A model of temporal response properties in primary auditory cortex

Susan L. Denham

Centre for Neural and Adaptive Systems, University of Plymouth,
Drake Circus, Plymouth PL4 8AA, UK

Abstract

The temporal response properties of cells in primary auditory cortex differ markedly from those observed sub-cortically, in particular the ability to synchronise firing to amplitude modulations is restricted to modulations below 10-20 Hz (Schreiner et al 1997). In the thalamocortical transformation of incoming signals a great deal of the temporal fine structure is lost (Creutzfeldt et al 1980), and the effects of a masker on a probe tone can be detected up to 400 ms after masker offset (Brosch and Schreiner 1997). What gives rise to these phenomena and can they be explained by some common mechanism? Explanations in terms of intracortical inhibitory circuits have been proposed but inhibition does not provide an adequate account, at least in the case of forward masking which is unaffected by the application of a GABA antagonist. On the other hand, simple threshold neural models cannot replicate such behaviour without some form of inhibition. The purpose of this investigation was to explore to what extent synaptic depression at thalamocortical synapses could account for the temporal response properties observed in primary auditory cortex. A model of synaptic depression is described and its behaviour investigated in a number of experiments. The model is shown to successfully replicate many experimental observations. In addition, the model also provides a novel account of the effect of subthreshold stimuli.

1. Introduction

Despite considerable experimental investigation, both psychophysical and physiological, there is as yet no clear understanding of how sounds are represented and recognised within the brain. In order to understand how complex sounds are encoded within cortex, and ultimately perceived, it is necessary to take into account the spectrotemporal nature of sounds. While the spectral analysis of sounds is quite well understood and the spectral organisation within most of the auditory system has been clearly documented, the way in which the temporal aspects of sound are encoded and used is less clear. In moving from the auditory periphery to the cortex, the typical time constants associated with processing tend to become progressively longer, and the upper limit on synchronisation to amplitude modulated sounds has been shown to fall in primary auditory cortex to roughly one hundredth of that observed in the auditory nerve (Schreiner *et al* 1997). In line with these physiological findings, there are also many aspects of auditory perception, such as the growth of loudness with duration and the effects of masking, which indicate that the auditory system performs some sort of temporal integration in processing incoming acoustic signals. However, the auditory system is also capable of fine temporal resolution, as evidenced by very short gap detection, double click discrimination, and also in the short latency and lack of jitter of onset responses in cortex (Viemeister and Wakefield 1991). This has been termed the resolution-integration paradox, i.e. how is it possible for a system to integrate information over long periods while retaining fine temporal resolution.

A simple linear model is unable to account simultaneously for all of these properties, since if the time constants are chosen to be long enough to account for the observed characteristics of temporal integration, they will be too long to allow the temporal precision also evident in cortical responses. A model that overcomes this problem is the 'multiple looks' model proposed by

Viemeister and Wakefield (1991). In this account it is suggested that a number of short samples of the stimulus are made, and their results combined by means of a weighting function. The model successfully explains much of the relevant psychophysical data, but does not attempt to identify the underlying physiological basis for the behaviour.

The hypothesis explored in this paper is that by incorporating synaptic dynamics within a neural model all of the phenomena outlined above can be explained. It has been shown that the dynamical properties of cortical synapses can significantly influence the temporal sensitivity of cortical circuitry (Abbott *et al* 1997). The key idea is that when synapses are repeatedly activated they do not simply respond in the same way to each incoming impulse and may develop a short-term depression or facilitation, depending on the nature of the pre- and postsynaptic cells, and on the characteristics of the particular synapse involved. In particular, thalamocortical synapses appear to be depressing (Thomson and Deuchars 1994); they are mediated by non-NMDA excitatory amino acids, depress rapidly and remain desensitised for some time. The suggestion that synaptic depression may explain the temporal characteristics of responses in cortex is not new and has been described in a number of other papers (Eggermont 1999, Chance *et al* 1998). For example, in the auditory system the upper limit and phase relationships of the modulation transfer function are well replicated by a model of synaptic depression (Eggermont 1999), and in the visual system, similar frequency response properties of cells in V1 were modelled by incorporating synaptic dynamics (Chance *et al* 1998). In both of these accounts, the behaviour of the model was expressed in terms of the firing rate of the postsynaptic cell. However, in order to explain some of the experimental results below, it was necessary to take into account the time course of the depletion and recovery of individual synapses and for this reason a model that explicitly represented the synaptic dynamics was used.

Such a model was presented by Tsodyks *et al* (1998, 1997) and was shown to replicate experimental results on the activity-dependent redistribution of synaptic efficacy (Markram and Tsodyks 1996). In fact, this model of the postulated dynamics of neurotransmitter release had already been proposed much earlier (Grossberg 1968, 1969), where it was derived from a set of psychological postulates and used to explain the excitatory transients in transmitter release after a rest period and related to the effects of synaptic depression, which had been observed experimentally by Eccles (1964). In the area of auditory modelling, a very similar model was also developed by Meddis (1986) to describe transduction in cochlear inner hair cells.

The remainder of the paper is organised as follows. Firstly the model is described and its behaviour explained. Then the model is then shown to replicate aspects of temporal processing observed in auditory cortex across a range of different experiments: including the loss of temporal fine structure in the thalamocortical transformation of incoming signals, the upper limit on synchronisation to modulation frequencies, the spectrotemporal course of forward masking, and the sensitivity to masker duration, the relationship between onset latency and the onset envelope of the stimulus, and the sensitivity of cells in primary auditory cortex to temporal manipulations of speech signals. In addition the model is shown to provide a possible explanation for the effect of subthreshold stimuli. Finally we discuss the implications of the model for auditory perception.

