

What do Cochlear Implants Teach us About the Encoding of Frequency in the Auditory System?

Johan J. Hanekom

Department of Electrical, Electronic and Computer Engineering
University of Pretoria

ABSTRACT

This article explores the coding of frequency information in the auditory system from the viewpoint of what has been learnt from cochlear implants. Cochlear implants may provide a window on central auditory nervous system function by creating the possibility to separate place and temporal information. An existing model of frequency discrimination in the acoustically stimulated auditory system is extended to include electrical stimulation. To be able to predict frequency difference limens for acoustic stimulation, an important assumption is that one spike per stimulus cycle is available, which may be provided by the existence of a volley principle. It is shown that to predict frequency difference limens for electrical stimulation of the auditory system, it must also be assumed that electrical stimulation causes desynchronization at a central auditory nervous system integration centre. With these assumptions, the model predicts the degradation in frequency discrimination that occurs for electrical stimulation. Finally, it is shown that cochlear implants have not yet proven conclusively that either rate-place coding or temporal coding is predominant in the auditory system.

OPSOMMING

Hierdie artikel ondersoek die kodering van frekwensie-inligting in die ouditiewe stelsel vanuit die oogpunt van wat geleer kan word uit kogleêre inplantings. Kogleêre inplantings kan 'n venster gee op die verwerking wat in die sentrale ouditiewe stelsel plaasvind, deurdat dit die moontlikheid skep om plek- en tydinligting te skei. 'n Bestaande model van frekwensiediskriminasie in die akoesties-gestimuleerde ouditiewe stelsel word uitgebrei om elektriese stimulasie ook in te sluit. Om frekwensiediskriminasie vir akoestiese stimulasie te voorspel is die belangrike aanname dat een aksiepotensiaal per stimulussiklus beskikbaar is, nodig. Dit kan bewerkstellig word deur die werking van 'n sarsiebeginsel. Dit word aangetoon dat frekwensiediskriminasie vir elektriese stimulasie voorspel kan word deur die verdere aanname dat elektriese stimulasie desinkronisasie veroorsaak by 'n sentrale integrasiesentrum. Gegee hierdie aannames, kan die model die verswakking in frekwensiediskriminasie vir elektriese stimulasie voorspel. Laastens word dit aangetoon dat kogleêre inplantings nog nie bo alle twyfel kon bewys dat of plek-tempo-kodering of tyd-kodering van frekwensie die dominante meganisme in die ouditiewe stelsel is nie.

Key Words: frequency discrimination, electrical stimulation, cochlear implants, inter-spike interval, phase-lock coding, rate-place coding.

INTRODUCTION

A long-standing question about frequency analysis in the auditory system is how frequency information is represented: is frequency coded as a temporal code or as a place code (Møller, 1999) or as both? Pure tones are represented as both rate-place information (rate-place coding) and temporal information (phase-lock coding) in the discharge patterns of auditory nerve fibres and the central auditory nervous system, but the extent to which the auditory system uses either representation is unknown.

Rate-place coding is a spectral analysis mechanism whereby the auditory system may combine firing rate information from nerves originating from different places in the cochlea to determine the stimulus frequency. This

method of coding of frequency operates over the entire stimulus frequency range, but is usually presumed to be dominant for the coding of high frequencies (above about 5000 Hz) (Kim & Parham, 1991; Moore, 1973).

Phase-lock coding is a temporal mechanism, wherein the auditory system presumably uses the synchronization of neural discharges to individual cycles of periodic stimuli as a cue to determine the frequency of a pure tone. Phase-lock coding is usually presumed to operate primarily at lower frequencies, since phase-locking is progressively lost as stimulus frequency increases above about 2500 Hz. No phase-locking is observed above 5000 Hz (Johnson, 1980; Rose, Brugge, Anderson & Hind, 1968).

It is possible that both coding mechanisms operate in parallel over a large range of frequencies. So far, neither

neurophysiological studies in animals, nor psychoacoustic experiments in humans have been able to determine to which extent the central auditory system uses either mechanism alone or both mechanisms simultaneously to determine the frequency of a pure tone (Moller, 1999; Johnson, 1980). Whilst previously this may have been regarded as a purely academic question, the development of cochlear implants has made it important to understand how information is coded in the auditory nervous system. This knowledge will influence the stimulation strategies used in cochlear implant speech processors.

Specifically, we need to know what information transmitted to the electrically stimulated cochlear nerve is perceptually significant. Two strategies used in current cochlear implant systems reflect two different approaches. In the Spectral Peak (SPEAK) strategy (Loizou, 1999), which is based on the rate-place mechanism, spectral peaks are extracted and presented to electrodes that are arranged tonotopically. In contrast, the Continuous Interleaved Sampling (CIS) strategy (Loizou, 1999) uses high pulse-rate stimulation to conserve temporal waveform information.

Moller (1999) reviews the roles of temporal and rate-place coding of frequency. He presents convincing arguments in favour of the phase-lock code. His arguments are based on the robustness of the code and on the effects of various kinds of pathology on the impairment of frequency discrimination and pitch perception. It is well established that frequency tuning in the auditory system is a function of sound intensity (Moller, 1999; Johnstone, Patuzzi & Yates, 1986). The location of maximal vibration of the basilar membrane shifts at higher intensities. However, the perception of pitch of pure tones is relatively insensitive to changes in intensity over large intensity ranges. Some models conjecture that it is not the spectral peaks, but rather the complete spectral profile, or the edges of the spectral profile that are used in frequency discrimination (Moore & Glasberg, 1986). Moller (1999) argues that these are just as sensitive to intensity. Moller also reviews data that suggests that impairment of spectral analysis in the cochlea does not affect speech discrimination noticeably, which suggests that spectral analysis might not be important for speech perception.

