nelson *et al.*, so we have denoted our data by arrows encompassing 70%–90%DR for “comfortably loud” and 35%–45%DR for the “medium level.” It can be seen that our results are broadly consistent with, although somewhat lower than, the values derived from the Nelson *et al.* paper. The difference may be due to the considerable variability in both subject groups.

Nelson *et al.* proposed a simple qualitative model in which the relative increase in excitation for higher currents is caused by the stimulation activating peripheral neural processes near threshold, and more distant axonal processes at higher levels. The (presumably) less dense residual peripheral processes would need a larger current increment to recruit more neurons, and have a less steep rate-intensity function, compared to the more dense axonal processes, which would need a smaller current increment to recruit the same number of neurons, and where there is a steeper rate-intensity function. This model is partially consistent with our

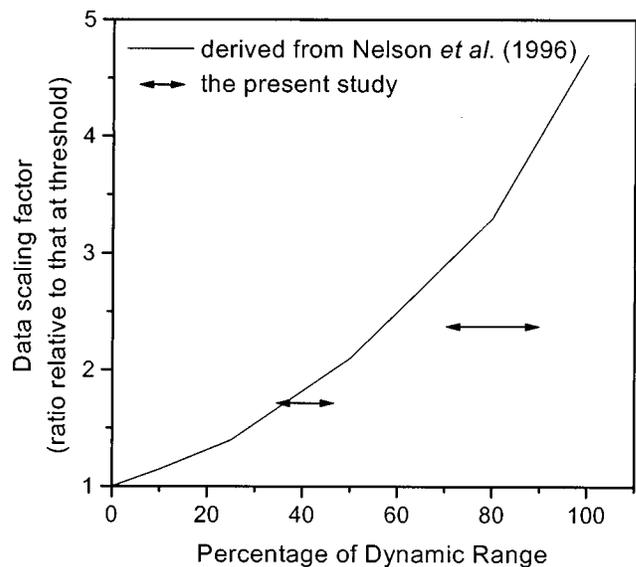


FIG. 8. The change in the power exponent ( $S$ ) of the current-to-excitation transformation (ratio relative to that at threshold) versus level within the dynamic range, derived from the data presented in Nelson *et al.* (1996) from seven subjects (solid line), along with the ratios of scaling factors (averaged over five subjects) from our experiment 1 (arrows encompassing the most likely regions of the dynamic range for medium and comfortably loud levels).

data. However, we propose that, in the “average” subject, the increase in relative excitation at higher levels is due more to increases in the number of neurons activated than any increase in the average spike probability.

In the individual implantee, however, there is a role for both factors in explaining the different patterns of loudness growth. Our subject 5 showed a very large increase in scaling factor between threshold (1.6) and comfortably loud (5.3) (see Table II), indicating a large increase in relative effect of current on excitation. At the same time, the average spike probability changed very little between threshold and comfortably loud levels, indicating that the increase in excitation was mostly due to an increased number of neurons activated. In contrast, subject 7 is atypical of our subject group, in that her scaling factor decreases above threshold, whereas her spike probability more than doubles between threshold and higher levels. This is consistent with the recruitment rate of neurons not being sufficient to maintain a low average spike probability, and in this case, the increase in spike probability would significantly contribute to the total increase in excitation with level. Her large dynamic range (see Table I) may therefore be due to low neural density at sites distant from the electrode.

Nelson *et al.* found that, in their subjects, a flat and very sensitive Weber function (corresponding in our experiment to a large and relatively constant scaling factor) was associated with a smaller dynamic range and superior ability to discriminate electrodes. Our subject 6 would fit this psychophysical description fairly well, and he has both excellent electrode discrimination and speech discrimination ability. However, our subject 7 has similar electrode and speech discrimination abilities to subject 6, but has a very large dynamic range and low sensitivity to level change (inferred from the small scaling factor). Thus the relationship between

intensity discrimination and electrode discrimination may be more complicated than that proposed by Nelson *et al.* More research is needed to define better what neural response characteristics are required for good electrode discrimination. For example, it is not known whether dense and small areas of neural activation would be easier to discriminate than large, partially overlapping regions.

### C. The central equal-loudness decision device

As mentioned earlier, the way the central decision device might store or integrate the outputs of the shorter sliding temporal integration window for periods of up to 200 ms was not specifically accounted for in our model. In the case of very low-rate stimulation (experiment 1, and the low rates of experiment 3), where the stimulation period is longer than the ERD, the window output fluctuates with the same period as the rate, and the simplifying assumption was made that the maximum of the output fluctuation is used to compare loudness. If the loudness-decision device used a running average (over about 200 ms) rather than the maximum output, this would have minimal effect on the model predictions for experiments 1 and 2, where the one- and two-pulse/period stimuli have the same period. However, in experiment 3, any “averaging” process would reduce the loudness estimation for low rates (<about 100 Hz) where the sliding window output fluctuates. Figure 5 shows that, for some of the data, there is a reduction in loudness as rates fall below 50–100 Hz, and this reduction is reflected in a reduction of the inferred average spike probability (Fig. 6). However, the loudness reduction is more likely to be due to an averaging process over longer periods by the central decision device.

