FIFTH INTERNATIONAL CONGRESS ON SOUND AND VIBRATION
DECEMBER 15-18, 1997
ADELAIDE, SOUTH AUSTRALIA

USE OF AN AUDITORY MODEL TO EXPLAIN THE MECHANISM OF PITCH DISCRIMINATION

Suhashini Shankar and Suzanne Hogg
University of Technology, Sydney, Australia

ABSTRACT

The number and nature of the mechanisms of pitch discrimination in the human ear is not clearly understood. To deepen our understanding of the mechanism involved in perceiving pitch, we have used the auditory model of Meddis and Hewitt (1991). We optimised the parameters and measured the jnd of the model for sinusoidal stimulus. Comparison of our experimental findings on the mechanism of pitch discrimination, reported earlier, and the performance of the model indicates the single mechanism operating in discriminating pitch.

INTRODUCTION

The mechanism of pitch discrimination remains a mystery for over a hundred and fifty years. After many years of research, there is overall agreement among scientists as to the preliminary mechanical analysis of frequency at the basilar membrane in the cochlea and a subsequent neural analysis. But our knowledge about the central neural pitch extraction mechanism is very limited. It is not yet clear whether the neural analysis is based on spatial and/or temporal pattern of nerve activity. Nordmark (1968) argued the possible existence of two different mechanisms operating in pitch determination. From his plot of least discriminable change in time interval against frequency he reasoned that rate mechanism (analysis in time domain, temporal or periodicity) operates at low frequency and place mechanism (analysis in frequency domain, spatial) at high frequencies. Pierce J (1991) went a step further determining the critical rate below which periodicity mechanism can be analysed in isolation of the place mechanism. Shankar et al. (1996) repeated Nordmark's experiment with digitally generated stimuli and more controlled experimental conditions and obtained similar result of that of Nordmark. Houtsma and Smurzynski (1990) carried out four experiments in pitch identification and discrimination of complex tones. Their qualitative analysis was in accord with both single and dual mechanisms. To broaden our understanding of the single/dual mechanisms of pitch discrimination we have used the auditory model of Meddis and Hewitt(1991). This study consists of optimising the parameters and finding the jnd of the model for a range of pure tones. The performance of the model was compared to our experimental results and we discuss the possible mechanism that operates in our auditory system.

There are many computational models of the cochlea developed over the past decade. Hermes (1993) gives a detailed account of the nature of different PDAs (Pitch Determining Algorithms) for different applications. However, these models have not been very useful in studying the mechanism of pitch discrimination, because, for these applications knowledge about the mechanism of pitch extraction is not necessary. The model of Meddis & Hewitt (1991) stands out of all the third generation models that, according to Hermes, are based on the theories of pitch perception and the knowledge of auditory periphery. This model has the ability to predict the results of the very important pitch perception experiments such as missing fundamental, ambiguous pitch, pitch shift of equally spaced harmonics, musical cords, repetition pitch etc. Hence we chose their model for our study of the mechanism of place and / or rate pitch.

Briefly, the model uses several stages to mimic the effects of sound transmission along the human ear. A band pass filter simulates the effects of outer/middle ear. This is followed by a bank of band pass filters that models the initial mechanical analysis along the basilar membrane. Mechanical to neural transduction is done by Meddis's (1986) Inner Hair Cell model. Common periodicity in the outputs of the filters is extracted by an autocorrelation function(ACF). Summary autocorrelation (SACF) aggregates the periodicity information to decide the pitch. Only the periodicity mechanism operates following the initial analysis along the basilar membrane of the cochlea. The agreement/disagreement of the response of the model to our results of pitch matching experiments by human listeners will throw more light on our understanding of the mechanisms of pitch discrimination.

MEASURE OF JND

In our pitch matching experiments with human listeners, subjects matched the pitch of the test tone with that of the reference tone by the method of adjustment. Reference and test tones of duration 800 ms, separated by a period of silence of 800 ms were presented at 45 dB.