Cochlear implants provide the opportunity to study the coding of frequency by the rate-place code and by the phase-lock code separately. The perception of frequency as encoded in the rate-place code alone may be studied by using a fixed stimulation frequency and varying the site of stimulation in the cochlea. To explore the role of phase-lock coding alone, stimulation at a fixed position in the cochlea may be used while varying the frequency of stimulation (Townshend, Cotter, Van Compernelle and White, 1987; Blamey, Dooley, Parisi & Clark, 1996; Dorman, Smith, Smith & Parkin, 1994).

This article further explores the coding of frequency in the phase-lock code, using neurophysiological and psychoacoustic data from auditory electrical stimulation as instrument. One motivation for this study stems from the strong arguments by Moller (1999) in favour of phase-lock coding, but the way in which frequency information should be encoded in cochlear implants is also of interest.

APPROACH

A modelling approach is taken here. Note that the models described below are mathematical models of neural

processing that were simulated on computer, and that no real neurons were used. A recent model of frequency discrimination in the acoustically stimulated auditory system (Hanekom & Krüger, 2001) is extended to include the electrically stimulated auditory system. This model provides satisfactory predictions of frequency discrimination in the normal (acoustically stimulated) auditory system. If, in addition, it is found that the extended model can predict psychoacoustic data for the electrically stimulated auditory system, strengthening the arguments in favour of a phase-lock code for frequency may be possible. The applicability to cochlear implant stimulation strategies will be explained in the discussion.

A short motivation for using a modelling approach is provided. The current model predicts psychoacoustic data from neurophysiological data. Apart from being useful to summarize current knowledge, models build understanding of the signal processing that the auditory system has to perform. Models also assist in revealing gaps in our knowledge, whether it is in unknown parameter values, or in our understanding of the signal processing. Models are of necessity idealizations of reality. It is, of course, always possible to build complex models with many free parameters that will be able to model any given set of psychoacoustic data. However, simple models with few free parameters that can accurately predict trends in psychoacoustic data builds confidence that the signal processing that the central auditory nervous system performs is understood.

Plausible models of acoustic frequency discrimination data should account for the absolute values of frequency difference limens (Δf 's) and explain the origin of the bowl shape of the curve of the Weber fraction ($\Delta f/f$) plotted as a function of frequency (e.g. Moore, 1973; Sek & Moore, 1995), without the need to manipulate many free parameters to fit the psychoacoustic data.

Several models exist to explain psychoacoustic frequency difference limens (Δf 's) for the acoustically stimulated auditory system. These models are based on either the extraction of frequency directly from one or more neural spike trains (i.e. a temporal approach) (Goldstein & Sruлович, 1977; Javel & Mott, 1988) or the rate-place code (Javel & Mott, 1988), or both mechanisms simultaneously (Siebert, 1970; Sruлович & Goldstein, 1983). One recent model (Hanekom & Krüger, 2001) is examined in this article and extended to incorporate auditory electrical stimulation data.

The model is based on the following. A listener's ability to discriminate between two signals is limited by neural noise, i.e. the random nature of the neural spike train (Siebert, 1970). Siebert was first to propose the notion that the difference limen in a discrimination task (e.g. frequency or intensity discrimination) is equal to the standard deviation in estimating the magnitude of the stimulus variable (e.g. frequency or intensity). The implication is that estimators may be designed to extract a stimulus variable from its noisy neurally encoded form. The difference limen can then be evaluated by calculating estimation variance.

Based on these ideas, Goldstein and Sruлович (1977) proposed a temporal model of frequency discrimination in which frequency is encoded in inter-spike intervals only. They demonstrated that with the combination of a small number of fibres, sufficient information is available to account for perceptually measured frequency discrimi-

nation thresholds. An extension of their 1977 model (Srulovicz & Goldstein, 1983) accounts for a wider range of psychoacoustic phenomena. The more complex extended model incorporated both temporal and rate-place cues. Like Moller (1999), they concluded that phase-lock coding is a more likely mechanism than rate-place coding for the frequency discrimination task. Wakefield and Nelson (1985) extended the temporal model of Goldstein and Srulovicz (1977) to include intensity effects.

Based on Moller's (1999) strong arguments in favour of a phase-lock code for frequency, and the success of the simple inter-spike interval based model of Goldstein and Srulovicz (1977) in predicting the shape and magnitude of frequency discrimination thresholds, a new model of frequency discrimination was presented in Hanekom and Krüger (2001). This model used a particularly simple description of the statistics of phase-locking, which enabled the authors to construct an optimal estimation mechanism by which frequency information can be extracted from one or more neural spike trains. The objective of the extension to this model is to include the statistics of spike trains evoked by electrical stimulation in order to make predictions about frequency discrimination thresholds in cochlear implants. The statistics of spike trains that result from electrical stimulation are quite different from acoustically evoked spike trains, as will be discussed later.

It should be noted that this paper does not present any hypotheses about the central representation of pure tones. The emphasis is on the interpretation of the frequency discrimination performance of an optimal estimator, presumably located somewhere in the central auditory nervous system, given the statistics of acoustically and electrically evoked spike trains.

METHOD

The phase-lock model of Hanekom and Krüger is extended to incorporate electrical stimulation. This is a mathematical model of neural processing which may be simulated on computer. This model has been considered in more detail in Hanekom and Krüger (2001) and is only briefly outlined here. The proposed extension for electrical stimulation is simple and detail is provided.

ESTIMATOR STRUCTURE

Phase-locking is the tendency of nerve spikes (action potentials) to cluster around multiples of the stimulus cycle at a preferred phase (a specific time relative to the onset of the stimulus cycle). It is assumed that these clusters have Gaussian distributions (Javel and Mott, 1988) of which the variance depends on the amount of phase-locking. Perfect phase-locking occurs when spikes always occur at the same phase. When spikes are also entrained to the stimulus (i.e. spikes occur at each stimulus cycle), calculating the stimulus frequency perfectly is very simple. Thus, the distribution of the spikes around the preferred stimulus phase is a source of noise.