2. The dynamic synapse model

The model described here is based on that of Meddis (1986). The dynamic synapse model characterises the synapse by defining a "resource", e.g. the amount of neurotransmitter in the synapse, a proportion of which can be in one of three states: *available, effective, inactive.* The dynamical behaviour of the proportions of the resource that are in each of these states is determined by a system of three coupled differential equations:

$$\frac{dx}{dt} = z(t) - a.x(t) - l.x(t) \quad (1)$$

$$\frac{dy}{dt} = b.w(t) - z(t) + ?[K - x(t) - y(t)] \quad (2)$$

$$\frac{dw}{dt} = a.x(t) - b.w(t) \quad (3)$$

where

$$z(t) = g.K.I(t).f[y(t)] \quad (4)$$

$x(t)$ is the amount of *effective* resource, and could be interpreted as the activated neurotransmitter within the synaptic cleft as a proportion of the total resource; $y(t)$ is the amount of *available* resource or free neurotransmitter in the synapse, and $w(t)$ is the amount of *inactive* resource, neurotransmitter being reprocessed. In addition, as in Meddis' model there is assumed to be a slow leakage of effective resource from the synaptic cleft the rate of which is determined by l, and manufacturing of replacement resource at a rate determined by r. The constant b determines the rate at which the inactive resource $w(t)$ is released to the pool of available resource on a continuing basis, and a represents the rate at which effective resource becomes rapidly inactive again subsequent to being activated.

The input signal $I(t)$ represents the occurrence of a presynaptic action potential and is set equal to one at the time of arrival of the presynaptic action potential and for a small period of time dt thereafter, and otherwise is set equal to 0. The instantaneous efficacy of the synapse is determined by the variable g, which can be interpreted as the fraction of available resource released as a result of the occurrence of the presynaptic action potential. It takes a value in the range zero to one.

The key idea behind the model is that there is a fixed amount K of total resource available at the synapse, a proportion of which is activated in response to presynaptic activity, rapidly becomes inactive, and is then subsequently made available again through reprocessing. Thus, if the synapse is very active, i.e. it is bombarded by a large number of action potentials occurring over a short period of time, the amount of available resource $y(t)$ is rapidly reduced. There must then follow a period during which the synapse can recover in order to respond fully once more. This process appears to replicate the experimentally observed characteristics of synaptic depression, for example as reported in (Markram *et al* 1997, Markram and Tsodyks 1996).

The EPSP at the synapse, $e(t)$, is computed from $x(t)$ in (1) using the following equation for the passive membrane mechanism [26]:

$$t_{EPSP}\frac{de}{dt} = g.x(t) - e(t) \quad (5)$$

As has been well documented, synaptic transmission is a stochastic process, postsynaptic EPSPs vary in amplitude and there is a relatively high probability of failure in transmission at depressing synapses (Thomson and Deuchars 1995). In Meddis' model stochastic transmission was simulated by allowing the probability of a postsynaptic EPSP to be a function of $x(t)$, the amount of effective transmitter. However, this can result in the synapse becoming completely depleted, and is not consistent with experimental data showing that the probability of failure is inversely related to the failure of the previous presynaptic spike to elicit an EPSP (Galarreta and Hestrin 1998). Here we simply model both the probability of successful transmission and the amount of transmitter actually released as a probabilistic function of the transmitter available for release, $f[y(t)]$. While this is a simple phenomenological model that does not attempt any detailed account of the underlying biophysical processes, it does approximate the observed behaviour and has the advantage of being relatively simple to understand. The model has been tuned to replicate data on rapid and longer term depletion (Galarreta and Hestrin 1998), time for recovery (Galarreta and Hestrin 1998), the relationship between probability of failure and presynaptic spike train frequency (Galarreta and Hestrin 1998), and the relationship between the frequency of the presynaptic spike train and the average EPSP amplitude (Markram and Tsodyks 1996). The model parameters, used throughout the simulations are shown in table 1.

3. The neuron model

The neuron model is described by the following system of equations, which has been adapted from a model described in (McGregor 1989):

$$t_E \frac{dE}{dt} = -E(t) + V(t) + G_K(t).[E_K - E(t)] \quad (6)$$

$$s(t) = 1, \text{if } E(t) \geq q(t), \text{else } s(t) = 0 \quad (7)$$

$$t_{G_K} \frac{dG_K}{dt} = -G_K(t) + h.s(t) \quad (8)$$

$$t_q \frac{dq}{dt} = -(q(t) - q_0) + s(t) \quad (9)$$

where, *E(t)* is the variation of the neuron's membrane potential relative to its resting potential, *V(t)* is the driving input found by summing all the synaptic EPSPs, $G_K(t)$ is the potassium conductance, divided by the sum of all the voltage-dependent ionic membrane conductances, E_K is the potassium equilibrium potential of the membrane relative to the membrane resting potential, *q(t)* is the firing threshold potential, q_0 is the resting threshold, *s(t)* is the variable which denotes firing of the cell, t_E, t_{EPSP}, t_q, and t_{GK} are time constants, and *g, c* and *h* are constant parameters.

In this system of equations, *s(t)* is set to 1 to signal the occurrence of an action potential, i.e. *E(t)* reaching a value above the firing threshold *q(t)*; otherwise *s(t)* is zero. Equation (9) is introduced purely to provide a refractory period. It allows representation of an absolute period and a relative period. For the first few milliseconds after firing the value of *q(t)* becomes very large, preventing any further firing. As *q(t)* decays between spikes, the threshold for firing decreases with time elapsed since the last spike. A further spike can occur therefore in this period if the value of *E(t)* is sufficiently large. When *s(t)* is zero, the potassium conductance term $G_K(t)$ decays to zero via equation (8). When *s(t)* =1, the value of G_K is increased instantaneously by an amount *h*, and then decays again. We have not explicitly modelled the action potentials generated when the cell fires, but in the simulations below generally use the spiking variable *s(t)* as the output from the model.