The difference in frequency (df) of the test and the reference tones when pitches were matched was measured at each stimulus frequency(f). We expressed jnd's as the least discriminable change in frequency and least discriminable change in time interval when comparing our results with that of Nordmark's. df/f was used as a measure of jnd when our results and those of Shower & Buddulph (1931) and Harris (1952) were compared.

We adopted a similar method of measuring the jnd with Meddis & Hewitt's model. Applying the model with the pure tone stimulus of a given frequency we determined the pitch as indicated by SACF. We increased and decreased the frequency of the stimulus by 1 Hz until the pitch changes (ie) until the first peak of SACF changes. For example, stimulus frequency of 980 Hz to 1034 Hz gave the peak of SACF at 1000 Hz. The difference between 1000 Hz and 980 Hz is referred to 'df - lower' and the difference between 1034 Hz and 1000 Hz as 'df - higher'. The following table summarises this calculation.

Stimulus frequency (Hz)	df - lower (Hz)	df - higher (Hz)	mean df (Hz)
1000	(1000-980) 20	(1034 - 1000) 34	27

Table 1: Calculation of jnd

The mean of df - lower and df - higher gives us the jnd 'df' for 1000 Hz. The values of dT (least discriminable change in time interval) or df/f can be calculated when we compare the model output with our results or with that of Nordmark or Shower and Buddulph and Harris.

CHOICE OF PARAMETERS

Ramped pure tone stimulus of duration 30 ms and intensity (SPL) 50 dB was allowed to pass through the following modules.

stimulus → (BM Filter Bank) → (IHC Meddis 86) → (autocorrelation) → (SACF) → pitch

The model provides three modes, linear, logarithmic and ERB for the centre frequency of the basilar membrane filter bank. We used ERB mode for both the centre frequency and bandwidth for the filters that simulate the mechanical analysis along the basilar membrane. Mechanical to neural transduction is modelled by Meddis Inner Hair Cell IHC - 86. The performance of the model was tested at two different sampling frequencies 20,000 Hz and 50,000 Hz and ERB densities 0.5 (16 channels) and 1.0 (31 channels). Our stimulus did not pass through pre-emphasis band pass filter that simulates the effects of outer/middle ear because the bandwidth of this filter can be taken broader than the frequency range that we were testing. The last 10 ms duration of the stimulus was used in calculating the ACF. The list of the parameters that were used is given in the appendix.

RESULTS

The following tables summarise the result.

Freq (Hz)	df higher (Hz)	df lower (Hz)	mean df (Hz)
250	0	4	2
500	9	0	4.5
1000	34	20	27
2000	92	96	94

Table 2: 20,000 Hz sampling frequency and 16 channels

Freq (Hz)	df higher (Hz)	df lower (Hz)	mean df (Hz)
250	0	2	1
500	0	5	2.5
1000	19	16	17.5
2000	34	30	32

Table 3: 50,000 Hz sampling frequency 16 channels

Freq (Hz)	df higher (Hz)	df lower (Hz)	mean df (Hz)
125	0.55	0	0.275
200	5	0	2.5
250	3	0	1.5
500	8	9	8.5
1000	26	28	27
2000	99	100	99.5

Table 4: 20,000 Hz sampling frequency 31 channels

We then calculated the least discriminable change in frequency and in time interval.

Freq (Hz)	dT, μs 20000,16	dT, μs 50000,16	dT, μs 20000,31
250	32	16	24
500	18	10	34
1000	27	17.5	27
2000	23.5	8	24.875

