Modelling Neuronal Mechanisms of the Processing of Tones and Phonemes in the Higher Auditory System

Johan P. Larsson

Tesi Doctoral UPF / 2012

Supervised by
Dr. Gustavo Deco and Dr. Ernest Montbrió
Department of Information and Communication Technologies

Barcelona, September 2012

Dipòsit Legal:
ISBN:

This text was typeset with LaTeX using the `memoir` class.

Denna avhandling är tillägnad min älskade Cati
och tillika mina kära föräldrar Elisabeth och
Lars-Göran, för all deras kärlek, tålamod och stöd

Acknowledgements

I wish to express my heartfelt gratitude to all those who have helped me in various ways during the many years of work I spent on this thesis. You know who you are and what you did for me, so names are not needed here. Thank you so much!

Abstract

Though much experimental research exists on both basic neural mechanisms of hearing and the psychological organization of language perception, there is a relative paucity of modelling work on these subjects. Here we describe two modelling efforts.

One proposes a novel mechanism of frequency selectivity improvement that accounts for results of neurophysiological experiments investigating manifestations of forward masking and above all auditory streaming in the primary auditory cortex (A1). The mechanism works in a feed-forward network with depressing thalamocortical synapses, but is further showed to be robust to a realistic organization of the neural circuitry in A1, which accounts for a wealth of neurophysiological data.

The other effort describes a candidate mechanism for explaining differences in word/non-word perception between early and simultaneous bilinguals found in psychophysical studies. By simulating lexical decision and phoneme discrimination tasks in an attractor neural network model, we strengthen the hypothesis that people often exposed to dialectal word variations can store these in their lexicons, without altering their phoneme representations.

Resumen

Se ha investigado mucho tanto los mecanismos neuronales básicos de la audición como la organización psicológica de la percepción del habla. Sin embargo, en ambos temas hay una relativa escasez en cuanto a modelización. Aquí describimos dos trabajos de modelización.

Uno propone un nuevo mecanismo de mejora de selectividad de frecuencias que explica resultados de experimentos neurofisiológicos investigando manifestaciones de *forward masking* y sobre todo *auditory streaming* en la corteza auditiva principal (A1). El mecanismo funciona en una red *feed-forward* con depresión sináptica entre el tálamo y la corteza, pero mostramos que es robusto a la introducción de una organización realista del circuito de A1, que a su vez explica cantidad de datos neurofisiológicos.

El otro trabajo describe un mecanismo candidato de explicar el hallazgo en estudios psicofísicos de diferencias en la percepción de palabras entre bilinguës tempranos y simultáneos. Simulando tareas de decisión léxica y discriminación de fonemas, fortalecemos la hipótesis de que personas expuestas a menudo a variaciones dialectales de palabras pueden guardar éstas en su léxico, sin alterar representaciones fonémicas.

Contents

1 Introduction

In spite of many decades of research into the function and structure of the primary auditory cortex (henceforth, A1), the intracortical connectivity that is responsible for a considerable part of observed experimental responses is still unknown. In a recent review, this issue was highlighted (Oswald, Schiff & Reyes 2006), and it has been studied both experimentally (Wehr & Zador 2003, Wehr & Zador 2005, Liu, Wu, Arbuckle, Tao & Zhang 2007, Wu, Arbuckle, Liu, Tao & Zhang 2008, Sadagopan & Wang 2010) and using models (Loebel, Nelken & Tsodyks 2007, de la Rocha, Marchetti, Schiff & Reyes 2008). In the present work we intend to shed light on this issue, and we do so by modelling in detail the (putative) neuronal correlates, at the level of the A1 circuit, of a well-known phenomenon in psychology called auditory stream segregation (a.k.a. auditory streaming, AS). Our motivation for taking this approach was twofold. Firstly, of the different experiments in the literature on A1 yet to explain, work on AS stands out for the availability of both psychophysical studies and neurophysiological ones, which gives us both a wealth of neuronal data with which to constrain our models and a bridge (albeit tenuous) to perception from our low-level modelling. Secondly, being a phenomenon studied across the dimension of time at multiple scales, AS allows us to encompass in our study not only neuronal processing at the millisecond scale, but also slower mechanisms requiring from tens of milliseconds to over a second to have an effect. In short, we can incorporate great detail while not losing track of the big picture, something very important in any venture exploring the brain.

To investigate the AS phenomenon in psychophysical experiments, a common stimulus used is a sequence of instances of alternating tones of two different frequencies, henceforth denoted A tones and B tones, respectively. Some studies use $ABAB...$ sequences, while others use so-called triplets, with a silent space in between, i.e., $ABA\_ABA\_....$ Such experiments have shown (reviewed in, e.g., Bregman 1990, Moore & Gockel 2002)

that there are two parameters in particular that have a strong influence on auditory perception: the difference in frequency between the two tones of the sequence (Δf) and the rate at which the tones are presented (Presentation Rate, PR). Two different percepts are possible. For large Δf values and high PR values, the stimulus is perceived as two different streams of sounds, one of A tones, one of B tones (*segregated percept*). In contrast, small Δf values tend to evoke a percept consisting of a single stream of alternating frequency tones (*integrated percept*) (van Noorden 1975, Bregman 1978, Moore & Gockel 2002). For the segregated percept to occur, some time needs to pass after the stimulus onset, referred to as the *build up* of streaming, (up to a few seconds; see Bregman 1978). After an initial time with the integrated percept, listeners report the emergence of the segregated percept. A recent study provided evidence of correlates of this build-up process in A1 (Micheyl, Tian, Carlyon & Rauschecker 2005). Along with earlier studies of neurophysiological correlates of streaming, specifically those discussing the spectral filtering phenomenon called *differential suppression* (DS) (Fishman, Reser & Steinschneider 2001, Fishman, Arezzo & Steinschneider 2004), Micheyl et al. (2005) provide motivation for elucidating the mechanisms in A1, which generate responses in that brain area to this kind of stimulus (however, for a different take on the relative importance of spectral filtering, as a correlate of streaming, see Elhilali, Ma, Micheyl, Oxenham & Shamma 2009).

While there exist some models addressing auditory streaming A1 correlates as manifested by the DS phenomenon (see section 1.3; Kanwal, Medvedev & Micheyl 2003, Denham & Winkler 2006), these fail to fully replicate experimental results and interpret DS as evidence of physiological forward masking produced by either inhibition (Kanwal et al. 2003) or synaptic depression (Denham & Winkler 2006) mechanisms. Crucially, like others before them (Fishman et al. 2001, Fishman et al. 2004), they thus considered that the neural response to the present tone (of a sequence) is *solely* affected by the response to the previous tone. In so doing, they neglected the rest of the stimulus history, an aspect of the problem which we shall show is key to its satisfactory solution.

So, to take on the task of modelling the neural manifestations of AS at the A1 level, we have constructed a model representing a portion of the primary auditory cortex (A1) spanning between zero ('local' model) and six octaves of the tonotopic axis. We employ two different established neural models, representing neural activity of a population of neurons or at

the single neuron level, respectively. The population model is a rate model, describing the average firing rate of a population of neurons (Wilson & Cowan 1972), whereas for the single neuron model we opted for a modified version of the Hodgkin and Huxley equations (Hodgkin & Huxley 1952, Soto, Kopell & Sen 2006, Pospischil, Toledo-Rodriguez, Monier, Piwkowska, Bal, Frégnac, Markram & Destexhe 2008). The motivation for the latter choice was that it enabled a detailed look at the magnitude and time course of synaptic conductances, voltages and currents in response to different stimuli, necessary in our view for full comparability between model results and detailed neurophysiological data. The rate model, in contrast, allowed us to quickly get our bearings in the sometimes daunting parameter space explored, guiding the more detailed simulations and providing food for thought as to what mechanisms were involved in generating the AS data. We stress here, however, that while the dominating goal during the work was to explain how the neural correlates of auditory streaming seen in neurophysiological experiments (Fishman et al. 2001, Fishman et al. 2004, Micheyl et al. 2005) arise at the A1 circuit level, during the literature search we came across other important results which needed to be accounted for by the same circuit. These included approximate co-tuning of excitation and inhibition in A1 (Wehr & Zador 2003, Tan, Zhang, Merzenich & Schreiner 2004, Wu et al. 2008, Tan & Wehr 2009) and forward masking at the same level (Calford & Semple 1995, Brosch & Schreiner 1997, Wehr & Zador 2005), but also both thalamocortical interface properties (Chung, Li & Nelson 2002, Rose & Metherate 2005, Bruno & Sakmann 2006, Wang, Spencer, Fellous & Sejnowski 2010) as well as intracortical connectivity characteristics (Kaur, Lazar & Metherate 2004, Liu et al. 2007, Wu et al. 2008, Sadagopan & Wang 2010), which were implemented as close to those seen in experiments as possible.

Taking the approach outlined above, we have managed to model an important body of research on the neurophysiology of A1, laying the groundwork for an integrated thalamocortical and intracortical network view of processing in the primary auditory cortex. We will explain in detail the results of this effort in chapters 2 and 3. Specifically, chapter 2 describes a novel mechanism that explains the DS phenomenon, which we refer to as *dynamic frequency selectivity*, and then goes on to show how three different neural models can reproduce the most important data of the literature on this phenomenon. Thereafter, chapter 3 tests the robustness of the mechanism to the introduction of an intracortical connectivity, rigourously fit to relevant neurophysiological data. The state of the art of the literature on this main part of the thesis is reviewed in the following (sections 1.1,

1.2 and 1.3).

Furthermore, in an independent part of the thesis, we address a research question of a more cognitive nature, within the field of psycholinguistics. Specifically, we develop a different model able to explain some results regarding the nature of phoneme and word perception in different categories of bilinguals, obtained by Sebastián-Gallés, Echeverría & Bosch (2005). This work is explained in detail in chapter 4, including the necessary background and an introduction to the relevant literature.

1.1. Psychophysics and perception: What is auditory stream segregation?

Surely the reader has heard of the problem of auditory stream segregation, or perhaps the colloquial term *cocktail party problem* or even the technical term *source separation problem* is more familiar. Be that as it may, these terms essentially all concern the same problem — how is a listener able to distinguish between different auditory sources in an environment? This phenomenon occurs frequently, sometimes without us even consciously thinking about it; we listen to a friend speaking in a noisy environment (such as the proverbial cocktail party); we focus on a solo instrument in orchestral music; we find ourselves seated just in between two conversational groups at a dinner party, and both conversations interest us — we switch back and forth trying to pick up a maximum of information. In these examples we employ attention to pick out the auditory 'stream' (typically emanating from the same source; also called auditory 'object', despite its temporal dimension) that interests us, but there is an additional phenomenon of interest related to auditory stream segregation (for short: auditory streaming), namely inadvertent switching between streams. A slightly different version of the last example above serves as an illustration — you pick *one* of the conversations and you're trying to listen to it only, but somehow every once in a while you involuntarily find yourself predominantly hearing the other conversation, missing parts of the one of your choosing. Annoying as it may be in this case, switching could be of more use in a hostile environment where an animal needs to monitor all sounds in order to detect predators and/or conspecifics. In fact, switching can also occur between hearing both sound sources as separate (and mostly being able to switch between them using attentional control) and hearing them as one, integrated, percept (bad in the case of conversations).

The scientific study of the perception of sequences of alternating auditory stimuli has a long tradition. Already more than half a century ago, researchers started mapping the dependence of what humans perceive when listening to such sequences on the basic experimental parameters of spectral distance or *frequency difference* (Δf) between the two constituent tones and the speed at which the sequence is presented, i.e., the *presentation rate*, or *PR* (Miller & Heise 1950, Miller 1947). They showed that while for low enough Δf, subjects perceive the sequence as a stream of two tones continuously alternating in pitch, there exist for each frequency f of the lower tone a Δf value above which the sequence is no longer heard as one stream, but rather two streams of unrelated tones of different pitch. This Δf value is in the study referred to as the *trill threshold*, and the value of $\frac{\Delta f}{f}$ is approximately constant across frequencies f of the lower tone, with a value of 0.15, or between two and three semitones. The dependence on *PR* was less explored, but experiments in Miller (1947) indicated that at slow enough *PR* the perception is of a two-tone stream, whereas for fast enough *PR* the listener perceives two streams.

This area of research has been active ever since these early beginnings. Notably, van Noorden (1975) devoted his PhD thesis to this topic, extending earlier studies by systematically varying the *tempo* of the sequences, i.e., their *PR* , and also examining the influence of subjects' attention on what they perceive. He introduced the concept of *temporal coherence*, i.e. the phenomenon when a listener perceives a sequence of tones as being connected, and a tone sequence evoking such coherence he refers to as a *string* of tones (the latter concept has later come to be called a *stream*, which is the term we will use in this work). One important finding of van Noorden (1975) was that two perceptual boundaries could be defined in the *PR* - Δf space, namely the *temporal coherence boundary* (TCB) and the *Fission Boundary* (FB), respectively. The latter is defined by the value of Δf (measured for all *PR* values) below which a subject's perception ineluctably becomes that of one stream of alternating tones, whereas the former is defined by the values of Δf (across *PR* values) above which said perception always becomes that of two slower streams, each containing tones of one single frequency. We reproduce here a figure from van Noorden (1975) which illustrates this (figure 1.1). The reader will note that the boundaries do not coincide anywhere in the parameter space, despite their obviously complementary nature. This is because the intervening region of parameter values corresponds to an *ambiguous* percept, the subjects reporting either fission or temporal coherence, depending on their attentional state. This is best explained in the words of van Noorden

himself (pp. 82-83, van Noorden 1975):

> It proved possible to determine a clearly defined boundary between temporal coherence and fission by taking the observer's attentional set into account. When the observer's attention is directed towards hearing the string ABAB.., this boundary is found at much larger tone intervals than when his attention is directed towards hearing the string A A or B B.

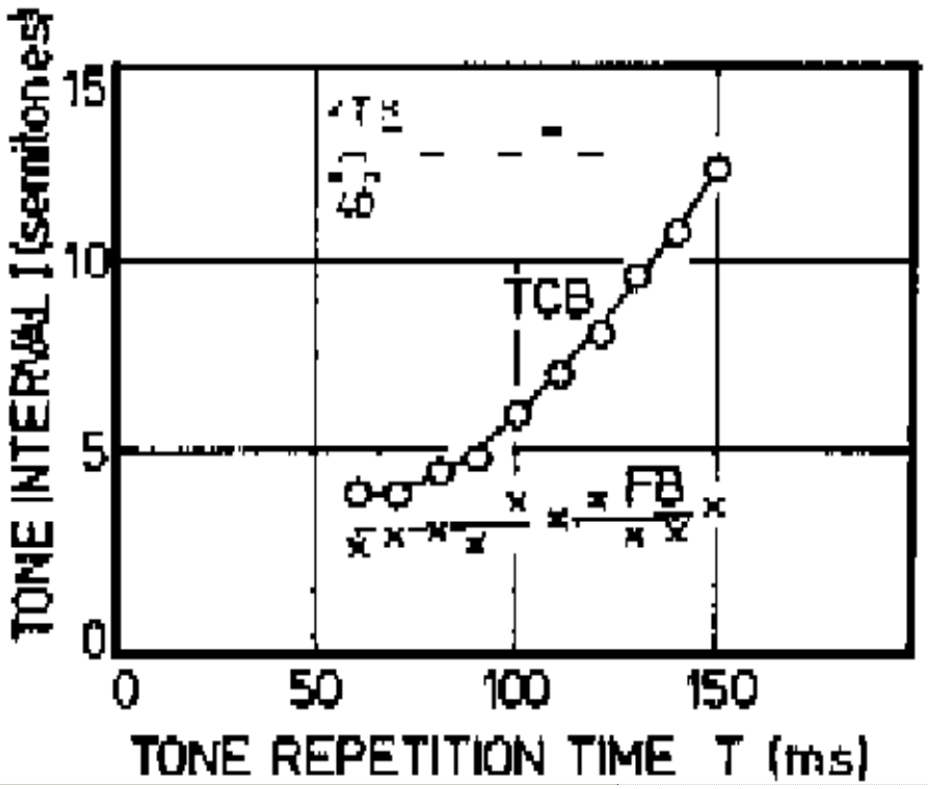


Figure 1.1: Reproduction of figure 2.7 of van Noorden (1975), showing the Temporal Coherence Boundary (TCB) and the Fission Boundary (FB) for a range of values of Δf (in semitones, ordinate) and PR ($T = \frac{1000}{PR}$ ms, abscissa)

Additionally, van Noorden and others have found that, apart from a mere difference in frequency, other differences in the characteristics of the two distinct tones of a sequence could evoke stream segregation,

for instance intensity (van Noorden 1975) and fundamental frequency of filtered harmonic complexes (Vliegen & Oxenham 1999). These and other stimulus manipulations affecting streaming were reviewed by Moore & Gockel (2002).

One phenomenon which we do not study in this thesis, but which nevertheless is important to mention, is that of switching. As summarized in Bregman (1990), switching is the voluntary or involuntary change of perception from that of one stream to that of two streams, or vice versa, even though there is *no* change in stimulus attributes. We will however not cover this topic further.