It is assumed that the auditory system can combine spike trains from a number of fibres to obtain a single spike train that has one spike per stimulus cycle. This idea is essentially the same as the volley principle of Wever (1949). Javel (1990) speculated that the great redundancy in auditory nerve fibre innervation of the inner hair cells

may exist to ensure that a spike occurs on every stimulus cycle. Superimposing a number of spike trains results in clusters of spikes, with cluster centres spaced approximately $1/f$ apart. If spike trains from more fibres are superimposed, estimates of the cluster centres become more accurate, resulting in more accurate estimates of the actual stimulus period. This is on the condition that fibres fire on exactly the same preferred phase, which is not necessarily true. Different fibres, tuned to slightly different frequencies, will all phase-lock to the stimulus, but each at its own preferred phase (Javel, 1990). Neurons have been found in the cochlear nucleus that may be able to implement a volley principle by combination of several auditory nerve inputs (Moller, 1999). To achieve this, the integration centre has to compensate for differences in the preferred firing phase. The auditory system may achieve this by variation in fibre length and variation in the strength of synapses (Cook & Johnston, 1999).

In the current implementation of the model, spike trains were not combined explicitly. It was assumed that one spike per stimulus cycle was available. Under this condition, measurements of inter-spike intervals used to estimate frequency are just noisy measurements of the actual period of the stimulus waveform. The problem of obtaining a good estimate of the stimulation frequency from these noisy measurements can be solved with an estimator that is often used in engineering applications, the Kalman filter (Kalman, 1960).

NUMBER OF FIBRES COMBINED

At high intensities, the combination of spike information from just a few nerve fibres will ensure the availability of one spike per stimulus cycle. At lower intensities, the combination of more nerve fibres is required to account for human frequency discrimination data. The probability of missing cycles decreases as the number of fibres to be combined increases. It is estimated from simulations that the current model requires the combination of around 100-200 fibres to ensure one spike per stimulus cycle at all frequencies for acoustical stimulation.

MODEL OF PHASE-LOCKING FOR ACOUSTIC STIMULATION

At high stimulation intensities, for fibres with characteristic frequency (CF) at or close to the stimulus frequency, spikes may occur on each stimulus cycle for low frequencies (lower than about 1000 Hz), although this is usually not so and cycles are often missed (Rose, Brugge, Anderson & Hind, 1968). Spikes can be very scarce at low intensities or when the stimulus frequency is far from the CF of a fibre.

It is assumed in the current model that the central processor integrates spike train information from a restricted area in the cochlea where the strongest activity is found. This corresponds to the average localized synchronized rate (ALSR) model of Sachs and Miller (1985). Thus, for a stimulus well above threshold as used in the current model, it is assumed that the most strongly stimulated fibres fire at their maximum rates.

A further assumption is that spikes cluster around a specific phase of the stimulus cycle according to a Gaussian distribution. The distribution of spikes is Gaussian with standard deviation

$$\sigma_n = \sqrt{k} \frac{1}{2\pi f} \arccos\left(\frac{G(f,A)-0.5}{G(f,A)}\right) \tag{1}$$

where $G(f,A)$ is the synchronization index, k is a scaling factor used to fit neurophysiological data, f is the stimulus frequency, and A is the stimulus intensity. This equation was derived in Hanekom and Krüger (2001).

The synchronization index $G(f,A)$ is a function of both frequency and intensity. $G(f,A)$ may be written as the product of two factors, $G(f,A) = G_1(f) G_2(A)$, where A is intensity in dB SL and f is frequency in Hz. $G_1(f)$ and $G_2(A)$ characterize the variation of the synchronization index with variation in frequency and intensity of the pure tone stimulus respectively. For acoustic stimulation $G_1(f)$ is given by

$$G_1(f) = \frac{0.85}{1 + \left(\frac{f}{3500}\right)^3} \tag{2}$$

and $G_2(A)$ is given by

$$G_2(A) = \frac{1.1 A^{0.3} H}{\sqrt{0.5(A^{0.3})^2 H^2 + K}} - 0.6 \tag{3}$$

Equation 2 and equation 3 are curve fits to typical values of the synchronization index as a function of frequency and intensity respectively. In equation 3, K is a sensitivity constant that controls the threshold of the model fibre, while H is a tuning constant that takes on a maximum value of 1 when the model fibre has CF at the stimulus frequency.

MODEL OF PHASE-LOCKING FOR ELECTRICAL STIMULATION

Some important differences exist between neural synchronization to acoustical and electrical stimulation.

First, electrically stimulated fibres exhibit phase-locked responses with a much higher degree of synchronization (Shepherd & Javel, 1999; Javel, 1990; Van den Honert & Stypulkowski, 1987; Hartmann, Topp & Klinke, 1984). This is demonstrated in period histograms where the peak is much narrower for electrical stimulation than for acoustic stimulation (e.g. Javel, 1990). Thus phase-locking occurs on a very precise phase of the stimulus signal (Hartmann, Topp & Klinke, 1984). For high frequencies (4-8 kHz), phase-locking is weaker (Dynes & Delgutte, 1992). The fibre still discharges regularly, but many stimulus cycles may be skipped, similar to the acoustic case. Phase-locking is maintained at high frequencies (10 kHz) for electrical stimulation, unlike acoustical stimulation which demonstrates no phase-locking above 5 kHz. The synchronization index for electrical stimulation is described by

$$G_1(f) = \frac{0.92}{1 + \left(\frac{f}{3500}\right)^3} \tag{4}$$

which is a curve fit to the synchronization index data in Dynes and Delgutte (1992) as shown in figure 1.