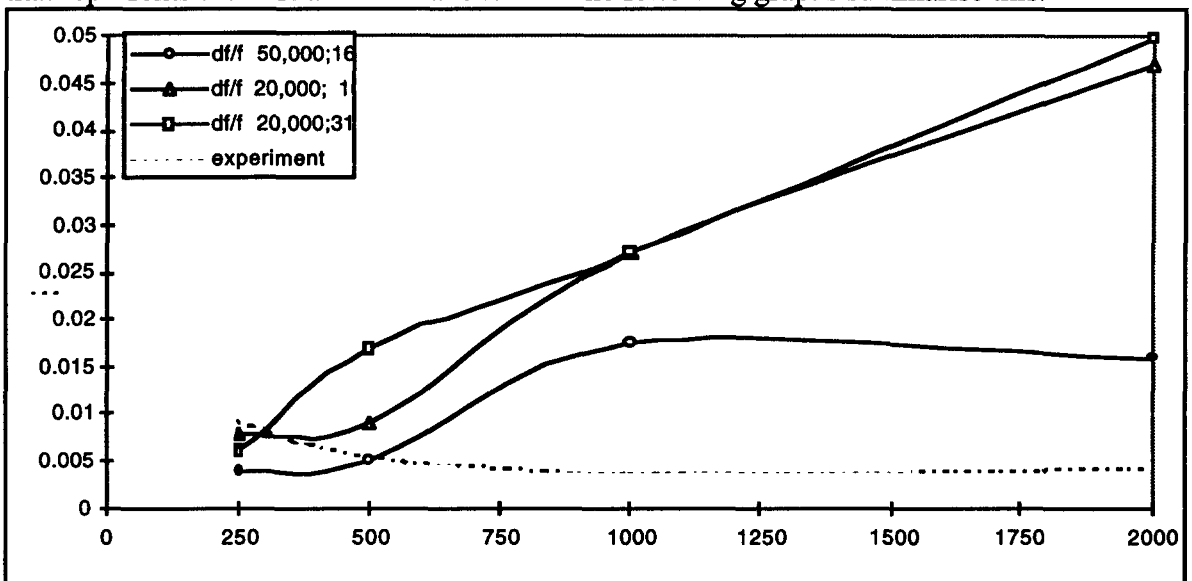
Table 5: Comparison of dT(μs) at different sampling frequency, and number of channels for different stimulus frequency

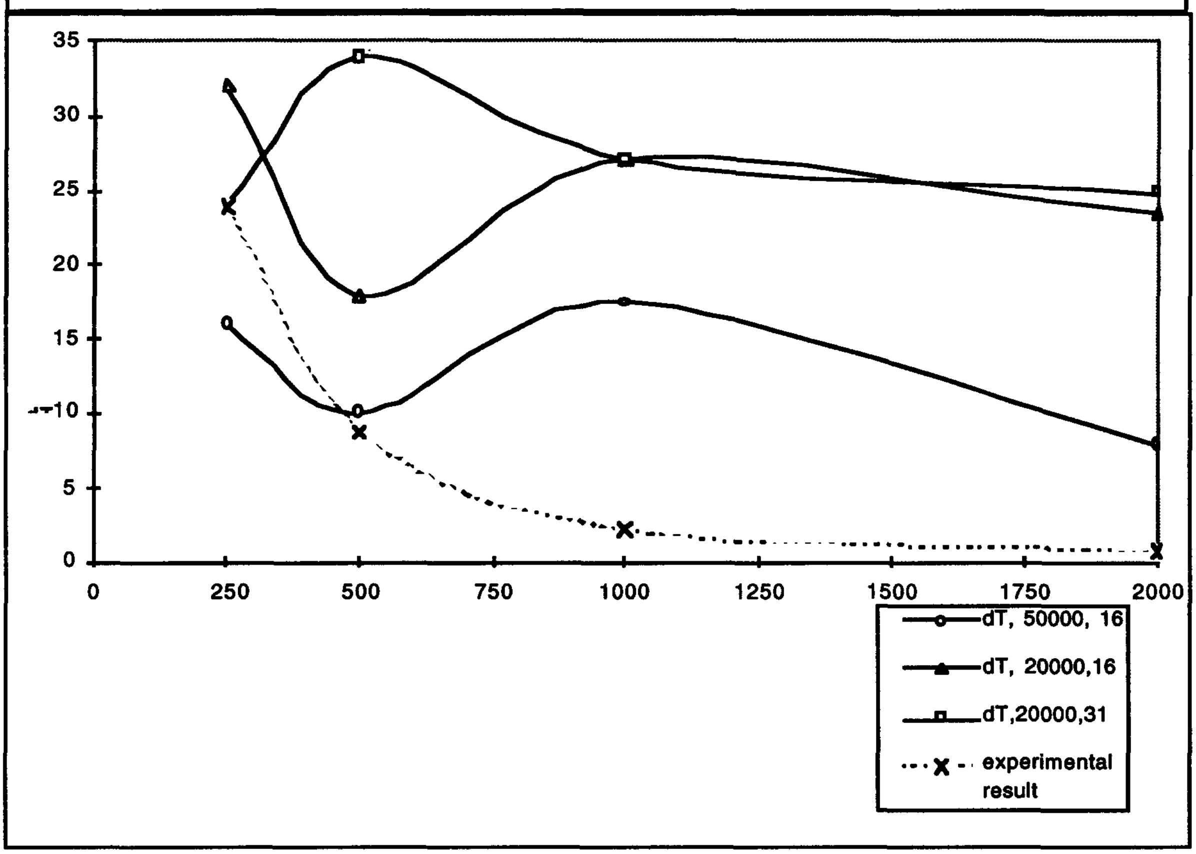
Freq (Hz)	df/f 20000,16	df/f 50000,16	df/f 20000,31
250	32	16	24
500	18	10	34
1000	27	17.5	27
2000	23.5	8	24.875