1.2. Neural basis/correlates: Where might streams (putatively begin to) segregate?

While the study of the auditory streaming phenomenon in humans has a long tradition, as we have seen, researchers have only recently begun to investigate the neural response to such tone sequences in neurophysiological studies performed on non-human animals. About a decade ago, Fishman et al. (2001) made extracellular recordings in the A1 of macaques while presenting sequences of alternating-frequency tones via headphones. The sequences' tones were referred to as A tones (of recorded unit's Best Frequency, BF) and B tones (of a non-BF frequency), respectively. The resulting measured spiking activity (Multi-Unit Activity, MUA, measured extracellularly) was analyzed with respect to primarily the speed of presentation of the sequences, PR, which took the values of 5 Hz, 10 Hz, 20 Hz and 40 Hz. As for Δf, it varied from 10% to 50% difference (of Hz), respectively, but only one value was used for each recording site. For what is to follow in the present work, it is important to note that the stimuli were presented after a long period of silence, and recordings were made from the onset of the very first tone and for a duration of approximately half a second. However, the response to the first tone was discarded in the analysis, and an average of responses to all the other tones (A or B separately) was used when calculating B/A response ratios. The results of Fishman et al. (2001) show that the responses to ABAB sequences have a clear dependence on the experimental parameters, parallelling that found in psychophysical studies such as those mentioned above. Particularly, what they refer to as *differential suppression* (DS), i.e., the phenomenon where responses to non-BF tones are more suppressed by increasing PR than are responses to BF-tones, shows the same trend as when human subjects in experimental studies tend to increasingly segregate sequences into two

streams as *PR* increases. Finding support in the studies on forward masking by Calford & Semple (1995) and Brosch & Schreiner (1997) (see section 1.2.2.2, below), the authors argue that differential suppression is most likely due to synaptic inhibitory processes being more powerful after a BF tone than after a non-BF tone, thus differentially suppressing the response to the subsequent tone in the two cases. They also mention afterhyperpolarization as a possible suppressive mechanism. In their discussion, Fishman et al. (2001) proffered a schematic model of how differential suppression could work, making use of an illustration of hypothetical response profiles across the tonotopic axis (reproduced here, see figure 1.2), evoked by A and B tones at different values of PR and Δf. In short, the model proposes that the net effect of DS is the segregation of A1 activity into isolated 'islands' of activity at different locations on the tonotopic axis, thereby possibly facilitating the direction of attention to one or the other of these loci (see caption of figure 1.2 for details).

In a subsequent study, Fishman et al. (2004) sought to remedy some limitations of their first study by measuring only once the responses had attained what amounted to a steady state, i.e., from the moment in which there were no discernable differences in amplitude between subsequent responses to tones of the same frequency (A or B tones, respectively). Basing this point in time on results from psychophysics (showing a so-called build-up time for stream segregation of several seconds; Bregman 1978, Anstis & Saida 1985), they let stimuli run for at least five seconds before recording. Further improving upon the design of experiments in Fishman et al. (2001), they independently varied, at each recording site, not only PR and Δf but also the tone duration, TD. This was made in recognition of the psychophysical result by Bregman, Ahad, Crum & O'Reilly (2000), showing that, apart from the influence of Δf, the Inter-Stimulus Interval (ISI $= \frac{1}{PR} - TD$) is, rather than only PR , the main factor that controls stream segregation (see figure 1.4(c)). With this setup, the authors get results similar to those of their previous study, demonstrating increasing differential suppression of non-BF tones with increases in Δf and PR (see figures 1.3 and 1.4). They also show, however, that increasing TD has a similiar effect to that obtained by increasing PR , since the ISI diminishes in both cases. They conclude (p. 1665, Fishman et al. 2004):

> The present findings suggest that stream segregation may be due to enhancement of frequency selectivity of A1 sites resulting from the effective sharpening of frequency tuning at the BF tonotopic location at brief interstimulus intervals.

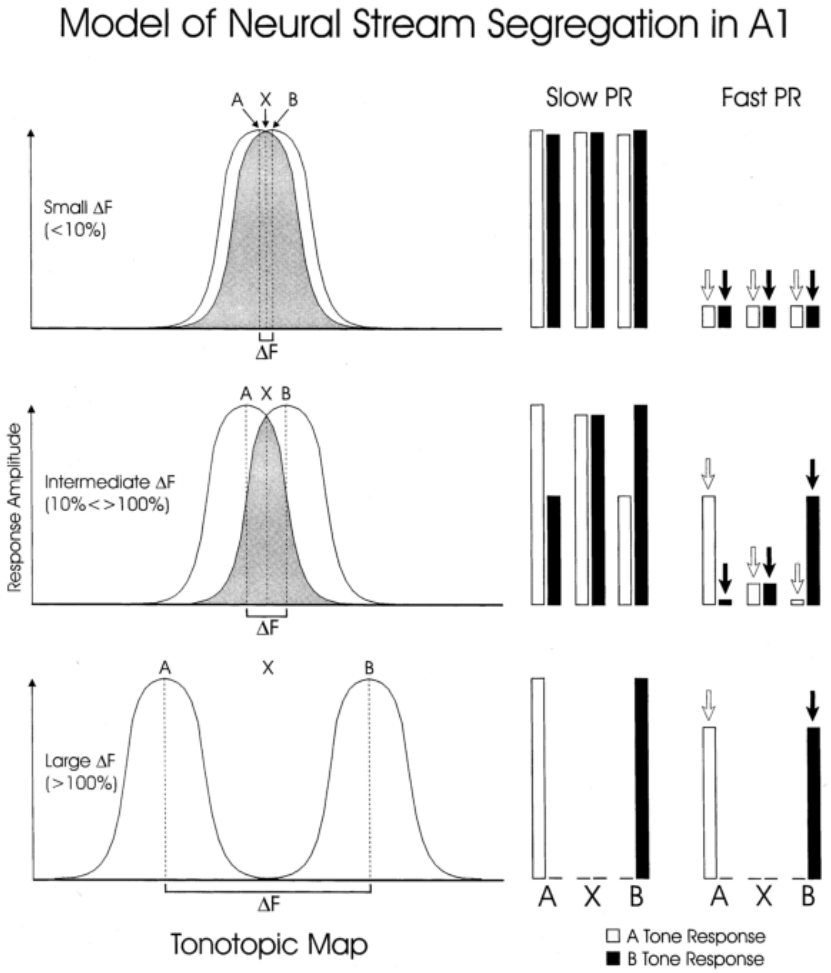


Figure 1.2: Adapted from figure 11 of Fishman et al. (2001), this schematic proposes a physiological model of streams segregation in A1. L.h.s and r.h.s. curves in each graph represent evoked activity across the tonotopic (TT) axis, by an A and a B tone, respectively. Bar plots represent hypothetical response amplitudes to A (white) and B (black) tones at locations A, X and B (see arrows above curves), respectively, at low (left) and high (right) presentation rates. Top panels: $\Delta f < 10\%$; Middle panels: $10\% < \Delta f < 100\%$; Bottom panels: $\Delta f > 100\%$. Crucially, at intermediate Δf (middle), low PR values yield a relatively uniform distribution of activity across the TT axis, whereas high ditto give responses to non-BF tones which are strongly suppressed in comparison to BF tone responses. This yields a segregation of activity into two 'bumps' at BF f_A and BF f_B sites (occurring at half of PR), to which attention then purportedly can be directed.

However, Fishman et al. (2004) maintain their schematic model based on forward suppression from one tone to the next, arguing that forward masking and auditory streaming as manifested in neural data may have a common cause (Calford & Semple 1995, Brosch & Schreiner 1997). As we shall show in the present work, this is in fact not the case.

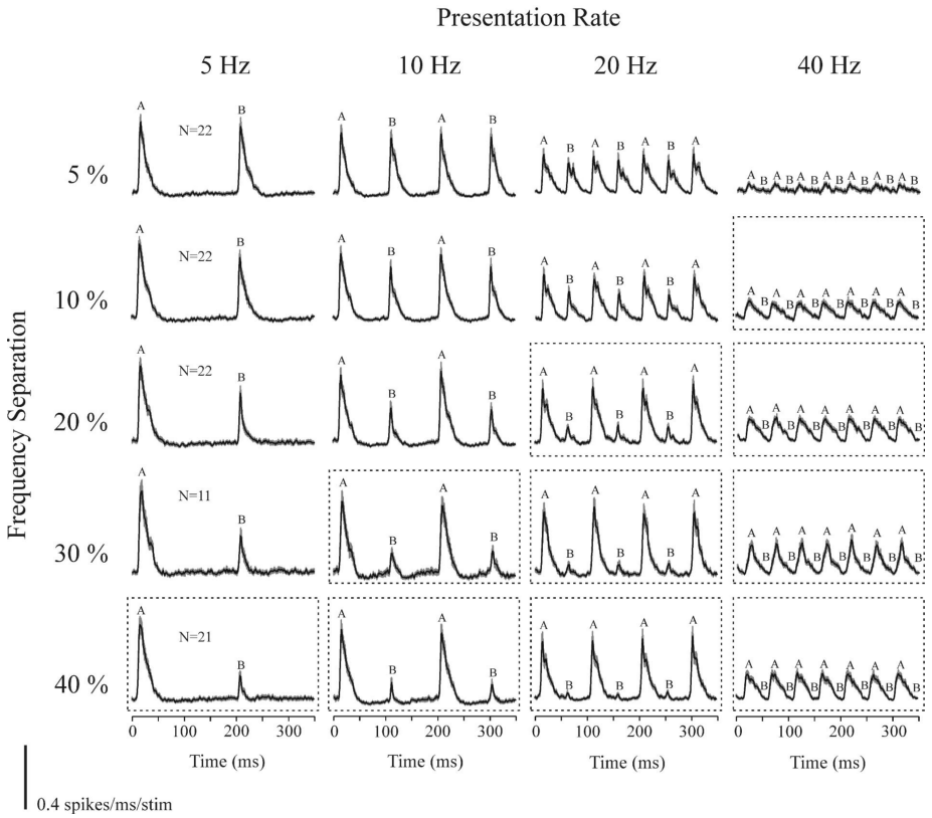


Figure 1.3: Adapted from figure 3 of Fishman et al. (2004). Here, we see the MUA response during 300 ms at steady state, to sequences of (A and B) tones of different PR and Δf values. This data constitutes the main result on auditory streaming neural correlates that we attempt to explain in this thesis. Legend explanation: N is number of recording sites averaged for each row; A and B on top of peaks indicate which tone elicited the peak; dashed frame indicates that that panel's B response is attenuated by more than 9 dB relative to its A response.

Later on, Micheyl et al. (2005) investigated also the so-called build-up phase of auditory stream segregation. They measured from A1 neurons of

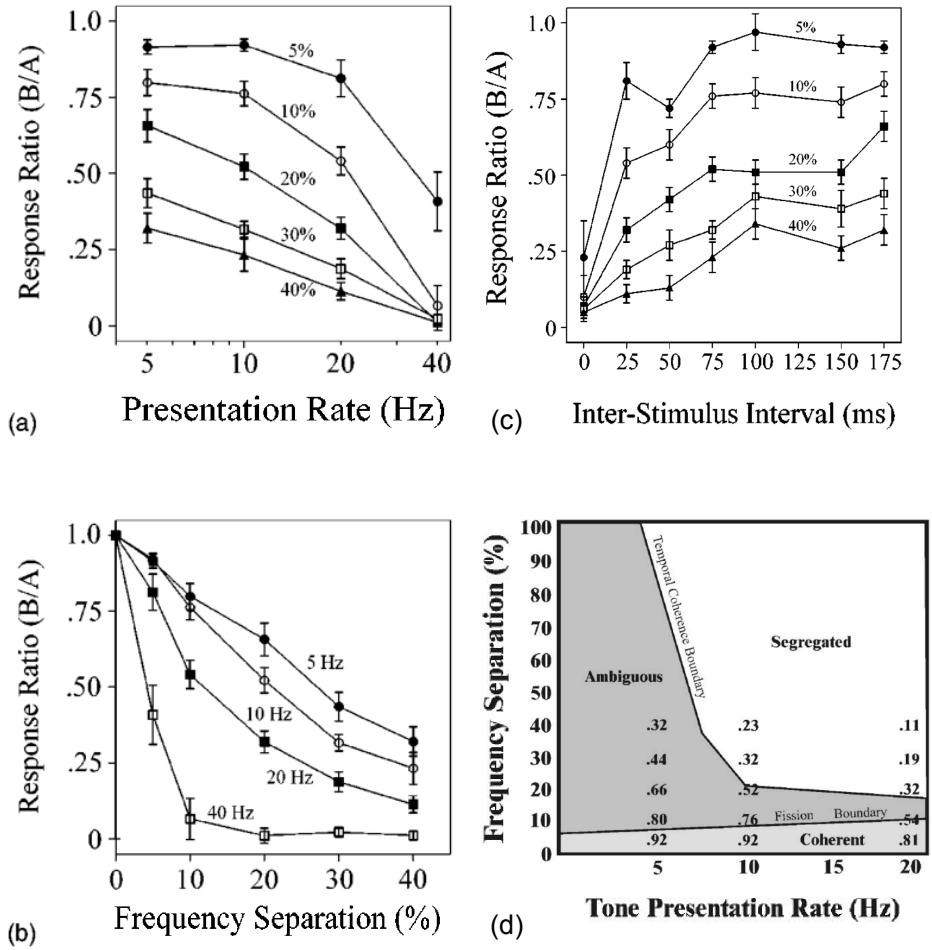


Figure 1.4: All panels adapted from Fishman et al. (2004) (subfigures (a)-(c) from figures 6.a,b,c and subfigure (d) from figure 7.a). Subfigures (a) and (b) show the average ratio of responses to B tones to those of A tones, for 25 ms tones (the main data on streaming correlates we seek to explain in this thesis). Subfigure (c) shows how, in Fishman et al. (2004), the temporal parameter affecting streaming is found to be ISI, dependent on *both PR* and *TD* (see main text). Finally, subfigure (d) superimposes the data from (a)/(b) (except that for *PR* =40 Hz) on a graph delineating the perceptual boundaries of stream segregation in humans (McAdams & Bregman 1979, van Noorden 1975). As Fishman et al. (2004) point out, the smallest response ratios are typically in the 'segregated' perceptual region, whereas the largest ones tend to be situated in the 'coherent' perceptual region.

macaques while presenting sequences of alternating tones to the animals. In parallel, they conducted studies with humans of the psychophysical phenomenon of auditory streaming. Crucially, using a simple threshold mechanism applied to the neural responses, the authors find a neural data explanation to the psychophysical data based on human participants' stated percepts across time, showing a direct link between data from primary auditory cortex and cognitive data.

Taken together, the above studies point to a prominent role for A1 in physiological auditory stream segregation (but for a possible role of the cochlear nucleus, see also Pressnitzer, Sayles, Micheyl & Winter 2008).

However, to really appreciate the importance of A1 in the processing of repetitive sound stimuli, one needs to pay close attention to its *input*, originating in the ventral medial geniculate body (MGBv) of the thalamus, and how this input is transformed on its way to cortex. The next section is dedicated to this topic.

1.2.1. Thalamocortical interface

Several studies in different preparations have shown that the response to amplitude-modulated stimuli is different in thalamus and auditory cortex. For example, both Creutzfeldt, Hellweg & Schreiner (1980) and Miller, Escabí, Read & Schreiner (2002) found that thalamus can respond to temporal modulation approximately twice as fast as that which elicits reliable responses in cortex (the classical results of the former study show that response falls off by half at rates of about 50 Hz in thalamus, but at only 25 Hz in cortex).

The shape of this thalamocortical 'filter' can be either low-pass or band-pass (Creutzfeldt et al. 1980, Kowalski, Depireux & Shamma 1996, Eggermont 1999, Kilgard & Merzenich 1998*b*, Kilgard & Merzenich 1999, Miller et al. 2002), differing across neurons and/or regions of the auditory cortex. Illustrative of this is figure 1.5, which shows the *mean* temporal modulation transfer function (tMTF) found in A1 by Kilgard & Merzenich (1999) (upper panel), and its standard deviation (lower panel). While the mean has a low-pass shape, the s.d. indicates a variety of individual tMTF shapes across neurons. Here, A1 responsiveness falls off even faster than in Creutzfeldt et al. (1980), but in general one may state that A1 responsiveness shows a marked decline in the range of 10 Hz to 25 Hz, approximately.

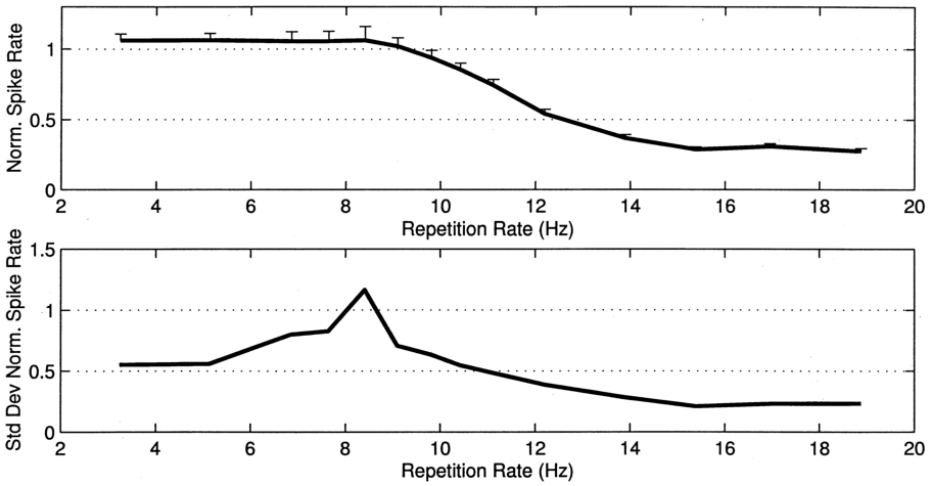


Figure 1.5: Illustration adapted from figure 5A and 5B in Kilgard & Merzenich (1999). The upper panel shows the mean normalized temporal modulation transfer function in A1, while the lower panel shows the standard deviation. While the mean showed a low-pass shape, the s.d. in the repetition rate range of about 4-12 Hz was always above 50% of the mean, peaking at above 100%, indicating a diversity in 'filter' shapes.