4. Simulation results

The response of the model was investigated in a range of experiments, as described below. However, for most of the experiments simulated, the thalamocortical signals were not recorded. In these cases the details of the acoustic stimuli used in the experiments were documented and it was decided to simulate the experiments using similar acoustic stimuli. For this reason a well-documented and tested peripheral model, DSAM (O'Mard *et al*), was used to generate signals characteristically found in auditory nerve fibre recordings in response to acoustic stimuli. The output from the peripheral model was reprocessed to ensure that the firing rate remains below about 200 Hz by enforcing a reasonable refractory period. Clearly this ignores the computations which occur in the rest of the sub-cortical auditory system. However, there does appear to be a fast veridical auditory pathway that transmits signals from the periphery to the cortex with relatively little alteration (Helfert *et al* 1991). While recognising that this simplification may result in a poor approximation of actual thalamic relay cell activity, it has the benefit of making the simulations tractable.

In the simulations, unless otherwise stated, the acoustic signals specified are processed by the DSAM peripheral model that includes an outer-middle ear transfer function, a gammatone filterbank, and an inner hair cell model. A simple stochastic spike generator model is used and a convergence of 20 inner hair cells to 1 auditory nerve fibre assumed. The spike trains are then processed to ensure that refractory periods are generally greater than 20 ms. However, when more than one spike occurs simultaneously, as is possible with a combinations of 20 spike trains per channel, then the refractory period is allowed to decrease in proportion to the extent of the coincidence. This has the benefit of not destroying the enhanced onset response generated by the inner hair cells.

4.1 Loss of temporal fine structure in the thalamocortical transformation of incoming signals

Differences between the response properties of thalamic and cortical neurons were investigated by Creutzfeldt *et al* (1980). Activity in thalamic relay cells and subsequent activity in paired pyramidal cells in primary auditory cortex (AI) was recorded, and it was found that even when thalamic activity was clearly synchronized to the stimulus up to 200 Hz, the paired cortical cell was unable to follow the details of the signal beyond about 20 Hz. The plots in figure 1 show the response of the model to spike trains generated to resemble typical thalamic activity in response to stimuli of the frequencies indicated. PSTHs of the total activity for 20 presentations are plotted both for the presynaptic spike trains and the model response. The model behaviour closely resembles that found experimentally. The model responds to details of the stimuli occurring at 10 Hz and to a lesser extent to details at 20 Hz, but for higher stimulus frequencies, the model only responds strongly at the onset of the signal. The reason for this is that at high frequencies successive presynaptic spikes arrive before the synapse has time to recover. This causes a strong depression of the synapse, resulting in the generation of very small postsynaptic EPSPs that are insufficient to raise the cell membrane potential above the firing threshold, as illustrated in figure 2.

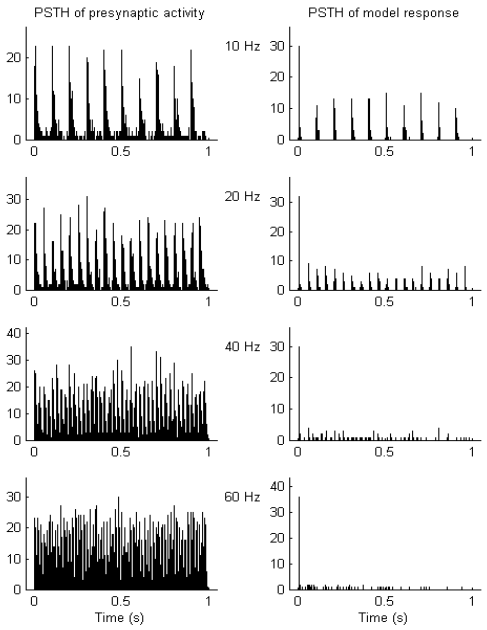

Figure 1. Simulation of the transmission of signals between thalamic relay and cortical pyramidal cells. PSTHs of input activity, left column, and model response, right column, for 20 repetitions, using a 2ms bin size. Spikes were generated probabilistically, resulting in the distributions labelled as 'presynaptic activity', and used as inputs to the model. This input activity resembles the activity in thalamic relay cells recorded experimentally in response to signals with periodicity indicated (Creutzfeldt *et al* 1980). The model qualitatively replicates the behaviour of paired pyramidal cells in AI, which showed almost no response except at signal onset when stimuli exceeded 20 Hz.

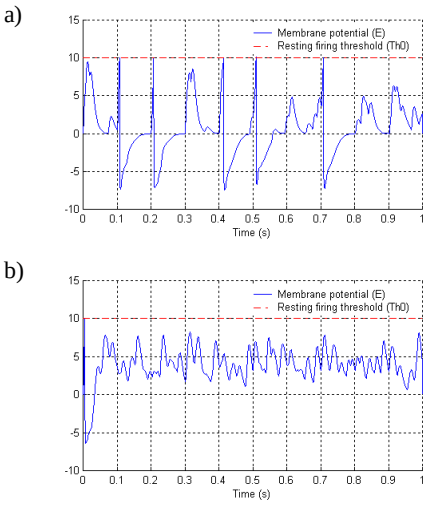

Figure 2. Example of the response of the model to a) 10 Hz and b) 40 Hz spike trains, showing how more rapid presynaptic spike trains prevent the recovery of the synapses and give rise to rather small EPSPs which are not sufficient to raise the membrane potential above the firing threshold.

4.2 Frequency response of the extended neuron model

The frequency response of the neuron model with a depressing synapse is illustrated in figure 3b. Although the synaptic dynamics were tuned to match those found experimentally in the somatosensory cortex (Markram and Tsodyks 1996), it is interesting to note that the model clearly responds preferentially to frequencies under 10 Hz, as is also found in AI (Schreiner and Urbas 1988). It seems to be the case that the dynamics of cortical depressing synapses may be quite similar across different cortical areas.