Secondly, the degree of entrainment is much higher for electrical stimulation than for acoustic stimulation. Spikes may occur on each stimulus cycle for low frequencies (below 1 kHz) (Javel & Shepherd, 2000; Javel, 1990).

Thirdly, there is no frequency selectivity for electrical stimulation (there is no basilar membrane filtering), so that all areas activated by the electrical stimulus generate action potentials, regardless of the stimulus frequency or the CF of the stimulated site. In addition, the statistical independence among spike trains in different fibres is lost, so that many fibres fire exactly in phase (Javel, 1990; O’Leary, Tong & Clark, 1995).

Lastly, synchronization increases with stimulus intensity (Shepherd & Javel, 1999). Spike latencies become smaller for more intense stimuli, and spike latencies are shorter for pulsatile stimuli than for sinusoidal stimuli. The probability of firing is a function of the stimulus strength, but the slope of this function is very steep (Van den Honert & Stypulkowski, 1987) so that for pulsatile stimulation at just around 6 dB above threshold, neurons fire at their maximum (entrained) rate. It is assumed that fibres are stimulated well above threshold, so that a value of $G_2(A)=1$ is used for electrical stimulation in the current model.

COMBINATION OF FIBRES FOR ELECTRICAL STIMULATION

If spike trains are combined according to a volley principle, then presumably the auditory system must compensate for differences in the preferred firing phase of different fibres as explained before. Fibres fire exactly in phase in electrical stimulation. Any mechanism that compensates for differences between fibres in acoustic

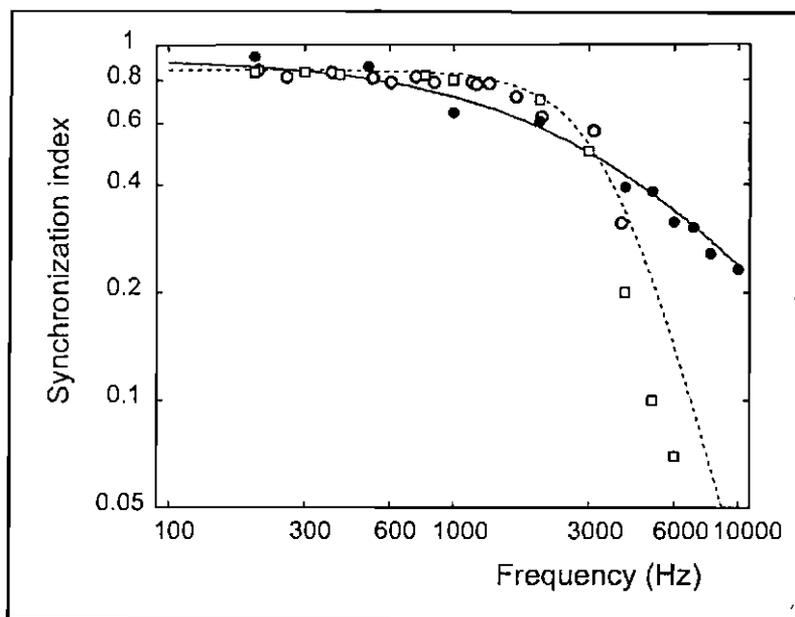


FIGURE 1. Synchronization index as a function of frequency for electrical and acoustic stimulation. The solid curve (electrical stimulation) was calculated from equation 4. The dotted curve (acoustic stimulation) was calculated from equation 2. Open squares are data from Johnson (1980) and open circles are data from Javel and Mott (1988) for acoustic stimulation. Filled circles are data for sinusoidal electrical stimulation from Dynes and Delgutte (1992).

stimulation may in fact disperse spikes around the preferred phase in electrical stimulation as measured at a central integration centre. So for electrical stimulation, it is assumed that phase-locked spike trains arriving at the central integration centre are desynchronized relative to each other. It is assumed that the central integration centre still generates one spike per stimulus cycle, but with larger variance in spike position around the preferred phase than for the acoustic case. The spike position variance is a function of the number of spike trains combined. It is assumed that the spike trains arriving at the central integration centre are statistically independent, so that the spike variances add. If it is assumed that spike trains from around 200 fibres are combined, the spike variance around the preferred phase will be two hundred times larger for electrical stimulation than for acoustic stimulation. Thus

$$\sigma_n = \sqrt{200} \sqrt{k} \frac{1}{2\pi f} \arccos\left(\frac{G_1(f) - 0.5}{G_1(f)}\right) \quad (5)$$

is used for electrical stimulation, with $G_1(f)$ as given in equation 4.

As further motivation for this argument, it is noted that injuries to the auditory nerve cause increases in neural conduction time (and thus temporal dispersion of neural activity). It is known that injuries to the auditory nerve affect speech discrimination ability more than cochlear injuries (Moller, 1999).

IMPLEMENTATION OF THE ESTIMATOR AND SIMULATIONS

Derivation of the Kalman filter equations falls outside the scope of this article, but details may be found in Hanekom and Krüger (2001). The discussion below is intended to elucidate the principles.

Essentially, to apply the Kalman filter, the modelling equations for the generation of spikes must be obtained in a certain format (the state space format, which is often used by engineers). This is relatively simple for the problem stated here. Once this has been done, the noisy measurements of the actual stimulus period may be applied as input to the Kalman filter. From an implementation viewpoint, the Kalman filter is then simply a set of equations solved iteratively to provide an estimate for the input frequency at each time instant.