### V. CONCLUSIONS

These series of experiments have measured and modeled the effect on loudness of varying the interpulse intervals within stimulus pulse trains with short-duration pulses (100  $\mu$ s and less). The temporal model of loudness perception, which incorporated peripheral refractory effects, a sliding central integration time window, and a central loudness decision device, was able to predict the data from the three experiments in a consistent way. The output of the model allowed comparisons among subjects of the average refractory characteristics of the activated neural population (in terms of average spike probability, and average refractory time). Since these factors were not correlated with stimulus level, it can be concluded that the most significant contributor to increase of total excitation with current is an increase in number of neurons activated, rather than an increase in their average spike probability.

The model predicted that the relationship between change in total excitation and change in current is not a constant power function across the entire perceptual dynamic range. Its prediction that, for most subjects, a specific dB change in current causes an increasing dB change in total excitation as the current increases is analogous to a prediction made from intensity difference limen experiments by Nelson *et al.* (1996).

In summary, the experiments have shown that it is possible to predict neural response behavior from psychophysical performance. This information is potentially useful both in understanding differences in implantees’ speech perception performance, and in devising more advanced ways of controlling the stimulus response in order to improve perception of complex signals such as speech.

### ACKNOWLEDGMENTS

This research was funded by the Human Communication Research Centre of the University of Melbourne, and the National Health and Medical Research Council of Australia. The authors wish to acknowledge the support and encouragement of Professor Graeme Clark and the patience and cheerful willingness of our subjects, without whom this research would not be possible. We also thank Chris Plack and Leslie Collins for their helpful comments on two earlier versions of the manuscript.

<sup>1</sup>In general, the duration of an integration window depends on the physical quantity being integrated. For example, in some acoustic models the signal intensity is integrated, and in others, the signal intensity is transformed by bandpass filtering and nonlinear compression before integration takes place. The more compression which is applied to the input signal, the smaller the ERD of the integration window which explains the experimental data. The integration window with ERD=7 ms used here (from Moore *et al.*, 1996) was derived from a model with a bandpass filter and nonlinear compression before integration. We have assumed that these initial stages model the transformation, due to cochlear processing, from acoustic intensity to units of peripheral excitation, and hence that the integration window in our model is operating on the same physical units as the acoustical model, when an analogous transformation of current intensity to excitation is applied before the integration window.

<sup>2</sup>These subjects used common-ground stimulation mode (also in their clinical speech processor maps) because there were suspected intermittent shorting problems between particular electrodes (distant from the ones used in this experiment). This electrode configuration will not produce uncomfortable sensations, even if shorting occurs.

<sup>3</sup>It is interesting to ask whether a model which did not include these refractory effects would be able to predict the data consistently across the three experiments. To investigate this question, we fitted the data from experiment 1 to a model where there were two parameters: an integration time-window ERD, which was subject but not level dependent, and the power exponent  $S$ . This model will always predict a monotonic decrease in current reduction with widening interpulse interval. For some subjects the goodness-of-fit (as measured by  $\chi^2$ ) was comparable, but in others (such as subject 7), where the data were clearly nonmonotonic, the goodness-of-fit was significantly worse. The derived parameters showed a much larger intersubject variability than for the model with refractory effects: values of ERD varied between 5 ms (subject 2) and infinity (subject 3), and values of  $S$  between 2.0 (subject 2, threshold) and 65 (subject 1, mid-level). In the no-refractory-effects model, the ERD and  $S$  values completely define the expected results for experiments 2 and 3. In only one case (subject 3) were the predictions of this model for experiment 2 reasonably close to the data. For the other three subjects the slope of the model prediction was too flat (subjects 1 and 4), or gave absolute values about a factor of 2 different from the data (subject 2, half the data values; subject 1, double the data values). The predictions for experiment 3 also failed to match the data accurately for three out of four subjects (1, 3, and 6), with the threshold current changes being underestimated and the comfortable-level changes being overestimated.

Buus, S., and Florentine, M. (1985). “Gap detection in normal and impaired listeners: the effect of level and frequency,” in *Time Resolution in Auditory Systems*, edited by A. Michelsen (Springer-Verlag, New York).  
Butikofer, R., and Lawrence, P. D. (1979). “Electrocutaneous nerve stimulation. II: Stimulus waveform selection,” *IEEE Trans. Biomed. Eng. BME-26*, 69–75.