For comparison the response of a neuron model without a depressing synapse is shown in figure 3a. Clearly such a model cannot replicate the behaviour observed experimentally without the addition of delayed inhibitory inputs which increase in strength with stimulus frequency. Alternatively, modeling the synapse as a low pass filter but with very low cut-off frequency could result in a similar frequency response, but would fail simultaneously to account for the short response latencies found in AI (Heil 1997). The benefit of the proposed model is that it can account both for the low pass frequency response and short onset latency (see section 4.7) within a single neuron model.

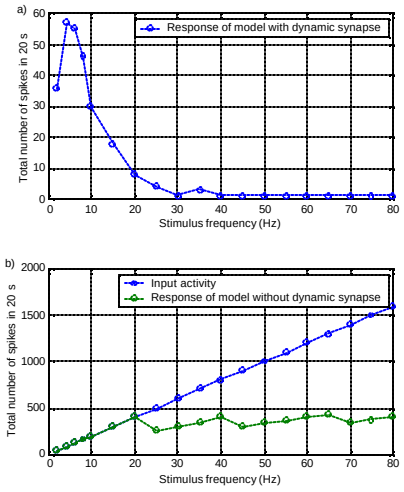

Figure 3. Frequency response of the model. The response of the neuron model a) with and b) without a depressing synapse to a spike train of the frequency indicated, simulated for 20 seconds. The plots show the total number of times the cell fired during the 20-second period. For higher frequencies there was generally only a response to the first spike in the train.

4.3 Best modulation frequencies

Rate modulation transfer functions were extensively investigated by Schreiner and Urbas (1988), who found that the best modulation frequencies in AI were generally below 15 Hz. This contrasts with the upper limit for best modulation frequencies found in the inferior colliculus which may be as high as 1000 Hz (Langner and Schreiner, 1988), and in the medial geniculate body of the thalamus where temporal resolution is maintained up to about 300 Hz (Rodrigues-Dagaeff 1989). More recently normalised rate modulation data for AI consistent with Schreiner and Urbas' data was presented (Kilgard and Merzenich 1998). To demonstrate the validity of the modeling approach taken, figure 4 shows a comparison between these experimental results and the response of the model to similar acoustic stimuli. As can be seen, the model response replicates the experimental results reasonably well.

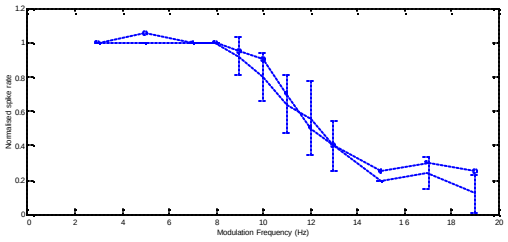

Figure 4. Response to repeated tones at the given repetition rates; model results 'o__o' and experimental results '__' (Kilgard and Merzenich 1998). Normalised repetition rate transfer functions are found using a stimulus consisting of 6 tones pulses at the repetition rate indicated and then calculating the mean response to the last 5 tones in the sequence divided by the response to the first tone; each tone has a duration of 25ms.

Interestingly, although a depressing synapse will need a recovery period in order to regain its full strength, this does not mean that the model is unable to account for the phenomenon of 'period doubling' which is sometimes observed (Kilgard and Merzenich 1998). Since depression is synapse-specific, it is quite possible for incoming activity at other synapses which have not yet become depressed to sufficiently raise the cell's membrane potential so that it fires once more during a single stimulus period. However, because of the cell's refractory period this is most likely to happen at low modulation frequencies, and not at higher ones. An

example of the model reproducing such behaviour is shown in figure 4.

4.4 The time course of forward masking

Although there are undoubtedly a number of factors that contribute to the phenomenon of forward masking, it is clear that the depression of thalamocortical synapses must contribute to the total effect. Explanations for forward masking have also been sought in terms of lateral or forward inhibition. However, it has been shown that masking continues to exist even in the presence of a $GABA_A$ antagonist and therefore even if inhibitory inputs have some part to play they cannot provide a full account (Brosch and Schreiner 1997). Both cortical forward masking and that evidenced behaviourally have been shown to last far longer than explainable in terms of peripheral adaptation (Relkin and Smith 1991, Calford and Semple 1995, Brosch and Schreiner 1997). The model clearly provides a mechanism for forward masking, since synapses that have been previously activated require time to replenish their transmitter stores and respond less strongly when depleted. The time course of synaptic recovery appears to be consistent with the time course of cortical forward masking. The tonotopic distribution of masking is also consistent with a model of forward masking in terms of the depression of thalamocortical synapses since it has been shown that masking is closely related to the receptive fields of cortical neurons (Brosch and Schreiner 1997). Figure 5 shows the depletion at synapses across the tonotopic axis in response to masking stimuli at the intensities indicated. The distribution and time course of synaptic transmitter depletion in this figure resembles Brosch and Schreiner's plots of the time course and distribution of masking.

It should be noted that the incoming activity generated by the peripheral processing due to the masking stimulus starts at 100 ms and only lasts for 30 ms, but this causes the long lasting depletion evident in the figures. The distribution of the transmitter depletion is directly dependent on the nature of the tuning of the filters in the peripheral model. The time course for recovery is determined by the extent of the depletion in the model synapses, which again results from the tuning characteristics of the peripheral filters. However, in the peripheral model processing is governed by much faster time constants which support very rapid recovery, and although there is some adaptation, such long lasting effects are not evident in the peripheral response. This suggests that cortical forward masking may be accounted for by a combination of the characteristics of peripheral processing together with depression of thalaomcortical synapses.