Spike trains from single fibres were computer generated using the model. Estimates were obtained for frequency by observing the spike train from a single modelled fibre under the assumption that one spike per stimulus cycle was available. Spikes were placed according to a Gaussian distribution with standard deviation σ_n . For acoustic stimulation, $G_1(f)$ of equation 2 was used in equation 1 to calculate σ_n , while $G_1(f)$ as in equation 4 was used in equation 5 for electrical stimulation.

The frequency difference limen Δf was then obtained by assuming it to be equal to the standard deviation in the frequency estimate, following Siebert (1970) and several other authors after him. The standard deviation in the frequency estimate was obtained by repeating the pure tone

stimulus of duration T many (typically 200) times and calculating the standard deviation of the frequency estimate at a specific time. This time was either at the end of the interval T or after 50 observations of inter-spike intervals, as will be explained in the discussion. Values of Δf were obtained as a function of stimulus frequency for both acoustic and electrical stimulation.

The equations describing the model were coded in Matlab, a computer language designed for doing mathematics. Simulations were run on a Pentium II personal computer under the Windows 95 operating system.

RESULTS

$\Delta f/f$ AS A FUNCTION OF FREQUENCY FOR ACOUSTIC STIMULATION

Figure 2 shows the normalized frequency difference limen ($\Delta f/f$) as a function of frequency for acoustic stimulation as predicted by the model. For comparison, $\Delta f/f$ for electrical stimulation is also shown. Frequency discrimination data for acoustic stimulation as measured by Sek and Moore (1995) are plotted on the same axis. The shapes of the two curves for acoustic stimulation are very similar, and both reach minima at 500 Hz. The absolute values of $\Delta f/f$ as predicted by the model correspond well to measured values across the entire frequency range, except at 10000 Hz.

Δf AS A FUNCTION OF FREQUENCY FOR ELECTRICAL STIMULATION

Figure 3 shows the frequency difference limen (Δf) as a function of frequency for electrical stimulation as predicted by the model. Simulation predictions are not shown as the

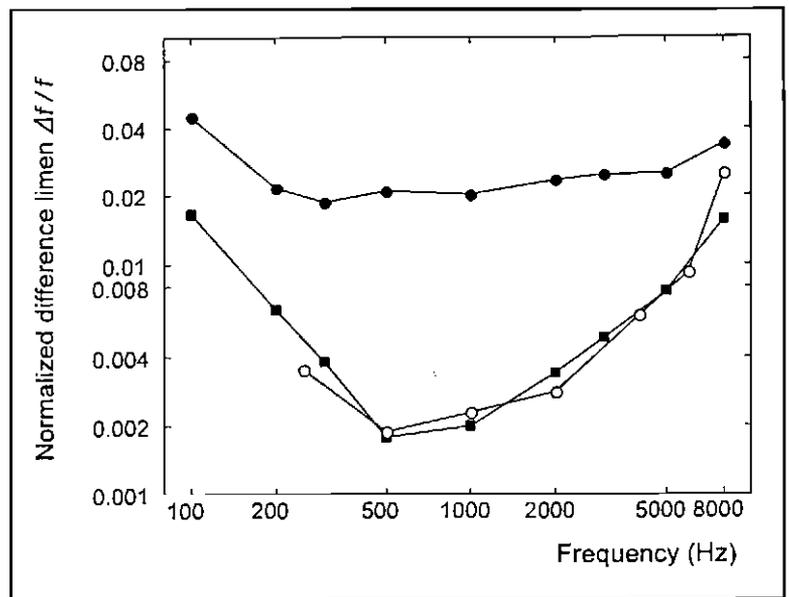


FIGURE 2. Values of the frequency difference limen Δf expressed as a proportion of frequency ($\Delta f/f$) are plotted as a function of the frequency of a pure tone stimulus on logarithmic axes. Filled squares are model predictions for acoustic stimulation, while open circles are the perceptual frequency discrimination data of Sek and Moore (1995). Filled circles are model predictions for electrical stimulation.

Weber fraction $\Delta f/f$ in this figure, as the measurement data was available as Δf . Frequency discrimination data as documented in Pfingst (1988) and Townshend et al. (1987) are plotted on the same axes. The shapes and slopes of the model prediction and the psychoacoustic data are similar. The absolute values of Δf as predicted by the model are smaller than measured values across most of the frequency range. At high frequencies, model predictions may be an order smaller than the psychoacoustic data. To bring the model and data into closer agreement, larger standard deviations may be used for spike position jitter. Alternatively, higher stimulus intensities shift the measured curves downwards (Pfingst, 1988). Unfortunately, it is unknown at which stimulus intensities the data were measured.

DISCUSSION

JUSTIFICATION OF ASSUMPTIONS

The current model rests on two important assumptions. First, it was assumed that the auditory system has some way to ensure one spike per stimulus interval across the entire frequency range. This assumption idealizes known neurophysiological data. More than one spike may occur per stimulus cycle in acoustic stimulation (Rose, Brugge, Anderson & Hind, 1968) and multiple spikes may occur in electrical stimulation (Javel & Shepherd, 2000). Both phases of the electrical stimulation waveform may evoke spikes (Van den Honert & Stypulkowski, 1987) and multiple spikes per phase may occur at higher frequencies of pulsatile stimulation (Javel & Shepherd, 2000).

Because of the way that the current model was formulated, it is a requirement that there is only one spike per stimulus cycle. The Kalman filter will regard more than one spike per stimulus interval as a source of noise. If a small percentage of cycles have either more than one spikes per cycle, or some cycles are skipped, the dominant inter-spike interval is still the period of the stimulus waveform and the central estimator will make the correct estimate (although with larger standard deviation in the estimate). With many cycles not conforming to the one spike per cycle assumption, the central estimator may make an incorrect estimate of the input frequency. A higher likelihood exists that this will happen for electrical stimulation, as spikes may occur on both phases of the stimulus waveform.