The fact that there is such a marked difference between responses at the thalamic and cortical levels renders responses in auditory cortex, and specifically in A1, particularly important in the study of neurophysiological correlates of tone sequence perception. In fact, this is easily realized when one looks at results from relevant psychophysical studies, such as figure 1.1 (van Noorden 1975), where the extremes of both the TCB and FB boundaries are at repetition rates of about 7 Hz and 17 Hz, respectively — the upper limit coincides approximately with the sharp decline of responsiveness in A1 we just highlighted.

A further issue of importance regarding thalamocortical input is *how* it is projected to the cortex. Addressing this, Miller, Escabí, Read & Schreiner (2001) made extracellular recordings in A1 thalamorecipient layers (LIIIb/LIV) and in the cat MGBv, with the objective of quantifying connectivity between the subcortical nucleus and the cortical area measured. When analysing convergence of thalamacortical projections, they found that if a ThC pair of neurons differed by more than 0.25 octaves in their BF, functional connections could not be observed. When instead looking at the convergence from the far edge of the receptive fields of thalamic

cells, they found very little functional connectivity for edges farther than 1/3 of an octave from the cortical cell's BF (in our study, as we shall see, we use this upper limit figure as the basis for our exploration of the effects of thalamocortical axonal spread, denoting the same by ν).

Furthermore, Miller et al. (2001) assessed the strength of the correlations (connections) using the measure of contribution, which is defined as the fraction of cortical spikes immediately preceded by a thalamic spike (1-10 ms lag). Obtaining a mean correlation strength of 0.045, they conjectured that about 20-25 thalamic inputs are required to fully activate a cortical cell. This number is similar to that found by Wang et al. (2010), who took a close look at thalamic drive of cortical responses (experiments and simulations). They found that there are 20-60 simultaneously driven thalamocortical synapses. The optimal corresponding theoretical range, when the reliability per spike was most energetically efficient, was found to be 20-40 ThC synapses active within a time window of 6-8 ms after stimulus onset. Interestingly, Bruno & Sakmann (2006) reach similar results in the somatosensory system. Seeing as there is some agreement on these numbers in the literature, we shall base our detailed model's thalamocortical input on them (see chapter 3).

Returning to Miller et al. (2001), they also found that inhibitory features of thalamic receptive fields are not transmitted with great fidelity, from which they concluded that (p. 157 in the paper)

> ...inhibitory [receptive field] subregions...are constructed or at least strengthened intracortically.

From their paper it is not clear, however, whether this is brought about by inhibitory to excitatory connections only, or also benifits from an excitatory to inhibitory connection boost of intracortical inhibition.

Finally, Miller et al. (2001), finding temporal modulation preferences in the cortex quite distinct from those of the thalamus (results which were very similar to those we summarized above), conclude (as have others) that it seems a functional transform is taking place between these two processing stations. One obvious candidate responsible for this transformation is synaptic depression, as we shall explore in the present work and others have before us (e.g., Denham 2001, Elhilali, Fritz, Klein, Simon & Shamma 2004, de la Rocha & Parga 2005). In the following section, we will review evidence showing that ThC synapses are depressing and that these therefore probably are responsible for the transformation observed at the ThC interface.

1.2.1.1. Experimental evidence for thalamocortical depression

Gil, Connors & Amitai (1999) found that ThC synapses are strongly depressing (more so than IC synapses) and that they have more release sites or functional contacts, on average, than IC synapses (7 vs $\sim$ 2). Further, their findings supported a distribution in release probabilities across synapses, demonstrated earlier by others (e.g., Huang & Stevens 1997).

Much more recently, Viaene, Petrof & Sherman (2011) performed an *in vitro* study on mice, showing that the MGBv sends input to A1 via depressing synapses in LIV. This is called driving input in the article. In contrast, LII/III get so-called modulating input, but not direct, driving, input. More specifically, in LII/III all kinds of responses were present, but the denominated Class 2 response was dominant (20 of 26 responses - 39 cells displayed no response at all). This was found to be synonymous with the aforementioned modulating response. These responses display small initial EPSPs (1-1.5 mV), paired-pulse facilitation and a graded activation profile. Of the remaining six responses, four were mixed (at low intensities they display paired-pulse depression, at higher intensities they display paired-pulse facilitation) and two were Class 1 (i.e., driving input). In LIV, in contrast, there were only Class 1 responses, exhibiting large initial EPSPs ($\sim$ 7.5 mV (A1), $\sim$ 15 mV (S1)), paired-pulse depression and all-or-none activation profile. Additionally, in LIV there was only ionontropic glutamate receptor activation, whereas in LII/III also metabotropic glutamate receptors were active. The EPSP size was correlated with synaptic bouton sizes, i.e., in LII/III boutons were smaller than in LIV.

Many more studies have shown that there is depression in thalamocortical synapses (e.g., Thomson & Deuchars 1994, Rose & Metherate 2005). Given all this data, in our search for a mechanism situated at the thalamocortical interface, able to explain the auditory streaming correlate data (chapter 2), we focussed increasingly on synaptic depression. Thus, we came across two modelling studies addressing related issues using a leaky Integrate-and-Fire model with *stochastic* synaptic depression, something we had not contemplated up to that point. Specifically, de la Rocha & Parga (2005) showed how, when a cortical neuron's membrane potential is situated below but sufficiently close to threshold (called the Fluctuation-Driven Regime, FDR), the added variance in the thalamic input current due to employing stochastic depression increases the output rate of that neuron. This effect is enhanced by correlations across the neuron's affer-

ents, but only in a limited range of input rates, creating a non-monotonic response in the neurons. Later on, de la Rocha & Parga (2008) used the same kind of model in a study of somatosensory cortex neurons' response properties when presented sequences of identical stimuli as input, study which inspired us to explore whether we could, adopting their settings in our conductance-based model with a minimum of parameter changes, reproduce their results and then expand upon them by simulating the response to auditory streaming type sequences of non-identical stimuli (tones). The results of this venture are detailed in chapter 2.

1.2.2. Neurophysiology of the primary auditory cortex (A1)

As we indicated at the outset, one important goal of our modelling has been to investigate the open question of what the connectivity of the A1 circuit (particularly of layer IV) may be. Once we decided to go about this by concentrating on the modelling of the differential suppression phenomenon, we of course needed to base our models firmly on neurophysiological data which could elucidate the relative role of inhibition and excitation in intracortical processing of tonal stimuli. This took on particular importance once we had found a minimal novel mechanism explaining differential suppression (chapter 2), as we looked into the robustness of our solution to the introduction of intracortical connectivity (chapter 3). We here review the most important literature pertaining to this issue.

1.2.2.1. Inhibitory and excitatory response to brief stimuli

Wehr & Zador (2003) look at *in vivo* whole-cell responses of A1 cortical neurons (anaesthetized rats) to the free-field presentation of pure tones (typically 25 ms duration, with a 5 ms 10-90% cosine-squared ramp). The question they address is which synaptic mechanisms are responsible for the tuning of individual neurons to intensity and frequency of input sounds. Specifically, they look at excitatory and inhibitory receptive fields and their relationship and interactions.

Several studies predating this one (references 1-5 in their study) had indicated, using extracellular methods, that excitation and inhibition were organized in a so-called mexican hat fashion - excitation in the receptive field centre and inhibition laterally. The results of Wehr & Zador (2003) counter that result, showing that there exists a *co-tuning* of excitatory and inhibitory conductances evoked in any one cell with changing stimulus parameters (frequency, intensity). The ratio of inhibition to excitation

tended to be fixed for a particular neuron (across the tonotopic axis), but when averaged across cells it varied (0.74 ± 0.07). This ratio value might in reality be a slight underestimation, since the authors state in their supplementary material that cable attentuation effects cause misattribution of a portion of the synaptic conductance to excitation.

As expected from the existence of a spiking threshold, the spiking had a narrower tuning (onesided range slightly more than 2 octaves) than synaptic conductances (onesided range appr. 3 octaves). Typically, excitation evoked a transient spiking response, which was quenched by inhibition after a few milliseconds.

Whereas Wehr & Zador (2003) argue that the dependence of the neurons' response on frequency can be explained solely by the presence of tuned excitation (added inhibition would simply lower the responses and slightly sharpen the tuning), they recognize that the long time course of two-tone (forward) suppression (Calford & Semple 1995, Brosch & Schreiner 1997) can not be accounted for by their data (inhibition too brief). They speculate that such suppressive effects might be inherited from pre-synaptic neurons or be due to short-term plasticity, not specifying which kind (thalamocortical or intracortical; later addressed by the same authors in Wehr & Zador 2005).

Looking at intensity tuning, Wehr & Zador (2003) conclude that inhibition is also not needed to explain the non-monotonic dependence of responses on increasing intensity, which is rather seen as inherited from synaptic inputs. Thus concluding that inhibition is necessary for neither frequency nor intensity tuning, they focus on its tendency to sharpen spike responses by always following excitation after a brief delay. Analyzing this delay, they arrive at an average of 2.4 ± 3.6 ms across stimulus conditions, and a clustering of most values in the 1 to 4 ms range. Whereas this range was quite stereotypical, the onset latency of the excitation varied, which indicated that (p. 444 in paper) "inhibition was locked to the onset of excitation rather than to the stimulus itself".

Interestingly, Wehr & Zador (2003) highlight the fact that the study by Zhang, Tan, Schreiner & Merzenich (2003) proposes long and variable delays between excitation and inhibition as a possible explanation of frequency-modulated (FM) sweep selectivity in the auditory cortex. This of course contrasts widely with the finding here of short stereotypical delays, and they speculate that this could be due to the difference in stimuli (pure tones versus FM sweeps). Finally, by simulating a simple Integrate-and-Fire model the authors demonstrate that if excitatory and

inhibitory conductances are balanced *on average*, this can indeed be a possible explanation of the apparently random firing of cortical neurons (the jitter in the output was larger than in the input), as proposed by Shadlen & Newsome (1998). However, using the same model, they also demonstrate that when imposing a balance such as that seen in their results, with a representative delay of 2.5 ms, one gets a highly transient spiking response where the input jitter is actually larger than that in the output.

In conclusion, they state that their data suggest that *balanced inhibition can sharpen neural responses in time, thereby actually reducing the randomness of cortical activity*.

Notably, Tan et al. (2004) conducted a similar study, published shortly after Wehr & Zador (2003), and while also demonstrating co-tuning of excitation and inhibition with input frequency, they found that inhibition had a much longer duration than excitation, in stark contrast with the results described above. Tan et al. (2004) noted this discrepancy and ascribed it to either age difference of the experimental animals in the two studies (older in the later study) or the fact that Wehr & Zador (2003) used ketamine anaesthesia, while Tan et al. (2004) employed pentobarbital, known to act as an agonist on Gamma-AminoButyric Acid A ($GABA_A$) receptors (Nicoll, Eccles, Oshima & Rubia 1975), thus increasing the duration of inhibition. In fact, this effect of pentobarbital is confirmed by Wehr & Zador (2005) in their subsequent study on forward masking in rats (they primarily use ketamine, but then demonstrate how the duration of inhibition is prolonged by up to 643 % when using pentobarbital), casting further doubt on these long-lasting inhibitory currents of Tan et al. (2004).

Higley & Contreras (2006) found co-adaptation of excitatory and inhibitory conductances in LIV in rat primary somatosensory cortex (*in vivo*, anaesthetized). This was most probably caused by thalamocortical depression, which in that case was equally strong for both fast-spiking inhibitory (FS) and regular-spiking excitatory (RS) cells. Inhibition measured in RS cells was found to be 2.5-3 times as strong as excitation, interestingly a number similar to that found in A1 by Wehr & Zador (2005).

Cruikshank, Lewis & Connors (2007) looked at the underlying synaptic basis for feed-forward inhibition in LIV in primary somatosensory cortex of mice (*in vitro*). They found that inhibitory interneurons respond more strongly than excitatory cells to thalamic input. This was shown to be due to differences in the cell types' synaptic properties, rather than in their intrinsic membrane properties. Specifically, although input resistance was greater in FS cells than RS cells, this was more than compensated by much stronger excitatory thalamic connections onto FS cells than onto RS cells

(mean value more than seven times greater). Furthermore, FS cells have greater innervation from thalamus than do RS cells, as well as slower decay kinetics.

Liu et al. (2007) use a novel approach to selectively silence intracortical currents while keeping thalamocortical currents intact. It has been known for some time that the $GABA_A$ agonist muscimol can cause presynaptic inhibition by activating $GABA_B$ receptors (Yamauchi, Hori & Takahashi 2000), and specifically it was shown in somatosensory cortex of mice that such receptors exist in the thalamocortical pathway (Porter & Nieves 2004). However, Liu et al. (2007) argued, earlier silencing studies had nevertheless employed muscimol to silence the cortex, thus inadvertently adding a $GABA_B$ inhibitory component to the thalamocortical current. To remedy this, the authors came up with a so-called pharmacological cocktail, which combined muscimol with the specific $GABA_B$ antagonist SCH50911 in order to counteract the mentioned effect of muscimol on $GABA_B$ receptors. They were thus able to, more cleanly than ever before, separate thalamocortical from intracortical contributions to the input evoked by tones in A1 of anaesthetized rats, as recorded *in vivo*. They found that the *former inputs determine the width of receptive fields in A1, whereas the latter are almost solely responsible for the actual frequency tuning along the tonotopic axis, shaping it with strong recurrent excitation.*

As we shall see, the Liu et al. (2007) study led us to look into using thalamocortical axonal spread ($\nu > 0$) in our simulations, thereby flattening the receptive fields in A1 (in our feed-forward model of chapter 2) and rendering intracortical recurrent excitation, and to a lesser extent, intracortical inhibition (Wu et al. 2008), crucial in shaping those receptive fields to re-establish the fit to streaming correlate data obtained when *not* using $\nu > 0$ (chapter 3). This issue was well illustrated by Wang (2007) when reviewing the study by Liu et al. (2007) (see reproduced figure 1.6).

Later, Wu et al. (2008) showed that the purported balance between inhibition and excitation in the cortex is at best approximate, in some contrast with results in Wehr & Zador (2003). In particular, performing *in vivo* whole-cell recordings in rat A1 to measure the response to 25 ms pure tones, they found that the frequency tuning curve was broader for inhibitory input than for excitatory input. This led to a sharpening of membrane potential responses of cells at the preferred frequency, medidated by the relatively stronger inhibition on the flanks of those cells' receptive fields. The results of that study, taken together with those of Liu et al. (2007), constituted an important body of neurophysiological constraints in

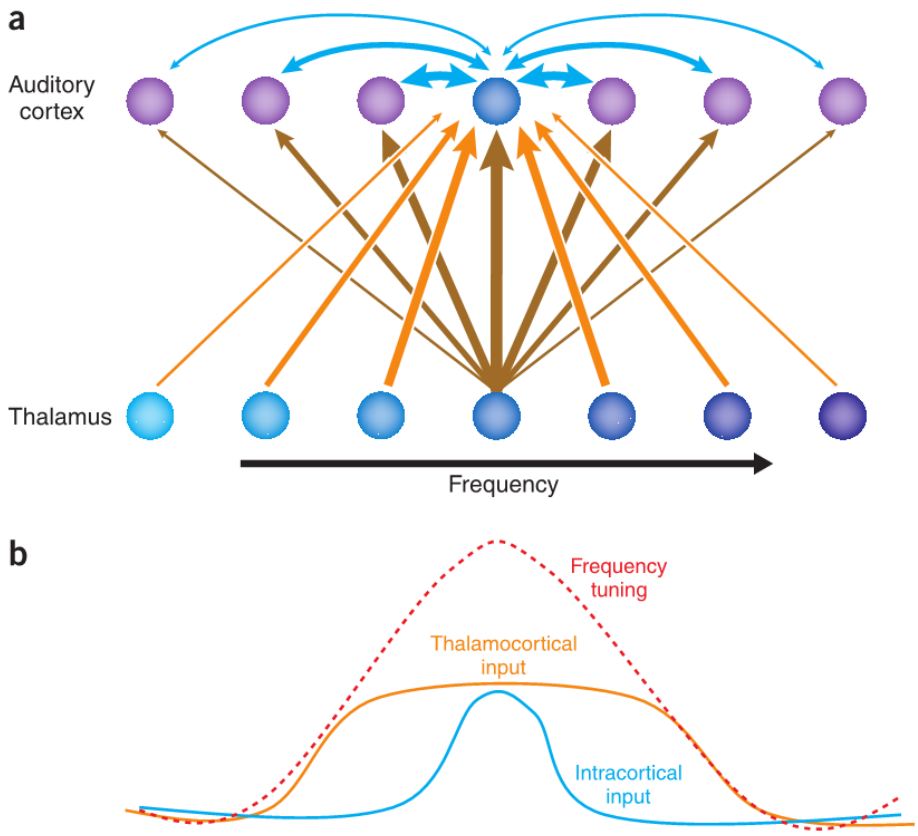


Figure 1.6: Adapted from figure in Wang (2007), this illustration shows how having a broad spectral input to any one neuron in A1 could be compatible with that same neuron displaying quite a narrow frequency tuning in its responses, by way of a strong recurrent excitation generated intracortically.

setting up our elaborate conductance-based model, contributing decisively to us finding a plausible network solution able to reproduce a number of results from experimental literature on A1 (see chapter 3). We will therefore return in more detail to the results of these two studies in that chapter, when describing how said network was set up.