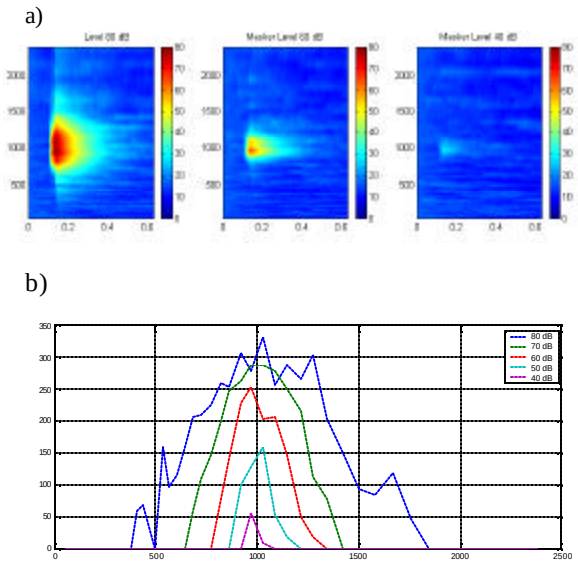

Figure 4. Distribution and time course of transmitter depletion at synapses across the tonotopic axis in response to a 1000 Hz masker of 30 ms duration at the intensities indicated. a) The colour scale indicates the percentage depletion relative to transmitter levels at the start of the masker for maskers of the three intensities indicated. b) Time for transmitter to recover to within 5% of initial mean levels after masker offset for each masker intensity indicated.

Another important aspect of the model is that it demonstrates that cortical forward masking could be dependent on presynaptic rather than postsynaptic activity. This offers a simple explanation for the puzzling experimental observation that masking is sometimes detected even in response to maskers that do not actually activate the target cell (Brosch and Schreiner 1997). If masking is a result of transmitter depletion of thalamocortical synapses, then it would be quite possible for such synapses to become depleted by thalamic activity even though there is

insufficient incoming activity to actually cause the cortical cell to fire, which is how the response to the masker was determined experimentally (Brosch and Schreiner 1997). Since these synapses would nevertheless be depleted, the probe tone could therefore be masked by the 'sub-threshold' masker. This is also consistent with the experimental finding that the suppression of inhibition did not substantially affect the forward masking observed (Brosch and Schreiner, 1997). However, it is possible that lateral inhibition, when present, could also have the effect of producing masking in response to 'subthreshold' maskers, since the masking stimulus may not be subthreshold for local inhibitory interneurons, even if it is so for the pyramidal cell being observed.

4.5 The effect of masker duration on forward masking

In psychophysical experiments it has been shown that the degree of masking is affected by the duration of the masker and masking increases with masker duration (Kidd and Feth 1982). This was also found to be the case by Brosch and Schreiner (1997) in their recordings in AI. Moreover, sensitivity to duration was observed even when the AI cell responded only at the onset of the masker, and although the effect of masker duration was noted, it was not suggested how this could occur. The model investigated here suggests a simple explanation, i.e. as long as there is some tonic incoming activity during the masker, then transmitter depletion at the thalamocortical synapses will be related to masker duration. Therefore, if as we hypothesize, the degree of masking is related to the degree of transmitter depletion at thalamocortical synapses, then the sensitivity to masker duration follows. The paper by Brosch and Schreiner (1997) did not include any detailed results on masker duration, so in figure 5 a similar experimental paradigm to that of Kidd and Feth (1982) is used. However, it should be noted that although the model clearly exhibits sensitivity both to masker duration, it is not clear exactly how the degree of transmitter depletion in the model can be related to the probe threshold shifts measured by Kidd and Feth (1982).

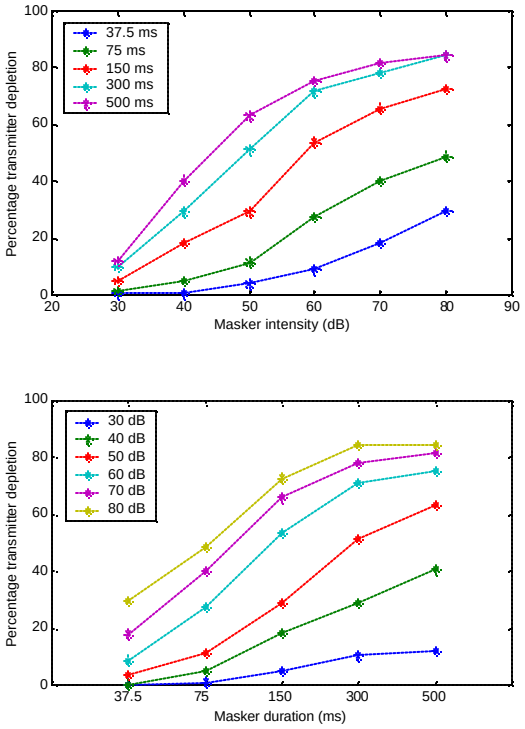

Figure 5. The effect of masker duration. The transmitter depletion relative to mean levels at the start of the masker is plotted for masker duration and intensities indicated. As can be seen the model is clearly sensitive to both masker duration and intensity, as found by Kidd and Feth (1982).

4.6 Disruption of synchronisation responses by subthreshold stimuli

In a recent paper Nelken and Yosef (1998) suggested that their experiments showed a correlate of comodulation masking release. Activity was record in AI in response to noise modulated at 10 Hz, and was found to synchronize to each noise pulse as expected. However, when a very soft, even subthreshold, continuous pure tone with frequency corresponding to the cell's best frequency, was added to the noise, then this synchronization was disrupted. In contrast, when the pure tone was added to an unmodulated noise then the response to the noise alone was indistinguishable from that to the noise plus tone. It was suggested that the cortex might therefore be able to detect masked sounds by means of their disruption of the more powerful masker.