Nonetheless, the model may still explain the observed frequency difference limens, because frequency discrimination measurements are differential and do not measure the absolute frequency perceived. The close correlation between the predicted and measured frequency discrimination thresholds suggests the possibility that a central representation of the pure tone exists that is equivalent to the one spike per stimulus interval assumption. This, however, is not what the model intended to prove. Rather, the intention was to show that frequency discrimination thresholds could be explained by spike position jitter in a phase-locked response. This is discussed further in the sequel.

The second assumption was that, because

many fibres fire in phase as a result of electrical stimulation, the net result at the central auditory estimator would be a desynchronization of spike trains, rather than improved synchronization. It is unknown whether data exist which supports this hypothesis. Available data seems to refute this notion. The cochlear nucleus (CN) exhibits greater response diversity than the auditory nerve (O'Leary, Tong & Clark, 1995). Some fibres display phase-locking to the stimulus, while the responses of other fibres are more complex. CN fibres that do phase-lock exhibit very little temporal dispersion of spikes for electrical stimulation (Javel & Shepherd, 2000).

However, as it is not known what the central representation of frequency is, to search for spike trains at the CN output that exhibits larger spike position jitter for electrical stimulation than for acoustic stimulation may be fallacious. It is known that temporal information on the auditory nerve is gradually transformed into a rate-place code at higher levels of the central auditory system, possibly at the level of the CN (Rhode and Greenberg, 1992). Many auditory afferents carrying a phase-lock code converge on CN cells. These fibres should provide at least one spike per stimulus cycle on the input to a CN neuronal assembly. The possibility exists that the phase-lock code may then be transformed directly into a rate-place code without the need for fibres firing at rates up to 5000 Hz. So, not enough is known to be able to prove or disprove the second assumption.

Neither assumption is unrealistic in terms of biological implementation and the results justify the two assumptions to some extent.

As a final comment, the possibility that the model predictions only hold for frequencies below 5000 Hz needs to be pointed out, as no phase-locking is observed at higher (acoustic) stimulation frequencies.

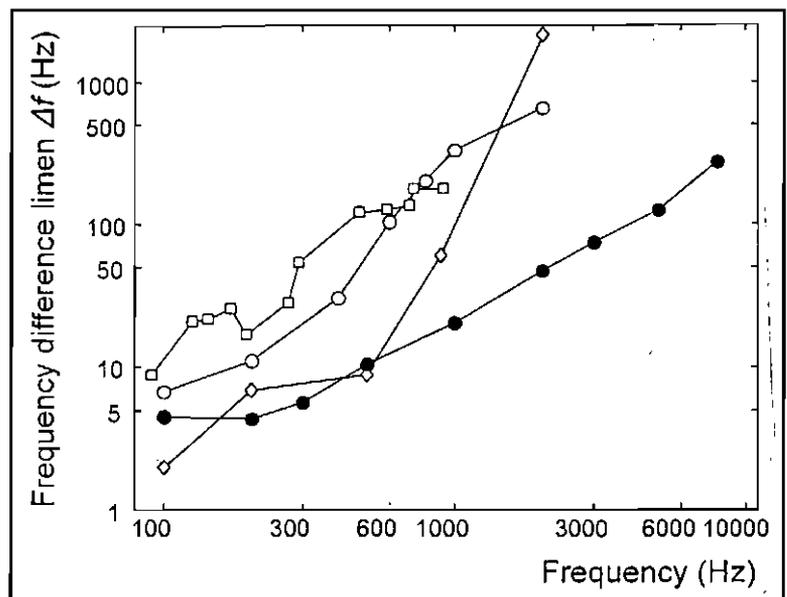


FIGURE 3. Values of the frequency difference limen Δf are plotted as a function of the frequency of electrical stimulation, on logarithmic axes. Filled circles are model predictions for electrical stimulation. Open circles and open diamonds are perceptual frequency discrimination data from two studies (for sinusoidal electrical stimulation) as reported in Pfingst (1988). Open squares are data for pulsatile electrical stimulation (Townshend et al., 1987).

THE ORIGIN OF THE SHAPE OF THE $\Delta f/f$ FREQUENCY CURVE

The Δf obtained is primarily a tradeoff between two parameters of the model: the number of observations and the spike jitter around the preferred phase of the stimulus cycle.

To account for psychoacoustic data below 500 Hz, stimulus duration T is limited to 100 ms so that the number of observations decreases with lower frequencies, which results in a growth in $\Delta f/f$ at lower frequencies consistent with psychoacoustic data. This choice for T is consistent with known auditory integration times (Eddins & Green, 1995). At these frequencies, $\Delta f/f$ is determined primarily by the number of observations available. At higher frequencies the number of observations in the 100 ms time interval grows. It was found that the number of observations needs to be close to $N=50$ to achieve the same $\Delta f/f$ values as the psychoacoustic data for acoustical stimulation. Larger N results in little further decrease in $\Delta f/f$. At higher frequencies (above 500 Hz), the spike jitter becomes a systematically growing percentage of the stimulus period. This plays the primary role in the growth of $\Delta f/f$ at these frequencies.

WHAT DO COCHLEAR IMPLANTS TEACH US ABOUT THE CODING OF FREQUENCY IN THE AUDITORY SYSTEM?

Psychoacoustic data from cochlear implants seem to refute the idea that temporal coding mechanisms are utilized by the central auditory system to extract frequency information from the neural spike train, as the frequency difference limens are much poorer than for normal-hearing listeners, even though there is much more synchronization to the stimulus waveform in electrical stimulation. The current model demonstrates (with reasonable assumptions) that a central auditory estimator that uses interspike intervals to calculate frequency may fare worse with electrical stimulation than with acoustic stimulation. This is consistent with psychoacoustic data. So at least the current model indicates that we cannot rule out temporal mechanisms as a mechanism for frequency coding.