Table 6: Comparison of df/f at different sampling frequency, and number of channels for different stimulus frequency

ANALYSIS

To compare the effects of the sampling frequencies and the number of channels on pitch discrimination we calculated the jnd in time (dT) and df/f. Our plot of dT vs f and df/f vs f show that the performance of the model agrees better with the experimental results when the sampling frequency is equal to 50,000 Hz and the number of channels along the filter bank that represents the basilar membrane is 16. The following graphs summarise this.





We chose this set of data from the above to compare to the results of our experiments. The comparison of our results with other researchers are discussed in Shankar et al (1996).

Freq (Hz)		**dT (micro seconds)**	
	experimental results	Nordmark	model (50,000;16)
125		14.2	
250	23.84444	5.4	16
500	8.670386	2.3	10
1000	2.164532	0.85	17.5
1500	1.237743		
2000	0.817127	0.38	8
3000	0.596062		
4000	0.783316	0.25	
8000		0.25	
12000		0.2	

Table 7: Comparison of the model performance to our experimental results and Nordmark's results

	df/f			
Freq (Hz)	Shower & Buddulph	Harris	experimental results	model 50000,16
62	0.0426			
125	0.0247	0.0034	0.0091	
250	0.0103	0.0031	0.0054	0.0040
500	0.0052	0.0022	0.0038	0.0050
1000	0.0036	0.0013	0.0045	0.0175
1500			0.0042	
2000	0.0019	0.0010	0.0042	0.0160
3000			0.0051	
4000	0.0023	0.0025		
8000	0.0029			
11700	0.0035			

Table 8: Comparison of the model performance to our experimental results and others

DISCUSSION

Jnd representation is one way of understanding the place / rate mechanism. Nordmark(1968) has discussed this in length. If place mechanism is the operant in determining pitch, jnd in frequency plotted against frequency (df vs f) is the correct representation. If pitch is determined by periodicity mechanism, plot of jnd in time interval (dT vs f) against frequency is a better choice. On this basis, Nordmark's plot of least discrminable change in time interval against frequency gave two distinct regions suggesting the possibility of two mechanisms operating in pitch determination. From his graph Nordmark concluded that for sinusoids at low frequencies, frequencies less than 2000 Hz, pitch is determined by time measuring process and at high frequencies it is by frequency measuring process. When we repeated

Nordmark's experiment with digitally generated stimuli, the two distinct regions in our plot of jnd in time interval against frequency were apparent.

The results of the model behaviour when pure tone stimuli were applied indicate a different picture. The model attempts to detect a common periodicity in the outputs of the filters. Only the periodicity mechanism is operant in the frequency range that we chose, 250 Hz to 2000 Hz. The plot of dT vs f for the model agrees almost with the experimental results. This suggests that periodicity is the one and only mechanism that determines pitch at all frequencies. The results of the model agrees reasonably to the results of Shower & Buddulph and Harris. Meddis (in private communication) indicates a unitary model of pitch perception from different perspective, from the analysis of complex tones.