Once again a simple explanation of these results is suggested by the model, which can easily replicate the experimentally observed behaviour as long as there is some tonic

thalamic activity in response to the pure tone. Because the activity in response to the pure tone continues through the silent gaps between the noise pulses, this prevents the recovery of the synapses between noise pulses and so the synchronized response is disrupted. This explanation is also consistent with Nelken's unpublished observations that the synchronized response to the noise alone was far more reliably obtained when the noise was trapezoidally modulated, than when sine wave modulation was used. Figure 6 shows Nelken's experimental results and the model's responses to similar stimuli.

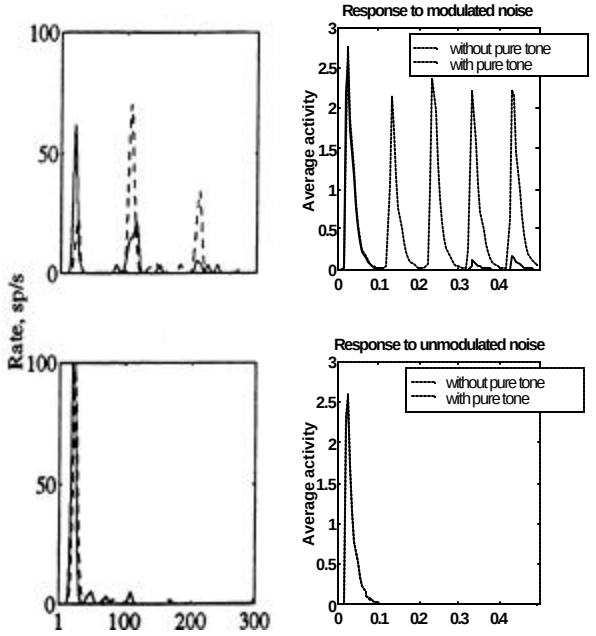

Figure 6. a) Response of neurons in AI (Nelken and Yosef 1998), left column, and the model, right column, to a wideband noise stimulus trapezoidally modulated at 10 Hz, without '---' and with '___' a continuous pure tone. b) Experimental and model responses when the noise is unmodulated.

4.7 Onset latency

Neurons in AI generally respond to the onset of stimuli and to transients in acoustic signals. The factors that influence the timing of the onset response are not well understood. However, it has been shown that for a linear rise function, the onset latency in AI is related to the rate of change of peak pressure, and is independent of rise time and plateau peak pressure (Heil 1997). In figure 7, it can be seen that the model's behaviour is very similar to that observed by Heil (1997). When the latencies are plotted against rate of change of peak pressure, then the latencies for different rise times superimpose quite closely. However, for stimuli that are close to the response threshold, this relationship does not hold so well; an effect also noted by Heil (1997). It was also found that when a cosine squared rise function was used for the stimulus onset envelope, then the onset response latency was related to the acceleration of plateau peak pressure, and this relationship is also clearly evident in figure 7, i.e. onset latency is similar for stimuli in which the initial acceleration of the stimulus envelope is the same, even when the final amplitude reached is different.

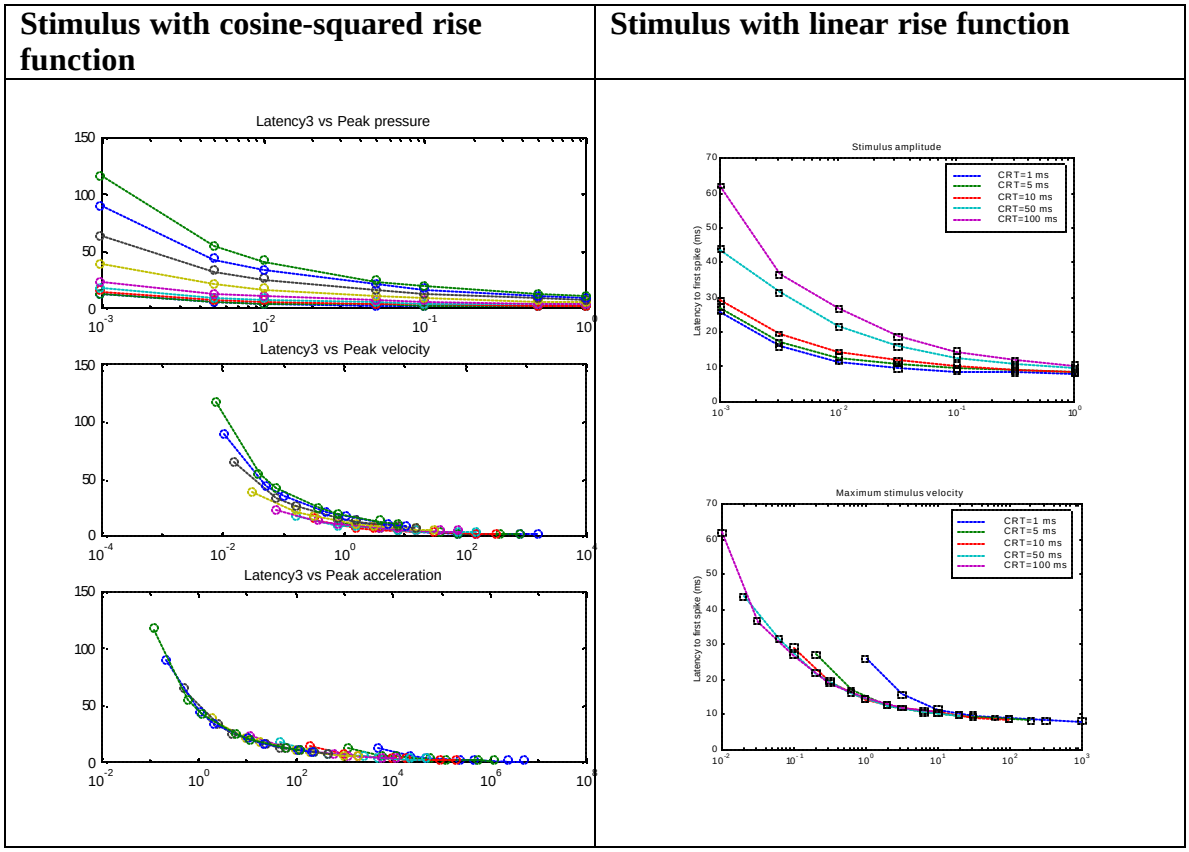

Figure 7. Onset latency for cosine squared and linear rise functions, plotted, as in (Heil 1997), as a function of peak pressure, maximum rate of change of the stimulus onset envelope, and maximum acceleration of the stimulus onset envelope.