It is known that cochlear implant signal processing strategies based on preserving the temporal pattern (e.g. CIS) are generally more successful than strategies based on vocoders (e.g. SPEAK) (Loizou, 1999), which supports the argument in favour of phase-lock coding.

Also, recent studies have shown that fewer channels in a speech processor can lead to equally good or better speech discrimination (Fishman, Shannon & Slattery, 1997), but if fewer than 4 to 6 channels are used, performance drops. The interpretation is that the actual number of independent information channels in an implant is probably not more than 4 to 6. Also, because higher stimulation rates can be achieved with fewer activated electrodes (Shannon, Adams, Ferrel, Palumbo & Grandgenett, 1990), the temporal characteristics of the signal are preserved better. Thus, evidence suggests that good spatial resolution is not achieved in cochlear implants, but also that preservation of the temporal waveform is important in cochlear implants.

Conversely, it has been shown in many pitch discrimination or electrode discrimination experiments (Nelson, Van Tasell, Schroder, Soli & Levine, 1995; Pfingst,

Holloway, Zwolan & Collins, 1999), where a fixed stimulation frequency was used on various electrodes, that cochlear implant users can discriminate between electrodes. Furthermore, pitch estimation experiments show that implant users can assign pitch to electrodes in a systematic fashion (Dorman, Smith, Smith & Parkin, 1994) which follows the tonotopical arrangement of the cochlea. Spikes are entrained to the stimulus in electrical stimulation (Javel, 1990), so if the phase-lock code was the only mechanism operating in frequency discrimination or pitch perception, stimuli on all electrodes would have had the same pitch. So electrode discrimination and pitch estimation experiments provide convincing arguments in favour of the rate-place code.

It is concluded that cochlear implants have not yet provided the final answers to the question of the coding of frequency in the auditory system.

IMPLICATIONS FOR COCHLEAR IMPLANTS

It is far easier to get high temporal resolution in electrical stimulation than it is to get high spectral resolution. Current spread from electrodes limit spectral resolution (Kral, Hartmann, Mortazavi & Klinke, 1998). New electrode designs may limit current spread (Cords, Reuter, Issing, Sommer, Kuzma & Lenarz, 2000), but certain physical limitations on electrode design remain. For example, maximum safe levels of charge density exist (Shannon, 1992). On the other hand, there are no basic technological limitations on increasing the stimulation rate. However, neural threshold adaptation may occur for high stimulation rates (above 400 pulses per second per channel), which suggests that higher stimulation rates may not be beneficial and may even degrade speech recognition performance (Javel & Shepherd, 2000). Still, the success of temporal pattern based strategies for cochlear implants like CIS is encouraging and warrants further study.

CONCLUSIONS

- (1) To be able to predict frequency difference limens for acoustic stimulation, an important assumption is that one spike per stimulus cycle is available, which may be provided by the existence of a volley principle. The volley principle may be implemented by the cochlear nucleus, where neurons have been found that can improve temporal precision by combination of a number of auditory nerve inputs (Moller, 1999).
- (2) An additional assumption is required in order to predict frequency difference limens for electrical stimulation of the auditory system. It is assumed that because many fibres fire on exactly the same phase of the electrical stimulation waveform, desynchronization results at a central auditory nervous system integration centre, which in turn leads to degradation in frequency discrimination.
- (3) Psychoacoustic data from cochlear implants show that both mechanisms for the coding of frequency information in the auditory system are equally likely. Thus, though cochlear implants may provide a tool to solve this problem, they have not yet provided the final answer to the question of coding of frequency in the auditory system.