Happel, Jeschke & Ohl (2010) address the issue of temporal and spatial (columnar) organization of responses in A1 to thalamic input. They qualitatively confirm the cortical boost (in amplitude and duration) seen by Liu et al. (2007) and find that horizontal excitatory connections are

both long-range (LII/III) and short-range (LIV). This latter result is of importance, since it confirms the sensibility of setting our intracortical excitatory spread (λ_E) to a low value, in order to reproduce other results from the literature, when we model the full tonotopic axis of A1's LIV (chapter 3).

Volkov & Galazjuk (1991) characterized responses of different cells in A1. They found phasic responses primarily in cells situated in layers III and IV. These cells also displayed a narrow frequency responsiveness, which the authors found to be shaped both by feed-forward and recurrent inhibition. Gittelman & Pollak (2011) show that the relative timing of inhibition and excitation is not as important as other aspects of the EPSPs elicited by sensory input, such as amplitude and EPSP shape, in determining the direction selectivity of a cortical cell. While this result contrasts with those of Wehr & Zador (2003), who argue for a stereotypical delay of 1-4 ms before the appearence of inhibition, it agrees more with Tan & Wehr (2009), whose values have a wider range and in some cases are even negative. In light of this, we did not pursue a more exact fit of a stereotypical delay when setting up our network (see chapter 3), but rather emphasized the range and relative strength of excitation and inhibition.

Isaacson & Scanziani (2011) recently reviewed what is currently known about inhibition in cortical circuits. Important issues they raise are, for instance, that inhibitory (GABA-releasing) neurons comprise about 20 % of all cortical neurons. Inhibitory neurons inhibit principal cells and are excited by the same, in a reciprocal connection pattern (feedback or recurrent inhibition). Afferent fibers from other parts of cortex or from subcortical nuclei, such as the thalamus, make excitatory connections onto both principal cells and interneurons. Some authors argue that these afferent connections are stronger onto interneurons than onto principal cells, but they typically studied somatosensory cortex (Gabernet, Jaddhav, Feldman, Carandini & Scanziani 2005, Cruikshank et al. 2007, Daw, Ashby & Isaac 2007, Hull, Isaacson & Scanziani 2009) (but for similar results in A1, see Levy & Reyes 2012, Schiff & Reyes 2012). As we shall see, this issue is important when attempting to reproduce different results from the neurophysiology literature on A1 in one and the same model (chapter 3).

Isaacson & Scanziani (2011) further point out that interneurons in turn inhibit principal cells (feed-forward inhibition), and that GABAergic interneurons inhibit each other, the latter connections being strongly reciprocal. We specifically include these inhibitory-to-inhibitory connections in our elaborate network model (chapter 3), and manage to reproduce recent neurophysiological results with these present, in contrast with other studies that do not include such connections (de la Rocha et al. 2008, Levy

& Reyes 2011). Finally, the authors of the review of cortical inhibition emphasize that the 'balance' between inhibition and excitation is only approximate, and is useful as a guiding concept but is not to be taken literally in terms of conductance shapes, for instance. This is in line with our reasoning in chapter 3 on this very issue.

1.2.2.2. Neurophysiological two-tone/-click forward masking

Calford & Semple (1995), in an *in vivo* study of two-tone suppression carried out in A1 of anaesthetized cats, found that all neurons that had a non-monotonic response-level function also had a non-monotonic masking-level function, implying that masking inhibition is dependent on the excitatory profile. Forward masking inhibition is probably generated at a high level of the auditory pathway, possibly at the cortical level. For most cells in the study, a 50 ms masker tone at CF elicited clear suppression of the response to a subsequent probe tone at CF, at probe delays <= 140 ms. Notably, the time between each masker-probe stimulus pair is not given in the study, which means one can not be sure whether recordings of responses were made with neurons in their habituated or recuperated state.

Brosch & Schreiner (1997) also studied forward masking in the primary auditory cortex of cat, using pure tones as stimuli. They found that the frequency range of inhibiting masker stimuli increased with masker intensity, i.e., the so-called masker tuning curve was V-shaped, like most receptive fields. The maximal duration of forward inhibition was in the range of 53 - 430 ms, with an average of 143 ms and a standard deviation of 73 ms. Their results show that cortical cells' excitatory response to stimuli depends on the temporal context of those stimuli. The more successive stimuli differ in spectral content, the higher the rate of repetition at which the neurons respond.

This issue was further looked into by Wehr & Zador (2005), using noise stimuli and recording in rat A1. They found that both spike responses and synaptic inputs remain suppressed for hundreds of milliseconds (to a 2nd click after 512 ms, neurons still only responded at 80% of their response to the first click), and since inhibitory conductances in their study rarely lasted longer than between 50 and 100 ms, the authors concluded that forward suppression must involve other mechanisms than inhibition, such as synaptic depression at either thalamocortical or intracortical synapses. It was further found that for click intervals of less than 128 ms, the response to the 2nd click was almost completely suppressed, on average. Correspondingly, conductances (exc./inh.) were suppressed for the same

intervals, further supporting an alternative mechanism to that of inhibition as responsible for forward masking. The time course of recovery was very similar for exc. and inh. conductances, on the one hand, and spiking activity on the other, which possibly points to a thalamocortical mechanism.

Two issues make this study different from other forward masking studies: they use clicks instead of tones and they present them at a higher intensity, 102 dB (other studies use about 60-70 dB). The fact that they use clicks is important for two reasons (in comparison to pure tones):

- Neighbouring A1 neurons get more input since a wider portion of MGBv is activated. This can contribute to the response of the measured neuron through intracortical connectivity, even without ThC spread.

- The measured A1 neuron gets more input from neighbouring MGBv neurons, if there is ThC spread, which augments response.

The above observations are important in case ThC synaptic depression is the main mechanism for forward masking, since both mechanisms of increasing input will increase depression, both its speed and its amount. This means one cannot compare directly to the other studies' results. Another complicating factor is that other studies used barbiturates instead of ketamine, substantially increasing the duration of inhibitory currents evoked by the tones. While this extra suppression might in fact make the studies more comparable, it further muddles exactitude. Finally, it is important to note that the time between click pairs is long enough for neurons/synapses to recuperate completely, which contrasts with other studies where responses are seen as habituated due to shorter inter-pair intervals (e.g. Brosch & Schreiner 1997).
All these considerations in the end led us to focus on Wehr & Zador (2005) in our efforts to reproduce and explain neurophysiological forward masking data in chapter 3.

1.3. Previous relevant modelling studies

There are several modelling studies carried out by others, that are relevant to this work. In the following, we will make a brief summary of the most important ones. We first include a fundamental study on the nature of synaptic depression (Abbott, Varela, Sen & Nelson 1997), because of its high relevance to our work (although technically referring to intracortical synapses, its results are equally valid for thalamocortical ones). Thereafter,

we describe the main studies focussing on A1, in two subsections.

In a combined experimental and modelling study, Abbott et al. (1997) investigated the effect of depression in afferent synapses on the response of a neuron. For a constant input rate, r, they found that the steady-state amplitude of the neuron's response was approximately proportional to $1/r$, for large r. With the added assumption of synapses adding linearly, this yielded that at steady-state, the total synaptic response of the neuron becomes independent of the constant input rate. However, similar to the psychophysical phenomenon described by the Weber-Fechner law (Fechner 1880), they found that the neuron maintained its sensitivity to sudden changes in the input rate, a change Δr evoking a response proportional to $\Delta r/r$. In other words, it produced a transient synaptic response proportional to the change in *percent* of the input rate.

Abbott et al. (1997) further showed that *with* depression, a 50% sinusoidal modulation of the input rate in half of a neuron's afferents (the other half's input being constant and of a markedly different rate) is reflected in its response, regardless of the absolute rate of the input in question. In contrast, *without* depression, small absolute modulations (e.g., 50% of 10 Hz, i.e., 5 Hz) presented simultaneously with a high constant rate in the other half of the afferents, are not reflected in the output.

Finally, Abbott et al. (1997) demonstrated that the tuning of a neuron's response to an arbitrary sensory attribute (x; such as frequency in A1) is degraded but not eliminated by adding depression to its afferents. Simultaneously, though, another effect of depression is to increase sensitivity to sudden changes in x. This, the authors explain, is because synapses that previously received low input rates and are relatively undepressed, can respond more vigorously to a transient change in input rate. In addition, they found that *with* depression, synchronous uncorrelated rate changes across many afferents produced a change in synaptic conductance that was independent of the number of afferents, n. In contrast, the same synaptic response falls off as $1/\sqrt{n}$ without depression.

1.3.1. Existing models of differential suppression in A1

Kanwal et al. (2003) conducted the first modelling study of this phenomenon, combined with neurophysiological experiments, in bats. They claimed that a model network incorporating lateral and especially recurrent inhibition was sufficient to explain their own experimental results as well as those of Fishman et al. (2001), and they described their mechanism as a purely forward masking affair, similarly to the argument of Fishman

et al. (2001) (at the BF f_A site, A tones evoke more excitation, which in turns evokes more recurrent inhibition, than do B tones — thus, the masking of B responses by A responses is more marked than vice versa, eventually suppressing responses to B tones at F_A). However, we intend to show in the present study that this is not the whole story, as using only inhibition as a mechanism would fail to account for other results with the same network (notably, the time course of recovery in physiological forward masking; Wehr & Zador 2005). We note, however, the experimental result of Kanwal et al. (2003) that the activity of a single neuron is not enough to (re-)produce neural correlates of auditory streaming experimentally (specifically, differential suppression), whereas summing the activity of on the order of 10 neurons is sufficient.

Later on, Denham & Winkler (2006) proposed a neural model with synaptic depression in its thalamocortical synapses (first proposed - see next section - in Denham 2001), which addressed the neural correlates of streaming as manifested in Fishman et al. (2004). While reproducing some data from that experimental paper, their model neglects to exhaustively explain the mechanism behind differential suppression. In particular, the view of the authors is that differential suppression is evidence of physiological forward masking (from one tone to the next), coinciding with the view of both Fishman and coauthors and Kanwal et al. (2003). The novelty of their study was to propose that this forward masking was produced by thalamocortical depression, and while the study by Denham & Winkler (2006) was therefore an important inspiration for us, we will go on to show in our work that this view of differential suppression as a manifestation of physiological forward masking is not an adequate explanation for the phenomenon.

1.3.2. Existing models of other aspects of A1 response properties

Notably, there exist modelling works looking into the issue of the thalamocortical interface transformation using synaptic depression in the thalamic afferents of A1. The first one of importance was Denham (2001), which qualitatively reproduced the data on this found by both Creutzfeldt et al. (1980) and Kilgard & Merzenich (1999). In that same study, Denham (2001) demonstrated that synaptic depression can explain several other phenomena found in both neurophysiological and psychophysical studies, most notably results on forward masking (Brosch & Schreiner 1997, Kidd & Feth 1982). In particular, it was shown that synaptic depression could account for the time course of forward masking and that the model's depletion

of thalamocortical synaptic resources during a tone explains the increased masking caused by longer tones (found in psychophysical experiments by Kidd & Feth 1982). Denham (2001) further hypothesizes that this same depletion could explain masking seen even when the masker did not evoke a spike response (Brosch & Schreiner 1997). An important motivation for the Denham (2001) study was the so-called resolution-integration paradox, and in the conclusions it is suggested that synaptic depression at the thalamocortical interface is a suitable candidate mechanism for explaining this, as it can maintain onset responses (resolution) while also providing the long time-constants needed for integration across time. This issue was again addressed in the study we turn to next.

In a combined experimental and simulation study, Elhilali et al. (2004) sought an answer to the question how it is possible for the response of auditory neurons to, on the one hand, follow very fast transients in the input (up to 200 Hz), but on the other hand, be limited in its response to slower modulations of about 20-30 Hz (i.e., the filtering mentioned earlier). In their experiments, using a stimulus exhibiting fast oscillations 'riding' on top of slower ones, they first showed that more than 70% of neurons (their response measured extracellularly) exhibited this phenomenon. They then proceeded to show that a model of synaptic depression (with a very fast time constant, $\tau_D = 65$ ms) and/or a simple cortical network of feed-forward thalamic excitation and intracortical inhibition can in principle exhibit this dual behaviour. This model served as part of the inspiration for using synaptic depression in modelling differential suppression in this work, although we of course emphasize more the aspect of biophysical realism in our models, especially in the conductance-based ones.

Saeb, Gharibzadeh, Towhidkhah & Farajidavar (2007) employ a three-layer feed-forward network of Integrate-and-Fire neurons to model changes in responses of the primary auditory cortex to sequences of tones, as seen in various experiments on auditory behavioural learning (Bao, Chang, Woods & Merzenich 2004, Kilgard, Pandya, Vázquez, Gehi, Schreiner & Merzenich 2001). The layers represent the subthalamic, thalamic and cortical levels of processing, respectively, and are connected by simultaneously depressing and facilitating synapses. The authors show that the cut-off frequency of the cortical neuron's frequency response can be moved up and down by varying the parameters regulating the level of facilitation in synapses of the two layer interfaces of the model. While this study is interesting and related to what we study through the use of depression and sequences of noise bursts as stimuli, it addresses a different issue,

behavioural learning in animals, which occurs at longer time scales with which we do not concern ourselves here.

In a different study, Loebel et al. (2007) apply a rate model to the study of various phenomena seen in A1, in particular excitation and inhibition co-tuning and two-tone forward masking. They make approximate fits of these two phenomena using *intracortical* depression, but never attempt to model differential suppression *per se*. We discuss their study further in Discussions and Conclusions (chapter 5).

Furthermore, de la Rocha et al. (2008) look at the underlying network architecture's influence on response tuning in A1, performing neurophysiological experiments and simulating a tonotopically organized rate model of A1. The authors essentially show that depending on the model parameters, particularly those governing the characteristics of the thalamic response and the inhibitory and excitatory input it gives rise to in cortex (displaying either co-tuning of excitation and inhibition or lateral inhibition, their so-called CON and LIN network setups, respectively), one may get broader (CON) or narrower (LIN) frequency tuning of responses, different temporal response profiles and monotonic (CON) or non-monotonic (LIN) dependence of spiking responses on input intensity. de la Rocha et al. (2008) cover a lot of ground with which we do not concern ourselves here, such as the intensity-tuning of cortical responses, which displays a variety of shapes in the literature. However, they also address some things overlapping with our study, notably the nature of exitatory and inhibitory responses in A1, and we studied their work in some depth while developing our own.

Finally, very recently, a few other relevant studies came out (Levy & Reyes 2011, Schiff & Reyes 2012), further studying the cortical microcircuit of A1 with emphasis on LII/III. We will briefly discuss these in chapter 5 (Discussion and Conclusions).

2 Proposed mechanism for explaining differential suppression

Part of the work of this chapter was presented by the author of this thesis at the Eighteenth Annual Computational Neuroscience Meeting, held in Berlin, Germany in July 2009[1]*, and a paper is about to be submitted for publication*[2].

2.1. Introduction

Neurons in the ventral medial geniculate body (MGBv) provide the main input to the primary auditory cortex (A1), in particular to layer IV (Viaene et al. 2011). While neurons in MGBv are able to sustain responses to sounds repeated regularly at rates up to 100 Hz, A1 practically fails to respond when the presentation rate (*PR*) goes much beyond 20 Hz (Creutzfeldt et al. 1980, Kilgard & Merzenich 1998*a*, Kilgard & Merzenich 1999, Eggermont 1999, Denham 2001, Eggermont 2002, Zhou & Merzenich 2008), corresponding to the perceptual boundary between rhythmic and continuous sounds.

Furthermore, at presentation rates below 20 Hz and for sound sequences composed of two tones alternating in frequency (see figure 2.1), the A1 neurons show a frequency-selective filtering known as *differential suppression* (DS). This filtering produces a refinement of the A1 neurons' frequency receptive fields, in such a way that the neurons become more selective to a tone's frequency as the tone's *PR* is increased (Fishman et al. 2001, Kanwal et al. 2003, Fishman et al. 2004, Bee & Klump 2004, Bee & Klump 2005, Micheyl et al. 2005). This phenomenon is believed to be a

[1] Abstract published: Montbrió, Larsson, Almeida & Deco (2009)
[2] Montbrió, Almeida, Larsson & Deco (2012)

neural correlate of auditory stream segregation, the process by which sound stimuli are organized into perceptual sound streams, reflecting individual sound sources (Bregman & Campbell 1971, Bregman 1990).

