Proceedings of The Institute of Acoustics

A COMPUTER MODEL OF AUDITORY NEURAL FIRINGS EXHIBITING THE SECOND EFFECT OF PITCH SHIFT

A.J. BROOKES AND R.P. WILLIAMS.
DEPARTMENT OF PHYSICS, UNIVERSITY COLLEGE CARDIFF.

Introduction:

The pitch of a complex tone, having a series of harmonics but no fundamental present, is heard to be that of the missing fundamental. This is the 'residue pitch' phenomenon and has been researched in detail by Schouten in 1940. At this time, the pitch of the residue was ascribed to the period of the waveform envelope at a particular point on the basilar membrane. If, however, all of the harmonics are shifted by the same frequency increment then the pitch of the residue is shifted in proportion even though the envelope period of the waveform remains unchanged. This shift was studied by de Boer (1) and it was suggested that the residue pitch corresponded to the inverse of the time intervals between major peaks in the waveform the deviation from the harmonic fundamental pitch thus predicted first effect shift which was evident as an extra, superimposed shift now called the 'second effect of pitch shift'.

In 1966, Schroeder (2) demonstrated that the second effect of pitch shift could be produced by phase modulation of the carrier frequency of the complex tone, and suggested mechanical and neural origins for this. Fischler and Cern (3) suggested a model, based upon motions of the basilar membrane showing that such phase effects were present and could produce the second effect. The predicted magnitude, however, was small compared with that shown in recent psychoacoustical experiments by Greenhough et al. (4). Later models have indicated that the second effect could be due to the effects of combination tones generated within the inner ear.

The aim of the model proposed in this paper is to demonstrate that the second effect could arise from neural sources. It is a simple computer model and has the advantage that it may be reproduced several times to produce a more complex system. The output from the model is in the form of inter-firing time and post-stimulus time histograms and closely resembles similar histograms plotted by Rose (5). Predictions have been made with the model concerning the effects of phase and intensity variations of the components of complex tones.

The Model:

The model is a simulation of nerve fibres in a localised area of the basilar membrane. The assumption is made that the waveform at such a point is similar to that entering the ear if a narrow-band signal is used. Most of the tests involved the use of three-tone complexes for this reason. No attempt has been made to include neural interactions such as inhibition and all effects are studied over a large period of time (5 seconds) so that time averaging can be used to simulate a larger number of fibres.

When a stimulus is applied to the model, the parameters which determine whether or not a firing may occur are as follows:

(a) Refractory time. This represents the time interval after a nerve impulse during which the neuron recovers and cannot carry a further impulse. Two periods of refractory time exist: absolute refractory time during which, irrespective of stimulus intensity, a neuron will not produce a conducted response, and the relative refractory period during which a sufficiently intense stimulus may incite a response. For simplicity in the model, the relative refractory period is ignored and the firing probability is set to zero for an absolute refractory time of about lmsec after a firing has occurred.

(b) Control of firing probability by simulus amplitude. Period histograms, such as those of Rose (5) indicate that the firing probability is directly related to the amplitude of the positive half of the waveform at any time. The model uses a 'rectifier' law for firing probability: linear for the part of the waveform above a set amplitude threshold and zero probability for all points below the threshold.

(c) Temporal integration. The recent history of the nerve and associated part of the basilar membrane is monitored and is used to adjust the response time of the nerve. The second effect of pitch shift with complex tones is assumed to be due to the changing position of the major peaks within the waveform envelope. Temporal integration represents the activity upon the membrane for a short time (of the order of a few millisecs) before any particular time. In the model it controls a time delay in the neural firing, the greater the recent activity, the greater delay imposed upon a firing.

Results from the model for a pure tone of a given intensity were matched to those published in (5). These results were reproduced merely by adjusting the parameter (b). At this point the firing probability threshold was set at zero so that the positive half of the waveform controlled the firing probability. The output corresponding to a stimulus consisting of 800 Hz and 1200 Hz was also compared to similar histograms in (5). This test indicated that the model was less sensitive to differences in the height of major peaks than was a real neuron. To correct this, the zero level threshold was raised.

Finally, three tone complexes were input into the model which was now consistent with all the data in (5).

<u>Output from the model</u>:

The output was in the form of inter-firing time histograms. Figure 1 shows the output for a simple pure tone stimulus. The interval between the peaks is an indication of the frequency of the tone. Figure 2 shows the output for a harmonic three-tone stimulus. There is a cluster of peaks about the residue frequency of 200 Hz.