- Carhart, R., and Jerger, J. (1959). "Preferred method for clinical determination of pure tone thresholds," *J. Speech Hear. Disord.* **24**, 330–345.
- Green, D. M., and Forrest, T. G. (1988). "Detection of amplitude modulation and gaps in noise," in *Basic Issues in Hearing*, edited by H. Duifhuis, J. W. Horst, and H. P. Wit (Academic, New York).
- Javel, E. (1990). "Acoustic and electrical encoding of temporal information," in *Cochlear Implants: Models of the Electrically Stimulated Ear*, edited by J. M. Millar and F. A. Spelman (Springer-Verlag, New York), Chap. 17, pp. 247–295.
- Levitt, H. (1971). "Transformed up-down methods in psychoacoustics," *J. Acoust. Soc. Am.* **49**, 467–477.
- Marquardt, D. W. (1963). "An algorithm for least-squares estimation of non linear parameters," *J. Soc. Ind. Appl. Math.* **11**, 431–441.
- Moon, A. K., Zwolan, T. A., and Pfingst, B. E. (1993). "Effects of phase duration on detection of electrical stimulation of the human cochlea," *Hearing Res.* **67**, 166–178.
- Moore, B. C. J., Peters, R. W., and Glasberg, B. R. (1993). "Effects of frequency on the detection of decrements and increments in sinusoids," *J. Acoust. Soc. Am.* **94**, 3190–3198.
- Moore, B. C. J., Peters, R. W., and Glasberg, B. R. (1996). "Detection of decrements and increments in sinusoids at high overall levels," *J. Acoust. Soc. Am.* **99**, 3669–3677.
- Moore, B. C. J., Glasberg, B. R., Plack, C. J., and Biswas, A. K. (1988). "The shape of the ear's temporal window," *J. Acoust. Soc. Am.* **83**, 1102–1116.
- Nelson, D. A., Schmitz, J. L., Donaldson, G. S., Viemeister, N. F., and Javel, E. (1996). "Intensity discrimination as a function of stimulus level with electric stimulation," *J. Acoust. Soc. Am.* **100**, 2393–2414.
- Oxenham, A. J., and Moore, B. C. J. (1994). "Modeling the additivity of nonsimultaneous masking," *Hearing Res.* **80**, 105–118.
- Parkins, C. W. (1989). "Temporal response patterns of auditory nerve fibers to electrical stimulation in deafened squirrel monkeys," *Hearing Res.* **41**, 137–168.
- Peters, R. W., Moore, B. C. J., and Glasberg, B. R. (1995). "Effects of level and frequency on the detection of decrements and increments in sinusoids," *J. Acoust. Soc. Am.* **97**, 3791–3799.
- Pfingst, B. E., and Morris, D. J. (1993). "Stimulus features affecting psychophysical detection thresholds for electrical stimulation of the cochlea. II: Frequency and interpulse interval," *J. Acoust. Soc. Am.* **94**, 1287–1294.
- Pfingst, B. E., Holloway, L. A., and Razzaque, S. A. (1996). "Effects of pulse separation on detection thresholds for electrical stimulation of the human cochlea," *Hearing Res.* **98**, 77–92.
- Plack, C. J., and Moore, B. C. J. (1990). "Temporal window shape as a function of frequency and level," *J. Acoust. Soc. Am.* **87**, 2178–2187.
- Plack, C. J., and Moore, B. C. J. (1991). "Decrement detection in normal and impaired ears," *J. Acoust. Soc. Am.* **90**, 3069–3076.
- Shannon, R. V. (1985). "Threshold and loudness functions for pulsatile stimulation of cochlear implants," *Hearing Res.* **18**, 135–143.
- Shannon, R. V. (1989). "A model of threshold for pulsatile electrical stimulation of cochlear implants," *Hearing Res.* **40**, 197–204.
- Shannon, R. V. (1993). "Psychophysics," in *Cochlear Implants: Audiological Foundations*, edited by R. S. Tyler (Singular, San Diego), Chap. 8, pp. 357–388.
- Stypulkowski, P. H., and Van den Honert, C. (1984). "Physiological properties of the electrically stimulated auditory nerve. I. Compound action potential recordings," *Hearing Res.* **14**, 205–223.
- Viemeister, N. F. (1979). "Temporal modulation transfer functions based on modulation thresholds," *J. Acoust. Soc. Am.* **66**, 1364–1380.
- Viemeister, N. F., and Wakefield, G. H. (1991). "Temporal integration and multiple looks," *J. Acoust. Soc. Am.* **90**, 858–865.