DS is measured by recording the neural responses of an ensemble of neurons of a particular best frequency (BF) in A1 to a sequence of alternating BF and non-BF tones, and then calculating the ratio of the non-BF tone response to the BF tone response. The neurophysiological basis of the DS phenomenon is poorly understood, and it is the subject of intense experimental investigations (Snyder & Alain 2007, Micheyl, Carlyon, Gutschalk, Melcher, Oxenham, Rauschecker, Tian & Wilson 2007, Elhilali et al. 2009, Bidet-Caulet & Bertrand 2009, Shamma & Micheyl 2010, Shamma, Elhilali & Micheyl 2011). In contrast, computational models of the DS phenomenon in A1 are scarce, and are either not consistent with recent neurophysiological experiments (Kanwal et al. 2003), or fail to explain the neuronal mechanisms responsible for the DS and to fully replicate experimental results (Denham & Winkler 2006).

Specifically, the suppression of non-BF responses has been interpreted in previous work as evidence of physiological forward masking, produced either by inhibition (Fishman et al. 2001, Kanwal et al. 2003, Fishman et al. 2004) or by synaptic depression (Denham & Winkler 2006). According to this interpretation, during the presentation of the present tone in a sequence, A1 neurons are suppressed differentially depending on whether (*only*) the previous ('masker') tone in the sequence elicited strong (BF tones) or weak (non-BF tones) spiking activity in their neighbouring cells (indirectly related both to inhibition and depression level).

However, recent experimental results using the forward masking experimental paradigm cast doubts on the role of inhibition, suggesting that suppression cannot come solely from cortical cells. Firstly, although inhibition could in principle contribute to the suppression observed within up to 100 ms after stimulus onset, since forward masking lasts more than 500 ms it has to be largely attributed to other mechanisms, such as synaptic depression (Wehr & Zador 2005). Secondly, inhibition measured in A1 intracellular recordings often lacks the necessary properties to support strong lateral inhibition: neurons most often receive balanced excitatory and inhibitory inputs activated by sensory stimulation (Wehr & Zador 2003, Zhang et al. 2003, Tan et al. 2004, de la Rocha et al. 2008, Wu et al. 2008, Tan & Wehr 2009). Thirdly, A1 LIV neurons are mainly driven by strong thalamic excitatory input, which results in robust spike responses (Liu et al. 2007, Wu et al. 2008, Viaene et al. 2011). This in combination with evidence for thalamocortical depression (e.g., Thomson & Deuchars 1994, Gil, Connors & Amitai 1997, Rose & Metherate 2005)

points to a role for ThC depression (Denham & Winkler 2006).

Moreover, mechanistic explanations of the DS phenomenon based on forward masking overlook a relevant part of the stimulus' history, since they consider the neural response to the incoming tone to be affected only by the response to the preceding tone. Tones in the DS phenomenon are separated by silent gaps of at most 200 ms, but it is known that auditory stimulus history strongly contributes to shaping A1 neural responses over periods of time up to seconds (Ulanovsky, Las, & Nelken 2003, Ulanovsky, Las, Farkas & Nelken 2004, Wehr & Zador 2005, Asari & Zador 2009). Furthermore, DS is observed in the steady-state of the neural responses evoked by continuous tone sequences (Fishman et al. 2004). These conditions strongly differ from those of the forward masking experimental paradigm, where masker sounds are presented after long periods of silence.

Here we propose a feed-forward mechanism for DS in the primary auditory cortex, based on thalamocortical synaptic depression and the threshold-nonlinearity of the A1 neurons' response. Thalamocortical synaptic depression is responsible for the suppression of neuronal responses at high stimulation frequencies in the auditory (Eggermont 1999), and somatosensory (Chung et al. 2002, Garabedian, Jones, Merzenich, Dale & Moore 2003) cortices. These filtering properties have been the focus of analysis of a number of computational studies (Eggermont 1999, Denham 2001, Loebel & Tsodyks 2002, Elhilali et al. 2004, de la Rocha & Parga 2008). Additionally, there exist theoretical works on forward masking in A1 based both on intracortical (Loebel et al. 2007) and thalamocortical (Denham 2001) synaptic depression. However, which are the synaptic, neuronal and network mechanisms shaping the neurons' frequency receptive fields during *repetitive* auditory stimulation using two tones of different frequencies? This is a very relevant auditory system research question, which is at the basis of the auditory stream segregation problem.

To address this issue, we chose to focus on the few essential ingredients sufficient for a model to be capable of reproducing the experimental data. In order to do that, we necessarily relaxed the demands on neurophysiological accuracy of the model, both in terms of its constituent parts and, to some extent, regarding the values of parameters that we *did* include in the model. Doing this required a process of trial and error, the result of which was a feed-forward model of an isofrequency patch of A1 consisting exclusively of excitatory neurons, leaving out inhibitory neurons, the tonotopic extension of cortex and intracortical connections. We find that with the joint effect of the spiking threshold of neurons (providing non-linearity of responses), the thalamocortical depression (establishing the steady state), the tonotopic organization of the thalamic response (making response in the modelled

patch of A1 tonotopic as well, via ThC connectivity), and a gradual increase or decrease of model response with experimental parameters (either by fiat or through crossneuronal variability in responses, depending on the level of modelling detail), we are able to reproduce the data of Fishman et al. (2004) for short tones. Specifically, the threshold explains the decline of response ratio with presentation rate; the depression explains the decline in absolute response to any tone with increasing presentation rate; the tonotopic organization of thalamic response explains the decline of response ratio with increasing frequency difference between a sequence's tones; finally, the gradual change of model response with experimental parameters was decisive in reproducing a smooth decline in both absolute responses and response ratios with increasing presentation rate and frequency difference. In the rate model this is simply imposed by design, through the use of a linear f-I characteristic (or transfer function; see equation (A.2) in appendix A) for the A1 population. However, when we simulate a whole population of conductance-based neurons of Hodgkin-Huxley type (see appendix B), this can not be imposed but relies instead upon a certain level of variability across neurons (whether intrinsically through varying neurons' resting potential, or through a spread in the values of thalamocortical synapses' parameters (e.g., U, τ_D, g_{CTh}), which shape the effect of external inputs impinging on neurons). We thus demonstrate that variability in the intrinsic qualities of single neurons and synapses substantially contributes to the solution.

In short, in this chapter we not only reproduce the data found by Fishman et al. (2004), but analyse the underlying possible causes of the DS phenomenon at the circuit level (something not attempted in, e.g., Denham & Winkler 2006) and arrive at a simple yet powerful mechanism that explains *how* stream segregation correlates arise at the level of the primary auditory cortex. We further confirm that our mechanism works as claimed in three different models, one rate model and two conductance-based models, without the need for intracortical connections. Finally, we discuss possible shortcomings of these models when introducing more realistic thalamocortical connectivity, laying the groundwork for addressing these issues in chapter 3.

2.2. Methods

We now proceed to describe the methods employed in this chapter. After first clarifying the most important nomenclature, we go on to present the three different models employed, one a rate model and two based on the

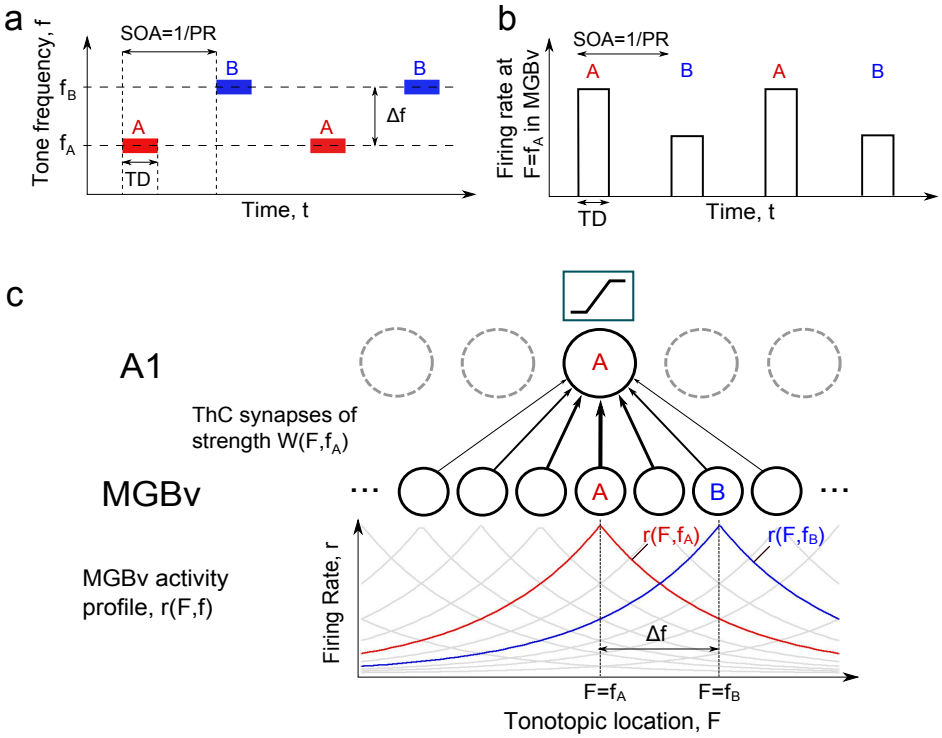


Figure 2.1: Description of the model. **a**: Schematic representation of the stimuli used to study auditory streaming, consisting of sequences of alternating A and B pure tones of frequency $f = f_{A,B}$, duration TD and presentation rate PR . **b**: The temporal dynamics of the firing rate at any particular location F of the thalamus (in this case $F = f_A$) is modelled as series of short (about 20ms) excitatory pulses of magnitude $r(F, f)$. **c**: Thalamic (MBGv) neurons connect to A1 neurons via thalamocortical (ThC) depressing synapses. A certain tone of frequency f elicits a response across MBGv modelled by the function $r(F, f)$.

Hodgkin and Huxley (HH) single-cell model (Hodgkin & Huxley 1952). We further describe how, in our modelling context, we choose to represent the tone sequence stimuli typical of auditory streaming experiments. Finally, we include a section on data analysis.

We emphasize already here that while we employ different models with different specifics as to parameter values and level of modelling detail, as described in the following, the schematic describing our model setup (see figure 2.1) is equally valid for both the rate model and the HH models. In addition, in appendix B the figures B.2 and B.3 illustrate specifics of the

tonotopically organized HH model, but it is not employed until chapter 3.

2.2.1. Nomenclature

Throughout our modelling study of A1, we will deal with frequency space at different 'levels' of description. First of all, there is the physical world, in which complex tones, or more generally, sounds, are made up of many frequency components. One such frequency component can be considered a sinusoid oscillating at a single frequency. We will refer to such a sinusoid, when presented in isolation, as a *pure tone* of frequency f. Secondly, the (complex) sounds of the physical world are decomposed into their underlying frequency components in the auditory system, starting already at the level of the basilar membrane of the cochlea in the inner ear. This decomposition is subsequently maintained at each level of neural processing of the auditory system, including the MGBv and A1. This phenomenon is referred to as tonotopy, being as it were a topographical representation of the tonal (i.e., auditory) world in neuronal space. We consider that the MGBv has such a tonotopic axis, a point on which is represented by the symbol F. For instance, the tone $f = f_A$ evokes activity in the thalamus which is centered at $F = f_A$ (its neuronal response is not modelled explicitly, but represented by a rate function). Finally, at the level of A1 there is also a tonotopic axis, along which locations are denoted by φ. Maintaining these clear distinctions is important for formulating the model with exactitude (see appendices A and B).

Using the above nomenclature, we will proceed to name some important concepts. In the experimental study of the auditory streaming phenomenon using series of alternating pure tones, one of the most important parameters is the tones' frequency difference, which we define as

$$\Delta f \equiv f_A - f_B. \tag{2.1}$$

In our study we will keep f_A, the frequency of the A tone, fixed, and vary f_B. It is important to note that equation (2.1) denotes an *experimental parameter*, and not a distance on either of the two tonotopic axes we consider in our model, at the thalamic and cortical levels of description. Such distances are described in terms of F and φ, respectively. Again, although these different quantities (f, F, φ) are obviously related, we make these distinctions for reasons of exactitude of discourse.

We further define the tone presentation rate, PR, as the inverse of the time elapsing between the onsets of subsequent tones in the sequence (the

stimulus onset asynchrony, *SOA*, here given in seconds). Thus,

$$PR \equiv \frac{1}{SOA}. \tag{2.2}$$

While *PR* strictly speaking has units of tones/s, we consider our sequences to be (albeit finite) periodic functions and thus employ units of Hertz (Hz), for simplicity.

Finally, although we use the term 'tone' to refer to external stimuli presented to our models of A1, we never use actual tones to generate our thalamic response rate, but consistently make use of Poisson processes for this purpose. Essentially, we do this since the peripheral auditory system has several processing centers below the thalamus which we do not model either (a black box made up of black boxes, as it were), and thus it was deemed easier to base the rate of response at the thalamic level on data from literature and thereafter artificially generate that rate. At an early stage of modelling, we did feed computer-generated sinusoids (pure tones) into DSAM (Development System for Auditory Modelling; see http://dsam.org.uk and O'Mard & Meddis 1997) and that software subsequently generated auditory nerve spike trains, which we then fed directly to our cortical model. However, since we also in that case skipped many levels of processing in the ascending auditory system (notably the cochlear nucleus and the superior colliculus), and seeing as results did not differ noticeably when we used Poisson rates instead (data not shown), we thenceforth decided to exclusively use the (more) artifically generated Poissonian input.

2.2.2. Rate model

We implemented a rate model in the programming language C (based on Wilson & Cowan 1972). All the details of this model, including analytical results on synaptic depression (valid for any neural model we use), can be found in appendix A. Here, we include a brief description.

We consider a model of a set of excitatory neurons in the primary auditory cortex (A1). The model describes the mean firing rate of this population of neurons and has a threshold-linear saturating input-output function. The neurons have a certain best frequency $f = f_A$ (see general model schematic in figure 2.1.a), and receive feed-forward exitatory input from the ventral Medial Geniculate Body (MGBv), via a set of thalamocortical (ThC) synapses that are considered to be depressing.

The auditory stimulus consists of a sequence of short tones alternating in frequency, as depicted in figures 2.1.a and 2.1.b (see appendix A for a detailed description and justification of the model). The sequence is presented

with presentation rate PR and frequency difference Δf , corresponding to the auditory streaming paradigm.

We further consider that the auditory stimuli alternately excite two different regions of the MBGv. In particular, A tones excite a region centered around $F = f_A$ in MBGv (red curve in figure 2.1.c), whereas B tones excite a region centered around $F = f_B$. These regions are highly overlapping, since the streaming stimuli have small frequency differences $\Delta f \sim [0, 1]$ octaves (McAdams & Bregman 1979), and the MBGv frequency receptive fields span approximately 2-3 octaves at the sound intensities typically used in the streaming experiments ($\sim$ 60 dB) (see, e.g., Liu et al. 2007).

2.2.3. Hodgkin and Huxley model

This model is described in detail in appendix B. Briefly, we made our own implementation of conductance-based Hodgkin and Huxley type neurons in the programming languages C and C++. We used kinetic equations from the literature (Soto et al. 2006, Pospischil et al. 2008) to define both regular-spiking (RS), excitatory, neurons and fast-spiking (FS), inhibitory, neurons, all based on the Hodgkin and Huxley (HH) formalism (Hodgkin & Huxley 1952). We used a population of RS and FS neurons to represent an arbitrary point along the tonotopic axis of A1 (Sally & Kelly 1988), representing a certain characteristic frequency (CF; the frequency of an external stimulus which elicits a response at the lowest intensity) and a certain best frequency (BF; the external stimulus frequency which elicits the strongest response). The two quantities BF and CF may sometimes differ for one and the same neuron, but throughout this study we shall make the simplifying assumption that they coincide and consequently use them interchangeably. Although we later concern ourselves with modelling the extent of the tonotopic axis using an array of populations, or 'columns' (see chapter 3), we here model *one* such population ($N = 1$), responding maximally to the A tone (i.e., with BF/CF f_A). Furthermore, in this chapter our inhibitory neurons are 'dormant', i.e., we do not activate them here (but see chapter 3), since we focus on the role of depression in shaping the response of a strictly excitatory neural population to sequences of tones.

We further implemented an elaborate input scheme, whereby we could periodically present tones (represented by a non-zero thalamic firing rate function, rather than the firing of explicitly simulated thalamic neurons) at a particular rate of presentation (Presentation Rate, PR). The thalamic firing rate function has both spatial and temporal modulation, is denoted

by $R_{\mathrm{Th}}(F, f, t)$ and is generated through an inhomogeneous Poisson process (see equation (B.41), found in section B.3, which thoroughly explains the input to the model). Resulting presynaptic spikes are then processed by thalamocortical synapses which display synaptic depression of either a stochastic or a deterministic nature (see section B.1.1), giving rise to currents entering our cell models, where they are converted into membrane voltages and subsequently action potentials (spikes). The presence of synaptic depression at thalamocortical synapses has been shown in several studies (notably Thomson & Deuchars 1994, Gil et al. 1999, Rose & Metherate 2005).