4.8 Response to speech

The frequency response of the model, illustrated in figure 2b, is very similar to speech modulation transfer functions, with frequencies around 4 to 6 Hz being the dominant frequency of the envelope of speech signals. Syllables in speech are generally, although not always, distinguished by an amplitude peak preceded and closed by amplitude trough. Therefore, when the model is stimulated by a speech signal, it has a tendency to fire at the onsets of syllables within the signal, as illustrated in figure 8. Synaptic depression may therefore give rise to a syllable-like segmentation of speech signals within AI. Such segmentation could occur in parallel across the tonotopic axis, independently within each frequency channel. This suggestion is consistent with the experimentally observed response to species-specific calls of neurons in AI, which tend to fire primarily at the onset of segments or syllables within calls, irrespective of the characteristic frequency of the neuron (Wang *et al* 1995, Creutzfeldt *et al* 1980). One effect this would have is to increase temporal synchrony across the tonotopic axis thereby promoting the grouping of related frequency components of a call, an effect previously modelled by Smith (Smith 1996). Synchronous activity is likely to be important for the effective transmission of signals to further processing centres which integrate information across frequency channels.

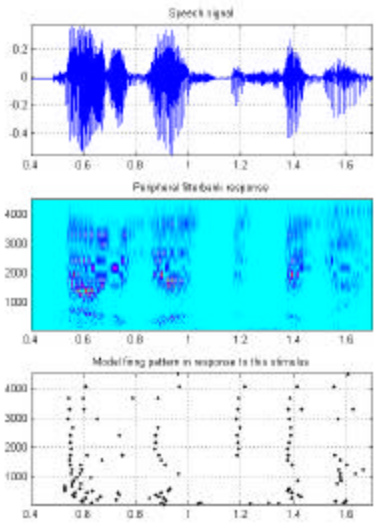

Figure 8. The response of the model to a speech stimulus.

In experiments in which species-specific calls were manipulated (Wang *et al* 1995), it was also shown that speeding up or slowing down the signal resulted in reduced responses. We suggest that the reasons for this differ between manipulations and that the behaviour of the model can help to explain these results. In the first case when the signal is slowed down, activity is still generated in response to syllable onsets, but since these occur at a slower rate, the total amount of activity per second decreases. In the second case, when the signal is speeded up, synaptic depression would prevent synchronisation to syllable onsets as effectively as for the control case. It seems reasonable to suppose that communication sounds have evolved to optimise their detection by cortex, and that the communication sounds that are used are those which are most salient within AI (Wang 1997). Hence, the similarity between the modulation transfer functions of speech signals and of those measured in auditory cortex is perhaps not surprising. Interestingly, although derived from very different principals, the behaviour of the model is very similar to the RASTA filter that was found to markedly improve speech recognition in noise (Hermansky and Morgan 1994).

5. Discussion

In this paper it has been shown how a model neuron which incorporates dynamic synapses responds to a number of different stimuli. The results seem to indicate that synaptic depression at thalamocortical synapses may explain a number of aspects of the temporal response properties of neurons in AI.

The nonlinearity of the dynamic synapse model allows it to behave in many situations like a low pass filter whilst also retaining a fast onset response. In response to repeated stimulation above 10 Hz, synaptic depletion prevents the cell from responding except at the onset of the stimulus. However, the synaptic dynamics are not slow and after a period of rest the synapse can respond with a large EPSP to the onset of a new stimulus, which can result in a response of short latency. The model's behaviour is therefore consistent with the generation of onset responses of short latency and with little jitter. Although the cell tends to respond only at the onset of stimuli, important processing can continue to occur in the dendrites throughout the duration of the stimulus. This allows the cell to exhibit sensitivity to stimulus duration, even when only responding at stimulus onset. In addition, some of the apparent nonlinearities of responses measured in AI, such as the influence of subthreshold stimuli or interaction between different components of a complex stimulus (Nelken and Yosef 1998), could be accounted for in this way.

Since synaptic depression operates at thalamocortical synapses which are the route through which sensory signals must pass in order to get to cortex, it seems likely that the dynamics of depressing synapses have a major role to play in sensory processing. Synaptic depression appears to result in a relatively infrequent sampling of the sensory inputs by cortex, where such information is presumably integrated into ongoing cognitive processes, and perhaps provides a physiological basis for the 'multiple looks' model of temporal integration (Viemeister and Wakefield 1991).

Another effect of synaptic depression is to greatly enhance the response to the onsets of signals. This could act to promote grouping across frequency channels. Synaptic depression effectively provides a kind of lateral inhibition acting in the temporal domain, which may help to increase the temporal contrast of stimuli (Brosch and Schreiner 1997). In the processing of speech signals this results in activity linked to the onsets of syllables, which may help to explain the perceptual salience of syllable onsets. It has also been suggested that although thalamocortical sensory signals on their own cannot elicit lasting activity, facilitation at pyramidal NMDA synapses might act to enhance the response to incoming signals of interest (Thomson and Deuchars 1994). This could provide a mechanism for the flexible processing of sensory signals, depending on factors such as previous experience or the current state of attention.

By taking synaptic dynamics into account in modeling these experiments, it has been possible to account for a number of previously unexplained results in a fairly straightforward way. On the basis of these investigations it is suggested that the dynamics of thalamocortical synapses may help to explain the temporal response properties observed in AI and in auditory perception.