Nordmark in his paper (1968) also discussed the distinction between group and phase frequency. Cochlea being a filter type analyser measures group frequency and this is dependent on the process by which it is measured. Time measuring process employed in the model (ie) periodicity detection is consistent with the definition of group frequency.

We are working on measuring jnd for the model using rectangular pulses. Behaviour of the model in comparison with our experimental results, Nordmark's and Flanagan and Guttman will give a better and more concrete picture of the mechanism of pitch perception.

REFERENCES

HARRIS, J.D(1952). "Pitch discrimination", *J.Acous.Soc.Am* **24**: 750-755

HERMES J D., " Pitch Analysis" in *Visual representations of speech signals* edited by Coor M, Beet S and Crawford M, 1993

HOUTSMA,A.J.M., and SMURZYNSKI,J(1990). "Pitch identification and discrimination for complex tones with many harmonics", *J.Acous.Soc.Am.* **87**:304

MEDDIS R and HEWITT J M(1991). "Virtual pitch and phase sensitivity of a computer model of the auditory periphery 1: Pitch identification." *J. Acoust. Soc. Am* **89** 2866 - 2882

MEDDIS R and HEWITT J M(1991). "Virtual pitch and phase sensitivity of a computer model of the auditory periphery 2: Phase sensitivity." *J. Acoust. Soc. Am* **89** 2883 - 2894

NORDMARK O J (1968) "Mechanisms of frequency discrimination" *J. Accous.Soc.Am* **44** 1533-1540

PIERCE, R. J. (1991). "Periodicity and pitch perception" *J. Acoust. Soc. Am* **90** 1889 - 1893

SHANKAR S, HOGG S, DERMODY P, SEYMOUR J. (1996) "Mechanism of pitch measurement as implied by discrimination studies" *Fourth international congress on sound and vibration*, Russia, 1996

SHOWER,E.G and BUDDULPH R (1931), "Differential pitch sensitivity of the ear", *J.Acous.Soc.Am* **3**,275-287

APPENDIX

Module specifications.

Par. file	Name	Description
PTone1.par	PureTone	stimulus generation paradigm.
PreEmph1.par	null	outer-/middler-ear filter model.
GammaT3.par	BasilarM_GammaT	basilar membrane filter model.
RPModuleIV.par	null	IHC receptor potential model.
Meddis86.par	IHC_Meddis86	inner hair cell (IHC) model.
SpikeGen.par	null	Auditory nerve spike generation
Dendrite1.par file.	null	Dendrite low pass filter parameter
MGFusiform.par	null	Neural cell model.

Miscellaneous parameters

2.50E-03	Ramp up rise time for signal (s).
1.00E-03	Bin width for the post stimulus time histogram (s).
1.00E-02	Period for autocorrelation functions.
2.50E-03	Time constant for exponential decay component.

Stimulus - pure tone

124	Stimulus frequency (Hz).
50	Stimulus signal intensity (dB SPL).
3.00E-02	Stimulus signal duration (seconds).
5.00E-05	Stimulus sampling interval, dt (seconds).

Gamma3T.basilar membrane filter module

4	Order of the gamma tone basilar membrane model
ERB	centre frequency mode for filter
80	Initial centre frequency of gamma tone filter bank
8000	final centre frequwency of gamma tone filter bank
0.5	erb density for centre frequency list
ERB	Band width (3 dB down) mode for filters.

IHC-Meddis 86 module

100	Permeability constant A (units per second).
6000	Permeability constant B (units per second).
2000	Release rate (units per second).
5.05	Replenishment rate (units per second).
2500	Loss rate (units per second).
66.31	Reprocessing rate (units per second).
6580	Recovery rate (units per second).
1	Max. no. of transmitter packets in free pool.
50000	Firing rate (spikes per second).