We simulated this model using two different approaches. The first uses a model where the necessary variability is achieved using a population of neurons, with intrinsic and/or synaptic properties (parameter values) differing across neurons. This model uses the deterministic, average, depression model, which consequently doesn't add any randomness to the outcome of a simulation. For all matters concerning how we set up the basic network for this approach (of one or more 'columns'), we refer the reader to section 3.3.1 of chapter 3, which contains details on how we at all times tried to be consistent with known literature when setting our parameter values.

Contrasting with the above, the second approach uses a smaller population of neurons (strictly used for averaging purposes, emulating different trials of the experiment by using different neurons in the model), each of which employs the stochastic depression model and consequently displays randomness in its responses due to this choice of synapses. In this second approach, we used the studies by de la Rocha & Parga (2005) and de la Rocha & Parga (2008) to setup the network, and then proceeded to expand on the results of those authors by studying sequences of sensory stimuli *differing* in their sensory attributes — while de la Rocha & Parga (2008) applied their model to the somatosensory modality and the study of same-attribute sequences, we of course study the auditory modality and auditory streaming stimuli, whose constituent tones differ in frequency. As we have seen in chapter 1, neither the underlying neurophysiology of the primary areas of these two modalities, nor the nature of their thalamocortical input, is all that different from each other, which warrants the use of this model for our purposes. In fact, the basic setup of our first approach is qualitatively similar to the second approach, which in turn strived to change as little as possible the parameters used by de la Rocha & Parga (2008).

There were however some additional features that needed to be implemented in order for us to be able to simulate the second approach.

Specifically, we introduced stochastic synapses (see section B.1.2), including a queue structure using pointers to handle the vesicle replenishment process; we used time-varying stimuli and inhomogeneous Poisson processes (acceptance-rejection method) to handle them; much work also went into habilitating the network and the data file handling for using Dirac delta-type inputs instead of continuously varying gating variables; finally, we implemented thalamic and cortical background input. Once implementation issues were resolved, our strategy for reproducing the data was to first identify the parameter values needed to place us in the fluctuation-driven regime (essentially reproducing figure 3A-D in de la Rocha & Parga 2005), for both deterministic and stochastic synapses. Thereafter, we focussed our efforts on fitting the results of the first part of de la Rocha & Parga (2008), which deals with repetitive pulse stimulation (i.e., $\Delta f = 0$ in our case). Once that was achieved, we made minimal adjustments to the parameters in order to reproduce the streaming data (with a mere lowering of q_{CTh}, it was enough, as we shall see).

Whenever we make use of cortical background input, we always employ the parameter values of de la Rocha & Parga (2008): Poisson processes generate excitatory and inhibitory pre-synaptic spikes at rates $r_{BG,E} = 5000$ Hz and $r_{BG,I} = 1000$ Hz, respectively. When one of these excitatory or inhibitory spikes arrives at a neuron (a 'synapse' for this particular purpose), it causes the release of a vesicle at any one of $M_{s,E} = 3$ and $M_{s,I} = 6$ functional contacts, respectively, with probabilities $U_{BG,E} = U_{BG,I} = 0.4$. Thalamic background input, when we use it, is set at 5 Hz, following the same paper. It of course enters through the same thalamocortical synapses as external input.

2.2.4. Representation of auditory streaming stimuli

As already mentioned, for the HH models, equation (B.41) is used to model the input of a stimulus. Specifically, the function $R_{\mathrm{Th}}(F, f, t)$ describes the spread of response in the ventral Medial Geniculate Body of the thalamus (MGBv; whose neurons, we emphasize, are not explicitly simulated) to an input from lower auditory centers, centered at the site $F = f$. To model the alternating tone sequences of Fishman et al. (2004), we start by denoting the two distinct tones A and B tones, of frequencies f_A Hz and f_B Hz, respectively. We then fix the A tone at position $F_A \equiv F_0$ (the centre of the tonotopic axis), and let the position of the B tone, F_B, vary with the value of the experimental parameter Δf. In our study, $F_B > F_A$, but results are equally valid for the case $F_B < F_A$, due to symmetry (in agreement with results in Fishman et al. 2004). For modelling the

short tones used in Fishman et al. (2004), we employ three different tone (stimulus) durations (TD values). When we use temporal modulation of the rate, we employ the function defined in equation (B.40) in some cases ($TD = 80$ ms, but most input discharged in the first 25 ms) and use a tone of $TD = 25$ ms with cosine-squared up and down ramps of 5 ms each in other cases (see results). When we have no temporal modulation (constant rate, used for easier comparison with analytical results), we take $TD = 20$ ms. The value of the integral across either stimulus is approximately equal, i.e., the same amount of input is apported by each type of stimulus, only slightly differently distributed in time. Note that the rate model is always simulated with a constant rate input representing each tone, as explained in detail in appendix A.

2.2.5. Data analysis

Both in the case of a local feed-forward network and in that of an extended recurrent network (chapter 3), we measure the activity in response to A or B tones in the population which represents cells with CF f_A, i.e., placed at position φ_A in the case of the extended network. Specifically, spikes are summed from the onset of a tone to the instant before the onset of the next tone, i.e., during periods of duration SOA, until the end of the simulation is encountered. We use a time bin size of 2.5 ms, unless stated otherwise. For presentability of summed activity we sometimes smooth the activity with a Gaussian filter having width σ_G the same as the bin size. For calculating the ratio of responses to B tones at φ_A to responses to A tones at φ_A, we sum the total activity in each SOA time slot, and average all responses to A tones and all responses to B tones across a period of time where steady-state depression has been reached. The response ratio (RR) is then equal to the B average divided by the A average. The data analysis for the rate model follows the same principles just outlined.

When reproducing data from de la Rocha & Parga (2005) and de la Rocha & Parga (2008), and subsequently applying that modelling approach to looking at streaming stimuli, we use the following quantities to discuss our results (de la Rocha & Parga 2005, de la Rocha & Parga 2008).

$$P_t(t) = \frac{\sum_{i=1}^{K} \frac{n_{rel,i}}{M_S}}{K} (\text{at instant } t) \tag{2.3}$$

$$P_t = \frac{1}{T_2 - T_1} \int_{T_1}^{T_2} P_t(t) dt \tag{2.4}$$

$$\mu_I = \frac{1}{T_2 - T_1} \int_{T_1}^{T_2} I_{\text{CTh}}(\varphi_A, t) dt \tag{2.5}$$

$$\sigma_I = \frac{1}{T_2 - T_1} \int_{T_1}^{T_2} I^2_{\text{CTh}}(\varphi_A, t) dt \tag{2.6}$$

$$\text{CSR} = \frac{1}{T_2 - T_1} \int_{T_1}^{T_2} \delta(t - t_k) dt, \tag{2.7}$$

where $t = T_1$ and $t = T_2$ are the starting point and the end point of the interval of analysis, respectively (values given in results), and $n_{rel,i}$ is the total number of vesicles released by synapse i at instant t. $P_t(t)$ is the instantaneous value of the probability of transmission, averaged over each neuron's synapses. All quantities are specific to one neuron — we then average these across the neurons of our population to emulate several trials.

2.3. Results

2.3.1. Understanding the role of depression

We have carried out an analytical study of the equations governing depression (see appendix A, section A.2). This analysis yielded exact solutions which give the value of our depression variable, d, at steady state, at four different (time) points in the sequence. We put 'time' in parenthesis, because once in the steady state our solutions are independent of *absolute* time since the sequence started. Rather, the four points are defined at four *relative* time points, defined within the span of two subsequent *SOA* periods comprising an A tone and a B tone, namely before and after the two tones (exactly defined in section A.2). We specifically make extensive use of the value of d before the A tone and the B tone, called d_0 and d_2 (equations (A.25) and (A.27)), respectively, in order to save simulation time and to quickly get a feeling for how depression controls the response of our model to the input. In order to use this tool, one needs to make some

assumptions regarding the input — it must be constant in time *during* each tone (the input rate may not be time-varying); the sequence must be uninterrupted (no silent periods interlaced) and use a fixed tone duration and presentation rate; and finally, thalamocortical input has no axonal spread ($\nu \equiv 0$) but projects to cortical neurons with the same BF/CF as that of the sending thalamic 'neuron'. These restrictions on the input need to be respected in full simulations aimed at a direct comparison with the analytical values, or else one must assess the difference in response due to any discrepancies in input characteristics.

We take four complementary approaches when using these steady-state values. The first makes use of analytics on its own to calculate a surrogate solution for the response ratio curves seen in Fishman et al. (2004), as we shall detail below. The second sets depression values to d_0 or d_2 (depending on the first tone presented) before starting a simulation, which may be very short (a single tone) or run for a while to acquire averages of responses to tones. The third approach is similar to the second, but instead of using the steady state values in the simulations, we set the start value of d to a value of our choosing at command-line[3]. Then, after running simulations for covering the range of possible values of d, we instead use d_0 and d_2 when analysing the data from those simulations, essentially 'reading out' the response of the network to a certain tone at any d value. Specifically, by computing d_0 and d_2 values for chosen values of the parameters of the solutions (A.25) and (A.27), respectively (PR, Δf, $R_{\mathrm{Th}}^{\max}$, TD, U and τ_D), we obtain the model response at each of them and can therewith calculate $RR\,(PR\,,\Delta f)$. Finally, in our fourth approach we make a fixed-point analysis of the response of our rate model to sequences of tones, simplifying further by showing that using an average value of the steady state depression is sufficient to approximate RR.

The results of our four approaches are described in the following section.

2.3.2. Assessing response through analytics and minimal simulations

The first approach using our analytical results to assess response at steady state is also the most straightforward one. Simply using the parameter values of interest, we calculate the steady state values of d_0 and d_2 for a range of Δf and PR (figure 2.2).

[3]We implemented an approximate solution for doing this also for stochastic synapses, by setting the right fraction of vesicles as released, properly distributed across all synapses of each neuron (with the downside of them never recovering, since their release time is unknown).

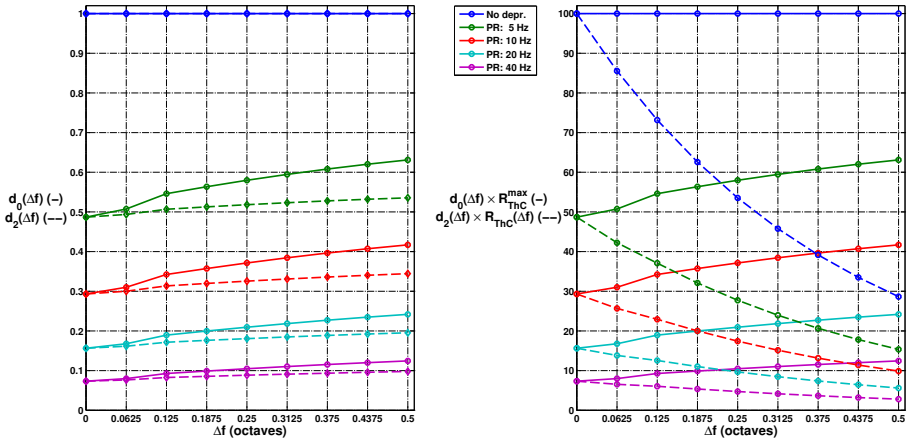
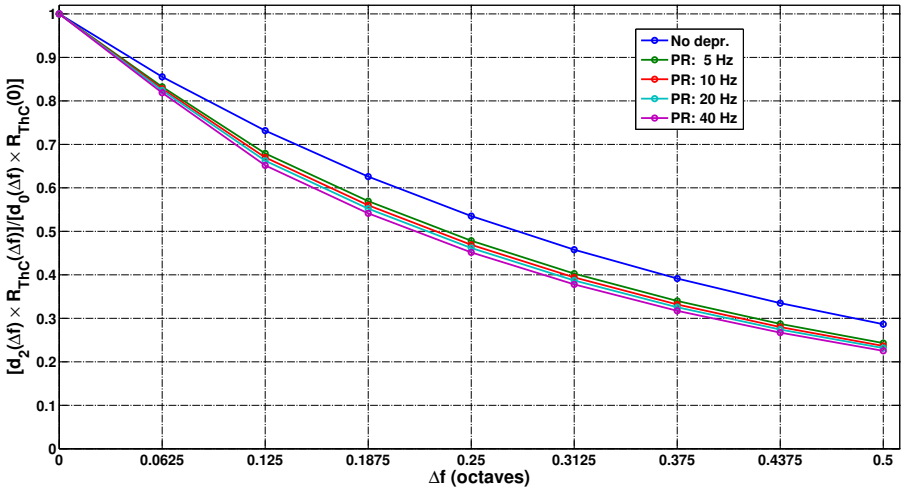


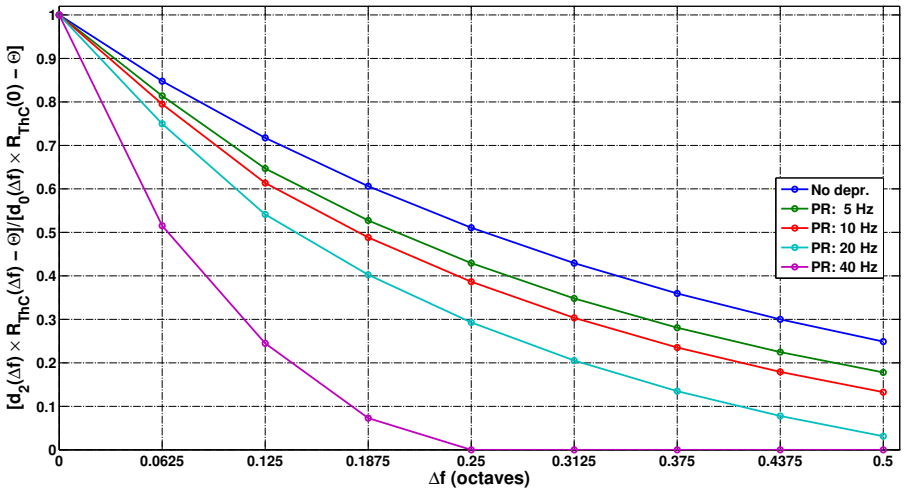
Figure 2.2: Making of ABAB... sequence 'tuning curve' - d_2 level multiplied by R_{Th} yields an approximation of the input evoked by a B tone at each Δf and PR , whereas d_0 level multiplied by $R_{\mathrm{Th}}^{\max}$ yields an approximation of the input evoked by an A tone. Left panel: d_0 and d_2 values' dependence on PR and Δf. Right panel: Input approximation as described above. Parameters: $U = 0.4$, $\tau_D = 0.5$ s, $TD = 20$ ms, $\mu = 0.4$ oct.

Then, by approximating the thalamocortical current at a synapse (equation (B.26)) with the product of the depression values and the input rate, R_{Th}, directly (instead of going via the synaptic gating variable of equation (B.30)) and making a crude threshold-linear 'neuron' by deducting from that current a threshold value (Θ, here in Hz), we arrive at an approximation of the response ratio RR which, depending on the value of Θ, can already show the rudiments of the data found by Fishman et al. (2004) (figure 2.3). It is of special interest to note the decisive role of the threshold in generating differential suppression. Thus, if we calculate RR without using a threshold (or, equivalently, using $\Theta = 0$ Hz), we get curves that closely follow each other across Δf with only slight differences in value (figure 2.3(a)). In contrast, once the threshold is 'activated', we see the clear tendency of differential suppression in the RR curves (figure 2.3(b); cf. figure 3 in Fishman et al. (2004), reproduced in our figure 1.4(b); $\Theta =$ 5 Hz).

In the second and third approaches, as it will turn out, the response of the neuronal population to increasing rates of input has a decisive impact on the ability of the population to model accurately the auditory streaming correlates. Specifically, the frequency of cortical response, averaged across



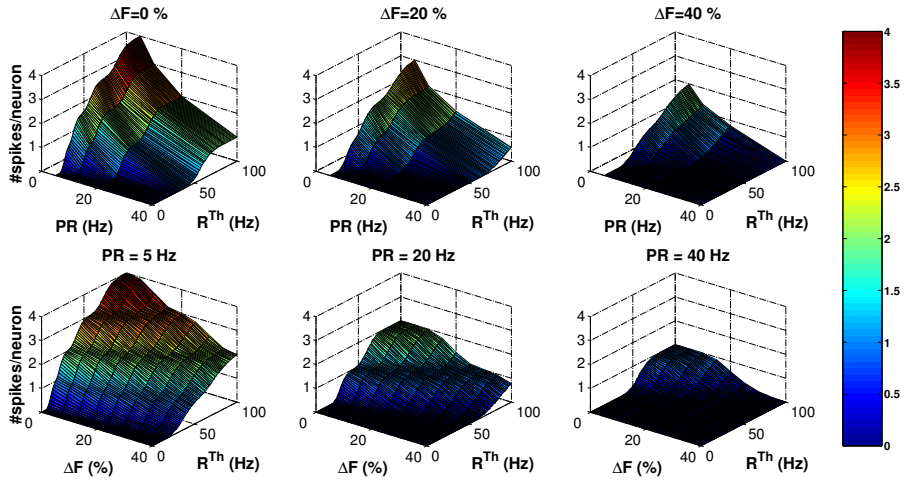
(a) With $\Theta = 0$, the shape of *RR* curves found by Fishman et al. (2004) is not discernable.