References

Abbott LF, Varela JA, Sen K, Nelson SB (**1997**). "Synaptic depression and cortical gain control", *Science*, Vol. **275**, 220-224.

Brosch M, Schreiner CE (**1997**). "Time course of forward masking tuning curves in cat primary auditory cortex", *J.Neurophys*., 923-943.

Calford MB, Semple MN (**1995**). "Monaural inhibition in cat auditory cortex", *J. Neurophys.*, **73**(5), 1876-1891.

Chance FS, Nelson SB, Abbott LF (**1998**). "Synaptic depression and the temporal response characteristics of V1 cells", J. Neurosci, **18**(12), 4785-4799.

Creutzfeldt O, Hellweg FC, Schreiner C (**1980**). "Thalamocortical transformation of responses to complex auditory stimuli", *Exp. Brain Res*., **39**, 7-104.

Eccles JC (**1964**). *The physiology of synapses.* New York: Academic Press.

Eggermont JJ (**1999**). "The magnitude and phase of temporal modulation transfer functions in cat auditory cortex", J.Neurosci, **19**(7), 2780-2788.

Galarreta, M., Hestrin, S. (**1998**). "Frequency-dependent synaptic depression and the balance of excitation and inhibition in the neocortex", *Nature Neuroscience*, **1**(7), 587-594.

Grossberg S (**1968**). "Some physiological and biochemical consequences of psychological postulates", *Proc. Natl. Acad. Sci, USA*, **60**, 758-765.

Grossberg S (**1969**). "On the production and release of chemical transmitters and related topics in cellular control", *J. Theor. Biol.*, **22**, 325-364.

Heil P (**1997**). "Auditory cortical onset responses revisited. I. First-spike timing', *J.Neurophys.*, 2616-2541.

Helfert EH, Snead CR, Altschuler (**1991**). "The ascending auditory pathways", in Altschuler RA, Bobbin RP, Clopton BM, Hoffman DW (ed.s), *Neurobiology of hearing: the central auditory system*, Raven Press.

Hermansky H, Morgan N (**1994**). "RASTA processing of speech", *IEEE Trans. On Speech and Audio Processing*, **2**(4), 578-589.

Kidd G, Feth LL (**1982**). "Effects of masker duration in pure-tone forward masking", *J. Acoust. Soc. Am.*, **72**(5), 1384-1386.

Kilgard MP, Merzenich MM (**1998**). "Plasticity of temporal information processing in the primary auditory cortex", *Nature Neuroscience*, **1**(8), 727-731.

McCabe SL (**1998**). "Cortical synaptic depression and auditory perception", ", *Proc. Nato ASI on Computational Hearing*, 111-116.

McGregor RJ (**1989**). *Neural and Brain Modelling*, Academic Press.

Markram H, Lubke J, Frotscher M, Sakmann B (**1997**). "Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs", *Science*, **275**, 213-215.

Markram H, Tsodyks M (**1996**). "Redistribution of synaptic efficacy between neocortical pyramidal neurons", *Nature*, **382**, 807-810.

Meddis R (**1986**). "Simulation of mechanical to neural transduction in te auditory receptor", *J. Acoust. Soc. Am.*, **79**(3), 702-711.

O'Mard LP, Hewitt MJ, Meddis R, "DSAM : Development System for Auditory Modelling", http://www.essex.ac.uk/psychology/hearinglab/lutear/home.html.

Nelken I, Yosef OB (**1998**). "Processing of complex sounds in cat primary auditory cortex ", *Proc. Nato ASI on Computational Hearing*, 19-24.

Relkin EM, Smith RL (**1991**). "Forward masking of the compound action potential: Thresholds for the detection of the N_1 peak", *Hear. Res.*, **53**, 131-140.

Schreiner, C. E., Mendelson, J., Raggio, M. W., Brosch, M., and Krueger, K. (**1997**). "Temporal processing in cat primary auditory cortex", *Acta Otolaryngol (Stockh); Suppl.* **532**, 54-60.

Schreiner CE, Urbas JV (**1988**). "Representation of amplitude modulation in the auditory cortex of the cat. I. Comparison between cortical fields", *Hear. Res.*, **32**, 49-64.

Thomson AM, Deuchars J (**1995**). “Diverse pre- and post-synaptic properties of fast excitatory synapses”, in Wheal H, Thomson AM (ed.s), *Excitatory Amino Acids and synaptic Transmission*, Academic Press.

Thomson AM, Deuchars J (**1994**). "Temporal and spatial properties of local circuits in neocortex", *TINS*, **17**(3), 119-126.

Tsodyks MV, Pawelzik K, Markram H (**1998**). "Neural networks with dynamic synapses", *Neural Computation,* **10**(4), 821-835.

Tsodyks MV, Markram H (**1997**). "The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability", *Proc. Natl. Acad. Sci, USA,* **94**, 719-723.

Viemeister NF, Wakefield GH (**1991**). “Temporal integration and multiple looks”, *J. Acoust. Soc. Am.*, **90**(2), 858-865.

Wang X (**1997**). " What is the neural code of species-specific communication sounds in the auditory cortex?", *Proc. 11th Int. Symp. on Hearing*, Grantham, UK, 456-462.

Wang X, Merzenich MM, Beitel R, Schreiner CE (**1995**). "Representation of a species-specific vocalisation in the primary auditory cortex of the common marmoset: Temporal and spectral characteristics", *J.Neurophys.*, **74**(6), 2685-2706.

Table 1. Model parameters used in the simulations.

Parameter	Value
dt	0.0001
α	130
β	4
λ	5
ρ	.08
K	1
γ	1
τ_{EPSP}	.002
τ_{E}	.005
τ_{GK}	.01
τ_{θ}	.001
E_K	-10
θ_0	10
η	100