REFERENCES

- Blamey, P.J., Dooley, G.J., Parisi, E.S., & Clark, G.M. (1996). Pitch comparisons of acoustically and electrically evoked auditory sensations. *Hearing Research*, 99, 139-150.
- Cook, E.P. & Johnston, D. (1999). Voltage-dependent properties of dendrites that eliminate location-dependent variability of synaptic input. *Journal of Neurophysiology*, 81, 535-543.
- Cords, S.M., Reuter, G., Issing, P.R., Sommer, A., Kuzma, J. & Lenarz, T. (2000). A silastic positioner for a modiolus-hugging position of intracochlear electrodes: electrophysiologic effects. *American Journal of Otolaryngology*, 21, 212-217.
- Dorman, M.F., Smith, M., Smith, L., & Parkin, J.L. (1994). The pitch of electrically presented sinusoids. *Journal of the Acoustical Society of America*, 95, 1677-1679.
- Dynes, S.B.C. & Delgutte, B. (1992). Phase-locking of auditory-nerve discharges to sinusoidal electric stimulation of the cochlea. *Hearing Research*, 58, 79-90.
- Eddins, D.A., Green, D.M. (1995). Temporal integration and temporal resolution. In B.C.J. Moore (Ed.), *Hearing*. San Diego: Academic Press.
- Fishman, K., Shannon, R.V., & Slattery, W.H. (1997). Speech recognition as a function of the number of electrodes used in the SPEAK cochlear implant speech processor. *Journal of Speech and Hearing Research*, 40, 1201-1215.
- Goldstein, J.L. & Sruлович, P. (1977). Auditory-nerve spike intervals as an adequate basis for aural frequency measurement. In E.F. Evans & J.P. Wilson (Eds.), *Psychophysics and physiology of hearing*. London: Academic Press.
- Hanekom, J.J. and Krüger, J.J. (2001). A model of frequency discrimination with optimal processing of auditory nerve spike intervals. *Hearing Research*, 151, 188-204.
- Hartmann, R., Topp, G., & Klinke, R. (1984). Electrical stimulation of the cat cochlea - discharge pattern of single auditory fibres. *Advances in Audiology*, 1, 18-29.
- Javel, E. (1990). Acoustic and electrical encoding of temporal information. In J.M. Miller and F.A. Spelman (Eds.), *Cochlear implants: models of the electrically stimulated ear*, New York: Springer-Verlag.
- Javel, E. & Mott, J.B. (1988). Physiological and psychophysical correlates of temporal processes in hearing. *Hearing Research*, 34, 275-294.
- Javel, E. & Shepherd, R.K. (2000). Electrical stimulation of the auditory nerve. III. Response initiation sites and temporal fine structure. *Hearing Research*, 140, 45-76.
- Johnson, D.H. (1980). The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones. *Journal of the Acoustical Society of America*, 68, 1115-1122.
- Johnstone, B.M., Patuzzi, R., & Yates, G.K. (1986). Basilar membrane measurements and the travelling wave. *Hearing Research*, 22, 147-153.
- Kalman, R.E. (1960). A new approach to linear filtering and prediction problems. *Transactions of the ASME - Journal of Basic Engineering*, 82, 35-45.
- Kim, D.O. & Parham, K. (1991). Auditory nerve spatial encoding of high-frequency pure tones: population response profiles derived from d' measure associated with nearby places along the cochlea. *Hearing Research*, 52, 167-180.
- Kral, A., Hartmann, R., Mortazavi, D., & Klinke, R. (1998). Spatial resolution of cochlear implants: the electrical field and excitation of auditory afferents. *Hearing Research*, 121, 11-28.
- Loizou, P.C. (1999). Signal-processing techniques for cochlear implants. *IEEE Engineering in Medicine and Biology Magazine*, 18, 34-46.
- Moller, A.R. (1999). Review of the roles of temporal and place coding of frequency in speech discrimination. *Acta Otolaryngologica (Stockholm)*, 119, 424-430.
- Moore, B.C.J. (1973). Frequency difference limens for short-duration tones. *Journal of the Acoustical Society of America*, 54, 610-619.
- Moore, B.C.J. & Glasberg, B.R. (1986). The role of frequency selectivity in the perception of loudness, pitch and time. In B.C.J. Moore (Ed.), *Frequency selectivity in hearing*. London: Academic Press.
- Nelson, D.A., Van Tasell, D.J., Schroder, A.C., Soli, S., & Levine, S. (1995). Electrode ranking of "place pitch" and speech recognition in electrical hearing. *Journal of the Acoustical Society of America*, 98, 1987-1999.
- O'Leary, S.J., Tong, Y.C., & Clark, G.M. (1995). Responses of dorsal cochlear nucleus single units to electrical pulse train stimulation of the auditory nerve with a cochlear implant electrode. *Journal of the Acoustical Society of America*, 97, 2378-2393.
- Pfingst, B.E. (1988). Comparisons of psychophysical and neurophysiological studies of cochlear implants. *Hearing Research*, 34, 243-252.
- Pfingst, B.E., Holloway, L.A., Zwolan, T.A., & Collins, L.M. (1999). Effects of stimulus level on electrode-place discrimination in human subjects with cochlear implants. *Hearing Research*, 134, 105-115.
- Rhode, W.S., Greenberg, S. (1992). Physiology of the cochlear nuclei. In A.N. Popper & R.R. Fay (Eds.), *The mammalian auditory pathway: neurophysiology*. New York: Springer-Verlag.
- Rose, J.E., Brugge, J.F., Anderson, D.J., & Hind, J.E. (1968). Patterns of activity in single auditory nerve fibres of the squirrel monkey. In A.V.S. de Reuck & J. Knight, *Hearing mechanisms in vertebrates*. London: J & A Churchill Ltd.
- Sachs, M.B. & Miller, M.I. (1985). Pitch coding in the auditory nerve: possible mechanisms of pitch sensation with cochlear implants. In R.A. Schindler & M.M. Merzenich (Eds.), *Cochlear Implants*. New York: Raven Press.
- Sek, A. & Moore, B.C.J. (1995). Frequency discrimination as a function of frequency, measured in several ways. *Journal of the Acoustical Society of America*, 97, 4, 2479-2486.
- Shannon, R.V. (1992). A model of safe levels for electrical stimulation. *IEEE Transactions on Biomedical Engineering*, 39, 424-426.
- Shannon, R.V., Adams, D.D., Ferrel, R.L., Palumbo, R.L., & Grandgenett, M. (1990). A computer interface for psychophysical and speech research with the Nucleus cochlear implant. *Journal of the Acoustical Society of America*, 87, 905-907.
- Shepherd, R.K. & Javel, E. (1999). Electrical stimulation of the auditory nerve: II. Effect of stimulus waveshape on single fibre response properties. *Hearing Research*, 130, 171-188.
- Siebert, W.M. (1970). Frequency discrimination in the auditory system: place or periodicity mechanisms? *Proceedings of the IEEE*, 58, 723-730.
- Sruлович, P. & Goldstein, J.L. (1983). A central spectrum model: a synthesis of auditory-nerve timing and place cues in monaural communication of frequency spectrum. *Journal of the Acoustical Society of America*, 73, 1266-1276.
- Townshend, B., Cotter, N., Van Compernelle, D., & White, R.L. (1987). Pitch perception by cochlear implant subjects. *Journal of the Acoustical Society of America*, 82, 106-115.
- Van den Honert, C. & Stypulkowski, P.H. (1987). Temporal response patterns of single auditory nerve fibers elicited by periodic electrical stimuli. *Hearing Research*, 29, 207-222.
- Wakefield, G.H. & Nelson, D.A. (1985). Extension of a temporal model of frequency discrimination: intensity effects in normal and hearing-impaired listeners. *Journal of the Acoustical Society of America*, 77, 613-619.
- Wever, E.G. (1949). *Theory of hearing*. New York: Wiley.