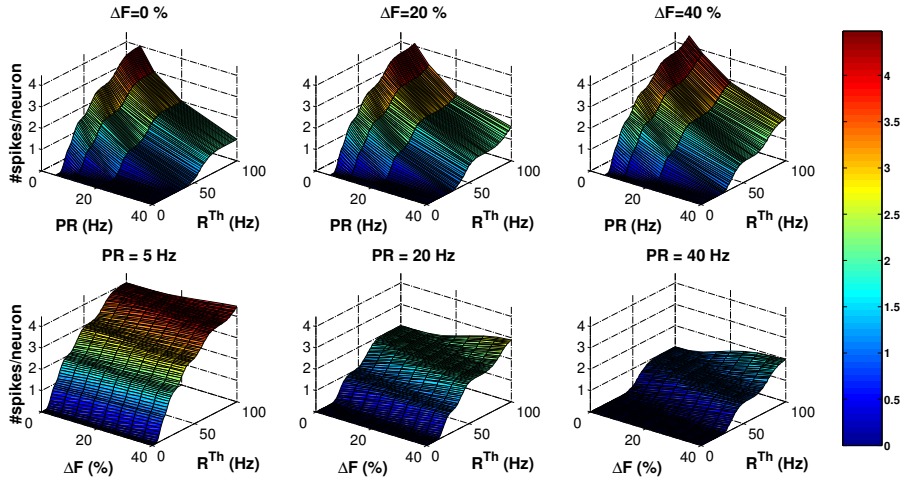
(b) With $\Theta = 5$ Hz, we already get *RR* curves reminiscent of those found in Fishman et al. (2004) (cf. figure 1.4.b).

Figure 2.3: Assessment of response ratio, *RR*, using steady state d values, input rate and an artificial threshold, Θ. Depicted is $\frac{d_2(\Delta f)\ R_{ThC}(\Delta f)-\Theta}{d_0(\Delta f)\ R_{ThC}(0)-\Theta}$ (using curves seen in figure 2.2) plotted against Δf(*PR* varies across lines). The importance of the excitability threshold for differential suppression is clearly illustrated: in (a), $\Theta = 0$ Hz and the *RR* values differ little across *PR*, whereas in (b), a small Θ of 5 Hz yields *RR* curves very similar to experimental results (Fishman et al. 2004). Other parameters as in figure 2.2.

the population, needs to be a monotonically increasing function of the input, without input intervals where the derivative of the function is zero. This is because synaptic depression, indirectly related to the experimental parameters Δf and PR, essentially 'moves' the response down the slope of the curve and such iso-response 'patches' can cause an equal response of the network to different values of those same parameters, something not seen (on average) in the data of Fishman et al. (2004). In the rate model, this constitutes no problem, as this 'transfer function' is imposed by the modeller (e.g., linear; cf. left panel of figure 2.8), but when modelling conductance-based neurons, one requires a variability of response across neurons in order to get such smoothness in this characteristic. We will now illustrate the problem and discuss necessary adjustments to parameter values to remedy it. Using the surfaces in figure 2.4, one can calculate the corresponding RR in two different ways, the first more approximate than the second. The fastest (approximate) way is to generate only the surface in figure 2.4(b), and make use of the fact that it contains the response to all possible input rates. Since different B tones are represented by different input rates, one simply picks out the response to B tones by reading out its value at that rate level. This is approximate since B tones are presented when depression is at d_2 level, not d_0, but as we shall see below, the difference is small. The second way requires both surfaces and picks the more realistic B response from the B surface of figure 2.4(a). The result of these two procedures, for the surfaces of figure 2.4, are displayed in figure 2.5(a) and figure 2.5(b), respectively. As can be seen in the figure, these two approaches lead to marginally different response ratios. However, they are obviously far from a fit of Fishman et al. (2004) data. If we now wanted to explore different parameter values to see if the fit improves, we would have to run another batch of numerous simulations to generate a new surface, which constitutes a limitation of this method. Another limitation is that for $\nu > 0$, the approximation of a B response by the response to A at the rate of B worsens. These limitations bring us to the advantages of our third approach, where we simply set d to a start value and simulate the response to a B tone. Systematically covering the range of possible d values, we thus create a surface such as that seen in figures 2.6(a) and 2.6(b). Both figures show the same surface, only the dots superimposed on the surface differ (for shape and colour coding, see caption). In (b), the responses to an A tone (i.e., B tone with $\Delta f = 0\%$) at different PR and Δf are shown, whereas in (a) the dots represent the responses to a B tone at those same experimental parameter values. The finesse here is that, while these dots are in part generated based on the data underlying the surface, we can make different projections of such dots onto the surface

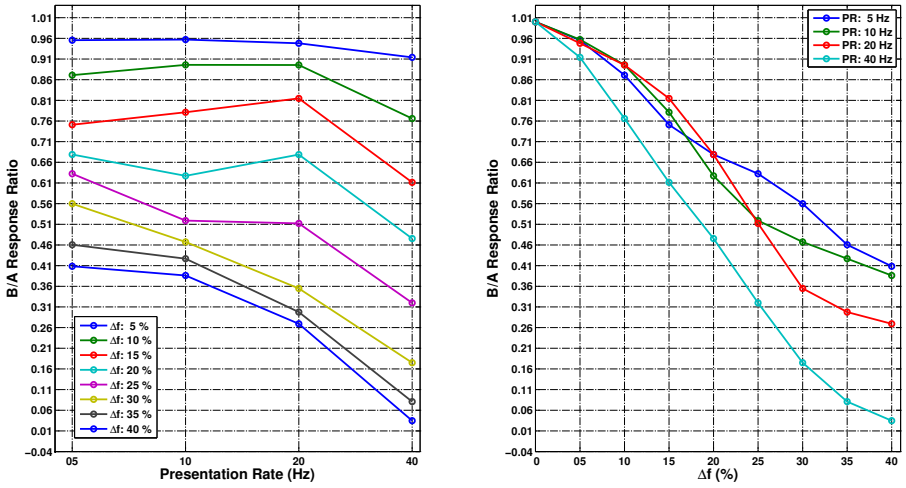


(a) B tone presented with depression set to d_2 values.

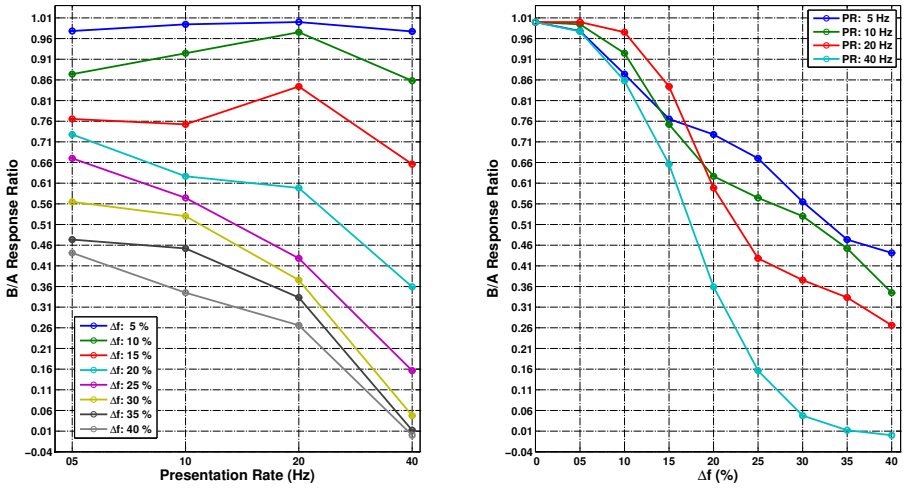


(b) A tone presented with depression set to d_0 values.

Figure 2.4: The 'transfer functions' in response to a B tone (a) and an A tone (b) of a population of $M = 100$ cortical neurons, each having $K = 100$ synapses, here with depression ($U = 0.4$, $\tau_D = 0.25$ s). The figure shows how the response of the network changes with the main experimental parameters, PR and Δf, as R_{Th} varies. Note that the inital value of depression is set to its respective steady-state values (cf. figure 2.2 and see subcaptions). Other parameters: $g_{\mathrm{CTh}} = 0.02 \pm 0.004$ mS/cm^2, $TD = 20$ ms.



(a) Response ratios calculated based on A responses from figure 2.4(b).



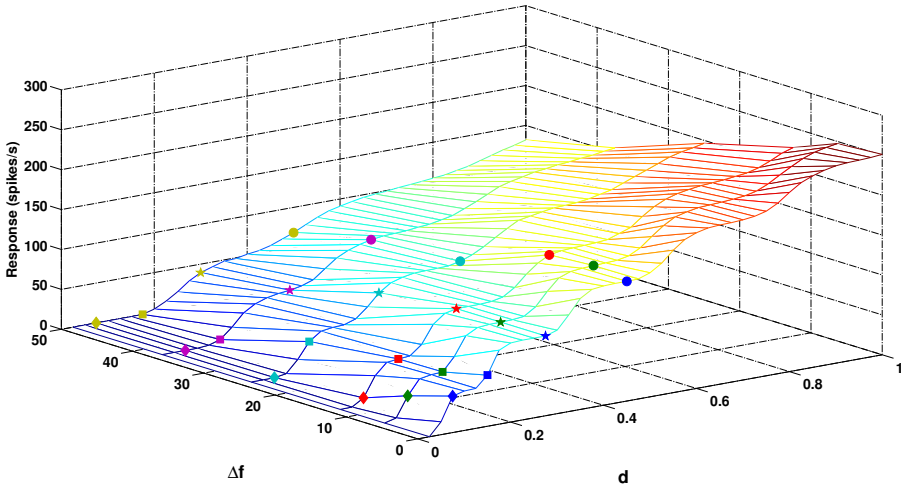
(b) Response ratios calculated based on responses to both A and B (figures 2.4(b) and 2.4(a)).

Figure 2.5: Response ratios calculated based on surfaces in figure 2.4 (see text and subcaptions).

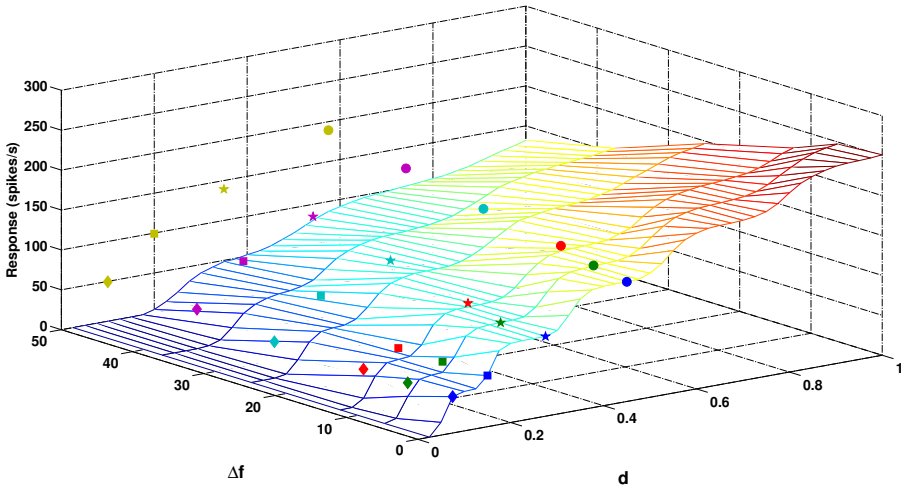
by changing U and τ_D in an analytical calculation yielding $d_0(PR, \Delta f)$ and $d_2(PR, \Delta f)$, respectively, thus rendering unnecessary further simulations in order to test other synaptic depression parameters. Specifically, the different d_0 and d_2 values thus obtained are used to place the different dots at the surface points correponding to the particular d, PR and Δf values (only limited by data mesh resolution). For the A tone, instead of placing all the dots at $\Delta f = 0$, we place them at the PR and Δf values used in the analytical calculation, for visibility reasons (but at the response rate of A ; hence, the dots can 'hover' above the surface). The only downside to this methodology is that it neglects the difference between the effect that the U value of the simulation (here: 0.4) *de facto* had during the presentation of the tone, and the effect that the U value of the analytics *would* have (τ_D has much less effect during a short tone than U). For instance, to generate the dots of figure 2.6 and subsequently calculate the response ratios of figure 2.7, we used $U = 0.4$ but changed τ_D from 250 ms (simulation) to 500 ms (analytics), which probably creates a very small error in comparison with using a surface generated with $\tau_D = 500$ ms. We were prompted to experiment with these values by the rather poor fit when the analytical and simulation depression parameters were identical (data not shown). Finally, if introducing thalamocortical axonal spread ($\nu > 0$), one would simply have to run a new simulation of a network with this feature, keeping in mind subsequently that analytical solutions for synaptic depression steady state values do not take into account such spread, increasing potential errors of this method.

Finally, in our fourth approach we formally propose a novel neuronal mechanism that produces differential suppression, based on the behavior of the synaptic depression variable that we have previously described, and then proceed to explain how this can easily be understood in the context of our rate model. So, in order to understand the differential suppression phenomenon it is useful to consider the problem as static. Indeed, it is possible to obtain a good estimation of the model's response to the time dependent input, $I_{Th}(t)$ of equation (A.3), considering only the synaptic depression variable once it has attained its steady state.

This can be understood by analyzing the simple equations governing the rate model (appendix A). The input function $I_{Th}(t)$ depends on time through the rates $R(F, t)$, and also via the synaptic depression function $d(t)$ (see equation A.3). Here however, instead of studying the effect of a time varying input (A.3) on the dynamics of the rate model (A.1), we solely consider the responses of a population of BF f_A to the constant inputs r_A and r_B.



(a) Response to a B tone at different Δf and PR values.



(b) Responses to an A tone at different Δf and PR values (projected to Δf of B tone).

Figure 2.6: Shown in both (a) and (b) is the surface which we get by simulating the response of a network to a B tone of different Δf values, with d_{start} covering the range of possible d values (B tone of $\Delta f = 0\%$ is of course an A tone). In the simulation, $U = 0.4$ and $\tau_D = 0.25$ s, while in the analytics we used the same U but set τ_D to 0.5 s. Circles, stars, squares and diamonds represent PR = 5, 10, 20 and 40 Hz, respectively, while blue, green, red, cyan, magenta and yellow markers stand for Δf = 0, 5, 10, 20, 30, 40 %, respectively.

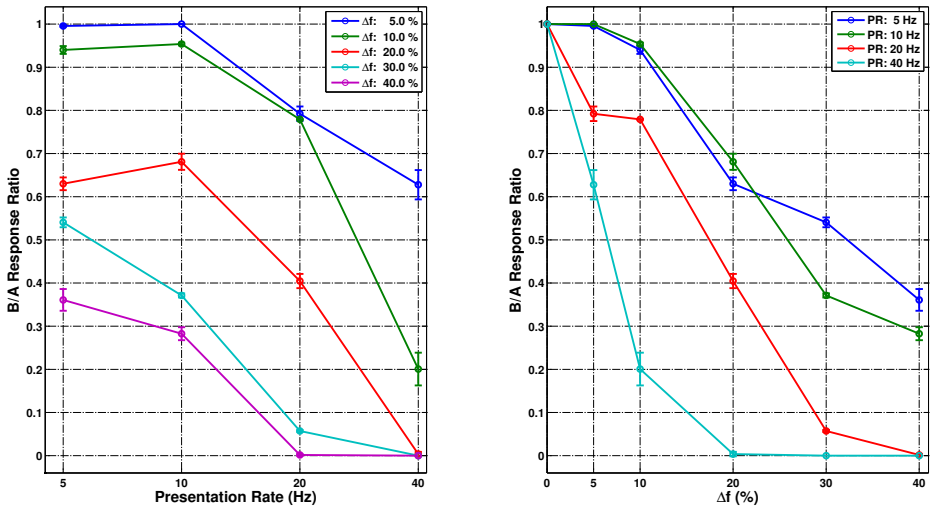


Figure 2.7: This plot shows the B/A response ratios obtained using the dots of figure 2.6(a) as B responses and the dots of figure 2.6(b) as A responses, respectively. By doubling τ_D from the value of the simulation data, we obtained a much better fit than when using $\tau_D = 250$ ms, as in the simulation.

As already mentioned, we first assume that the synaptic depression variable is in the steady-state regime. This function decays with fast kinetics during the input presentation (illustrated in figure A.2 of appendix A). However, we will assume that the value of the synaptic depression variable d is constant during the presentation of a tone. Moreover, for the inputs used in this study, we show in appendix A that we can approximate $d_A \approx d_B$, and consider a single mean value, $\bar{d} \equiv \frac{d_A + d_B}{2}$, of the synaptic depression variable, independently of the tone which is being presented.