If all three tones are raised by a small frequency increment, the major peak in the histogram moves to indicate a new residue pitch which demonstrates the first and second effects. As the complex is shifted further, it becomes increasingly anharmonic and the relative heights of the peaks in the histograms change until the original major peak ceases to be the largest. This effect represents a change to a different set of effective harmonic numbers and is observed psychoacoustically. It is possible to isolate the second effect as shown by Greenhough et al (4) by taking the outer two tones of a three-tone complex and shifting them equally and oppositely about the middle tone. Figure 3 shows the effect upon the histograms when this is done. Again as the complex becomes increasingly anharmonic, the relative heights of the maxima change. The mean frequencies represented by the major peaks are shifted slightly consistent with the second effect. It may be noticed from figures 2, 3 & 4 that sets of peaks are clustered around multiples of the fundamental period, and when the mean pitches of these subharmonic peaks are calculated, they all show proportional shifts. The second effect of pitch shift shown by the model is controllable in magnitude over a small range by adjusting the time delay and behaves in a manner which is similar to that observed experimentally in psychoacoustics: there is no great change in its magnitude with a change in the modulation frequency (whilst retaining a constant carrier frequency) as is experimentally shown in (4).

A COMPUTER MODEL OF AUDITORY NEURAL FIRINGS EXHIBITING THE SECOND EFFECT OF PITCH SHIFT

Predictions of the model:

Since the structure of the waveform is a factor determining the delay in firing (via temporal integration), then it would be expected that any structural changes would cause changes in the magnitude of the second effect. For instance, phase effects, changes in intensity and modulation-index changes should all cause changes in its magnitude. Conflicting evidence (6,7) for such effects has been published so our psychoacoustical research was directed to such phenomena.

A waveform with 100% modulation would be expected to show a larger second effect than one with less modulation due to the larger variations in the time delay factor. Thus, since a change in the phase of the centre tone of a three tone complex changes the modulation index then one would expect a related change in the second effect. Ritsma in 1969 (6) indicated that large pitch changes occurred between the fully amplitde-modulated complex tone and the quasi-frequency modulated (centre-tone phase 90°) complex tone, but these effects may be attributed to a change in the prominence of certain peaks in the waveform. Wightman (7) produced results conflicting with those of Ritsma indicating no pitch change for the harmonic case. No results seem to exist for any variation in the second effect for anharmonic signals.

Similarily, as the overall intensity of the anharmonic complex is changed, a pitch change should be detectable. Predictions by the model for such effects are, at present, qualitative since the stimulus intensity versus firing rate laws have not been thoroughly investigated at this time. However, the model would predict a second effect increase with an increase in intensity.

Psychoacoustical Investigations:

Computer-generated complex tones were relayed via electrostatic headphones to listeners in a sound-proof booth.

1) Phase change experiments: the centre tone phase of the complex was varied in steps of 30 degrees from 0° to 360°. Tones with several modulation and carrier frequencies were chosen to investigate the effects on different parts of the basilar membrane. Pitch matching was accomplished with the use of a test tone of a similar timbre to that of the complex stimulus. Results are shown in figure 6 which should be compared with model predictions in figure 5.

2) Intensity variation experiments: the intensity of the stimuli were varied in 5dB steps from the listener's threshold of hearing up to a point where distortion products impaired measurements. Again, several modulation and carrier frequencies were used. Figure 7 shows the results from this part of the experiment, four different plots are shown for different component shifts.

Discussion:

Although the model is simple in terms of parameters, its output and predictions matched psychoacoustical and neurological results closely. An interesting result was that the minima for the second effect versus phase graphs rarely fell at 90° but usually occurred around 120°. Observers also commented that this was the point at which the tone was the most difficult to match. This result may be compared with the results of Buunen & Bilsen (8).

References:

1. de Boer E. (1956) Doctoral dissertation.University of Amsterdam.
2. Schroeder (1966) Journal of the Acoustical Society of America 40, 79-81.
3. Fischler & Cern (1968) J.Acoust.Soc.Amer. 44, 1379-1385.
4. Greenhough, Williams & Wright (1973) J.Sound Vib. 26, 277-282.
5. Rose J.E. (1969) in Frequency Analysis and Periodicity detection in hearing. Ed Plomp & Smoorenburg (1970).

A COMPUTER MODEL OF AUDITORY NEURAL FIRINGS EXHIBITING THE SECOND EFFECT OF PITCH SHIFT.

6. Ritsma R.J. (1969) in Frequency Analysis and Periodicity detection in Hearing (1970).
7. Wightman F.L. (1973) J.Acoust.Soc.Amer. 54, 397-406.
8. Buunen & Bilsen (1974) in 'Facts and Models in Hearing' Ed Zwicker & Terhardt.