Therefore, we are interested in studying the *fixed points* of the rate model (A.1)

$$m^{\mathrm{SS}}_{A,B}(PR, \Delta f) = [I_{A,B}(PR, \Delta f) - \theta]_+. \tag{2.8}$$

for the two types of inputs

$$I_A(PR, \Delta f) = W_0 \; \bar{d}(PR, \Delta f) \; r_A, \tag{2.9}$$

$$I_B(PR, \Delta f) = W_0 \; \bar{d}(PR, \Delta f) \; r_B, \tag{2.10}$$

where we consider the limit of infinitely narrow thalamocortical projections, reducing equation (A.4) to $W_0 \delta(\varphi - F)$ (using $\lim_{\lambda \to 0} e^{-|x|/\lambda}/\lambda =$

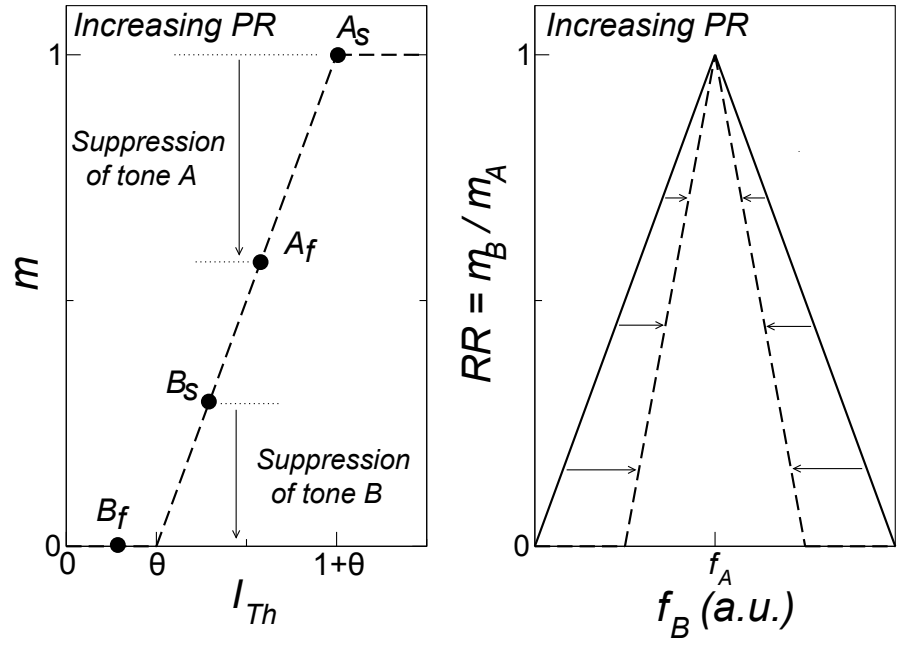


Figure 2.8: Scheme describing the proposed mechanism that produces differential suppression of best frequency (BF) and non-BF tones. Left Panel: Transfer function (dashed line) of a population of BF f_A as a function of the thalamic input I_{Th}. The circles represent the fixed points of the population's rate equation (A.1). When a tone A is presented with slow PR (fixed point A_s), the population's response is maximal. A non-BF tone B at low PR (fixed point B_s) elicits less response, that depends on the absolute value of the tones' frequency difference Δf. Increasing PR produces a decrease of I_{Th}, due to synaptic depression, that reduces both A and B population responses (fixed points A_f and B_f, respectively). The excitability threshold of the φ_A population, θ, produces the total suppression of the B responses at fast enough PR. Right Panel: Schematic representation of the response ratio (2.11) as a function of the frequency of the B tone. At low PR the response ratio is broad. Increasing PR produces the narrowing of the population's receptive fields (see text).

$2\delta(x)$), i.e., a synaptic input strength of W_0 at φ_A (through integration in equation (A.3)). It is important to stress that, in general, the fixed points of the rate equation (A.1) with fast-varying inputs (compared with the time scale τ of the rate equation) will not provide valuable information about the actual system's dynamics. However, in our case the study of the fixed points (2.8) provides important information that permits us to understand the mechanism producing differential suppression.

Figure 2.8 represents schematically the transfer function of a neural population with best frequency f_A , and four fixed points. Two of them represent BF-tones, that are being presented either in a slow sequence (fixed point A_s) or in a fast sequence (fixed point A_f). The tones presented at low presentation rates (PR) produce less thalamocortical synaptic depression (i.e., higher values of $\bar{d}$), and therefore the response of the neuronal population at φ_A will be high. Increasing PR will result in a reduction of the population's response due to the decrease of the mean synaptic depression function, $\bar{d}$ (see figure A.3).

The other two fixed points correspond to non-BF tones. Indeed, at slow presentation rates (fixed point B_s), the response is lower than that elicited by an A tone, due to the frequency difference $|\Delta f|$. Eventually, at fast enough PR , the input evoked by non-BF tones will not be strong enough to reach the excitability threshold of the population of neurons with best frequency f_A, causing the response to the B tones to be suppressed (fixed point B_f).

Therefore, it is possible to have a suppressive effect on responses to non-BF tones at fast presentation rates, due to the depressing nature of the thalamocortical synapses and the thresholding property of A1 neurons. As we discussed previously, Fishman et al. (2004) used the response ratio of the A and B responses in order to measure the differential suppression. In our case, the relative response of the B tones with respect to the best frequency tones (i.e. the A tones) can be simply obtained by calculating the quotient of the fixed points (2.8) as

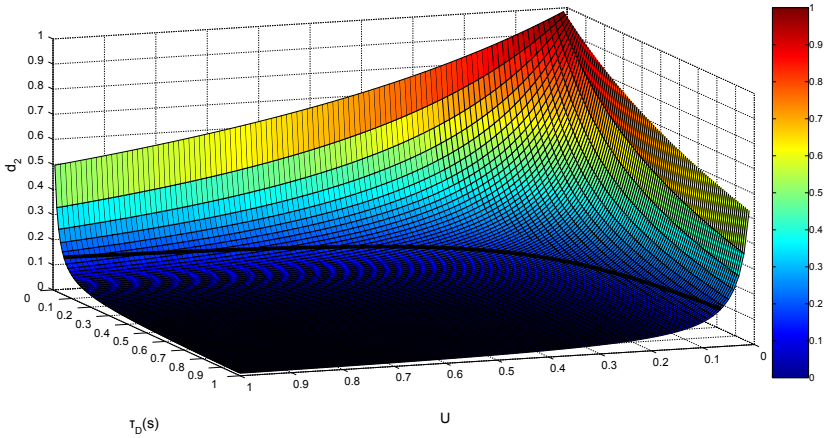
$$RR(PR, \Delta f) \equiv \frac{m_B^{\mathrm{SS}}}{m_A^{\mathrm{SS}}} = \frac{[W_0\ \bar{d}\ r_B - \theta]_+}{[W_0\ \bar{d}\ r_A - \theta]_+} \tag{2.11}$$

This expression confirms the sensibility of our approach when making an approximation of this ratio using only depression level and input rate (figures 2.2 and 2.3).

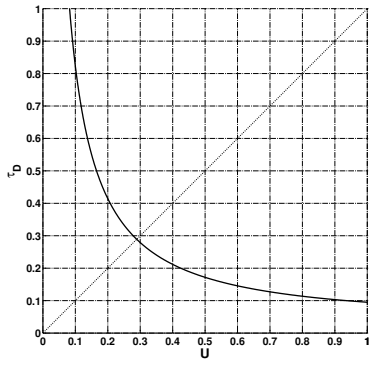
2.3.2.1. The importance of the synaptic depression parameters U and τ_D

As it was obvious from our preliminary simulations that the values of the parameters U and τ_D have a direct impact on the ability of our models to reproduce the neurophysiological data, we decided to perform an analysis with the objective of better guidance when setting those values. The initial idea was that one could be oriented by a crucial characteristic of the data, such as the Δf value at which the suprathreshold response to a B tone at φ_A first becomes (practically) zero, for $PR = 40$ Hz (between $\Delta f =$ 10% and 20%). Since for a certain network setup one reaches this point for a specific steady-state value of d, one could thus extract that d_2 (for B tones) value from a simulation with the setup of interest and then study which values of U and τ_D would result in that d_2 value, using our analytical solutions (appendix A). For instance, to generate figure 2.9(a), which shows an iso-d_2-value line (black; for $d_2 = 0.15$) superimposed on a surface of steady-state d_2 values as a function of τ_D and U, we first deduced that at $d_2 = 0.15$ we *just about* get a spike response for $PR = 40$ Hz and $\Delta f =$ 10% in a feed-forward network simulation using $TD = 20$ ms, $U = 0.4$, $\tau_D =$ 0.3 s. Then, we used the expression (A.27) to calculate the surface of the figure. Crucially, any point on the iso-line (re-plotted for better visibility in figure 2.9(b)) would yield $d_2 = 0.15$, which means there is a relative freedom in the choice of U and τ_D values, as long as their interrelationship (as defined by the line; in figure 2.9(b) U and τ_D are almost mirrored on the diagonal, for example) is respected. This affords us the possibility of picking U and τ_D from along such an isoline. However, this analysis is only valid for the steady-state, since the dynamical trajectory of the d variable varies with U and τ_D, but for our purposes in modelling streaming correlates it is useful.

We emphasize here the importance of this result, as it provides a tentative explanation for two things; firstly, it might explain why the rather widely distributed values (see Introduction) of these quantities, i.e., utilization of synaptic resources and their rate of recuperation, respectively, when measured in the literature, should constitute no 'problem' *per se* for neural processing outcome; secondly, it explains why we find streaming correlate solutions with different neural network models (e.g., rate model and HH model), that differ in the values of these two parameters — the specifics of each model as to input and processing can condition the ideal choice of those values (present chapter and chapter 3).



(a) 3D plot showing $d_2(U,\tau_D)$. The black line shows all possible combinations of U and τ_D which yield d_2=0.15.



(b) The black line from (a) plotted in 2D for better visibility. Note the near symmetry with respect to the diagonal.

Figure 2.9: Figure (a) shows an iso-d_2-value line (black; for $d_2 = 0.15$) superimposed on a surface of steady-state d_2 values as a function of τ_D and U, while figure (b) shows that same iso-d_2-value plotted in 2D. Crucially, any point on this line would yield $d_2 = 0.15$, which means there is a relative freedom in the choice of U and τ_D values, as long as their interrelationship (as defined by the line) is respected. This affords us the possibility of picking a d_2 value for which the response to a B tone at φ_A should be zero (for, say, $\Delta f = 20$ %, $PR = 40$ Hz) and then performing this analysis in order to set U and τ_D, picked from along such an isoline. Note how U and τ_D values are almost interchangeable (isoline almost symmetric with respect to diagonal).

2.3.3. The role of the input function

It is important to realize that, when modelling the response to tones, the spectral shape we choose for our input function is of the utmost importance. As this is something others have noted (see, e.g., de la Rocha et al. 2008), it is no surprise to us to find that the same is true in our case. Specifically, when modelling AS, we experimented with a certain spread in thalamocortical projections, our parameter ν. We found that when we had $\nu = 0$, modelling the effect of *RR* falling off with Δf was easier than when using $\nu > 0$. This was especially true at $\Delta f \leq 10$ %. To illustrate the problem, we have prepared figure 2.10, which shows normalized input in four different cases, along with the approximated percentages of maximum input arriving at φ_A in cortex (centered on $\Delta F = 0$), integrated across sending axons. Above left we have the simplest case, the exponential input function with $\mu = 0.4$ oct. Above right, we have the same input function, but here the percentages have changed according to our implementation of thalamocortical spread ($\nu = 0.33$ oct.). The latter illustrate the problem clearly, as the difference in total input to φ_A diminishes when ν is nonzero. While this is normal and expected, in terms of our modelling objectives it created the need to be explain how, e.g., Fishman et al. (2004) could show differences in response at such small Δf. While this could be due to averaging enough trials (they use 75, we typically less), we decided to also explore other explanations. The first was to include a modulation of g_{CTh} with ΔF. The lower right and lower left panels of figure 2.10 show the case of exponential and Gaussian modulation, respectively (width ν); clearly, as the corresponding percentages show, this remedies somewhat the problem, but for $\Delta f = 5$ %, we are still very close to the input at $\Delta f = 0$, especially for Gaussian modulation. To give an idea of how the input changes with the modulation, we have plotted the modulation function (pink line) and the result of its convolution with the original input function for each Δf (other colours - this is purely illustrative, as rate and conductance are never actually directly multiplied). We will discuss other approaches to solving this dilemma in chapter 3, notably using intracortical recurrent excitation.

2.3.4. Rate model results

2.3.4.1. Model's response to sequences of a single tone

Figures 2.11a and 2.11b show the firing rate of an A1 model neuron (with best frequency f_A), elicited by a sequence of pure tones of a single best frequency (BF) (panel a) and a single non-BF (panel b), respectively.

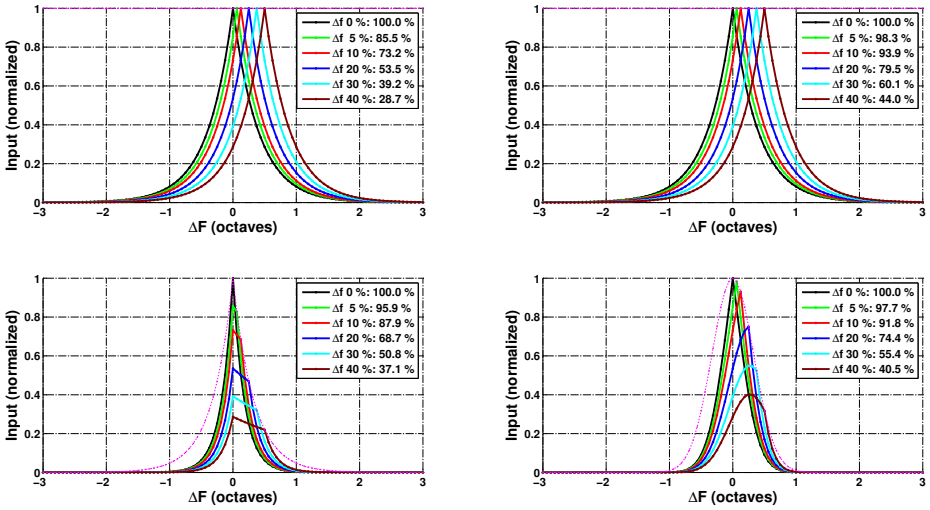


Figure 2.10: The figure shows normalized input (for a Δf of 0, 5, 10, 20, 30 and 40 %, respectively) in four different cases, along with the approximated percentages of maximum input arriving at φ_A in cortex (centered on ΔF= 0), integrated across sending axons. Above Left: the exp. input function with μ = 0.4 oct, ν = 0.0 oct. Above Right: the exp. input function with μ = 0.4 oct, ν = 0.33 oct. Lower Left: as in Upper Right, but with exponential modulation of g_{CTh} across tonotopic space (width ν = 0.33 oct.). Lower Right: as in Upper Right, but with Gaussian modulation of g_{CTh} across tonotopic space (width ν = 0.33 oct.). To give an idea of how the input changes with the modulation, we have plotted the modulation function (pink line) and the result of its convolution with the original input function for each Δf (purely illustrative, as rate and conductance are never actually directly multiplied). Note how, when $\nu > 0$ (Upper Right), percentages change, diminishing differences in total input across Δf values). We can remedy this in part by also modulating g_{CTh} across the TT axis, but for $\Delta f = 5$ %, we are still very close to the input at $\Delta f = 0$, which means the cortex probably 'solves' this problem using intracortical mechanisms such as recurrent excitation.

Additionally, the same panels show the instantaneous fraction of available neurotransmitters (synaptic depression variable d), averaged across the ThC synapses at $F = f_A$.

The A1 neuron stimulated at its BF recieves strong excitatory input and thus the ThC synapses at $F = f_A$ become strongly depressed. In contrast, non-BF tones produce less synaptic depression at this location. The total excitatory input to the A1 neuron is proportional to d (see equation (A.3)) and thus the presence of synaptic depression degrades the neuron's frequency selectivity (Abbott et al. 1997). Indeed, figure 2.12a shows the broadening of the frequency tuning curve of the A1 neuron when ThC synaptic depression is included in the model.

Additionally, note that in the presence of synaptic depression the tuning curves become sensitive to the presentation rate, since the higher the presentation rate, the more ThC synapses are depressed. Specifically, figure 2.12a also shows how tuning curves broaden around its maximum as PR is increased, reducing the frequency selectivity of the neuron.

2.3.4.2. The streaming paradigm: Differential suppression

Figure 2.11c shows the response of the model neuron to a series of tones alternating in frequency, the auditory streaming stimulus. The sequence is composed of a BF tone ($f = f_A$), and a non-BF tone B (of the same frequency as the B tone of figure 2.11b). The presentation rate is $PR = 20$ Hz, as in figures (2.11a,b). Note that in the present case there is a strong improvement of the neuron's frequency selectivity, if the neuron's responses to A and B tones (m_A and m_B) are compared *within* the sequence.

In contrast with figures 2.11a and 2.11b, now the neuron's response to the A tones increases, whereas the response to B tones (non-BF) is strongly suppressed.

Additionally, as we discussed previously, an increase of PR produces a decrease of the amount of neurotransmitters in the synapse $F = f_A$, and thus a reduction of the neuron's response to both BF and non-BF tones. Eventually, at some PR, the excitatory input produced by non-BF tones is not strong enough to reach the neuron's excitability threshold. Panel 2.11d shows that the neuron's response to the B tones has been totally suppressed at $PR = 30$ Hz, and thus the neuron becomes only selective to A tones.

The response ratio

$$RR = m_B/m_A \tag{2.12}$$

is considered to be a neural correlate of streaming (Fishman et al. 2001, Fishman et al. 2004, Micheyl et al. 2005). Figure 2.12b shows the quantity

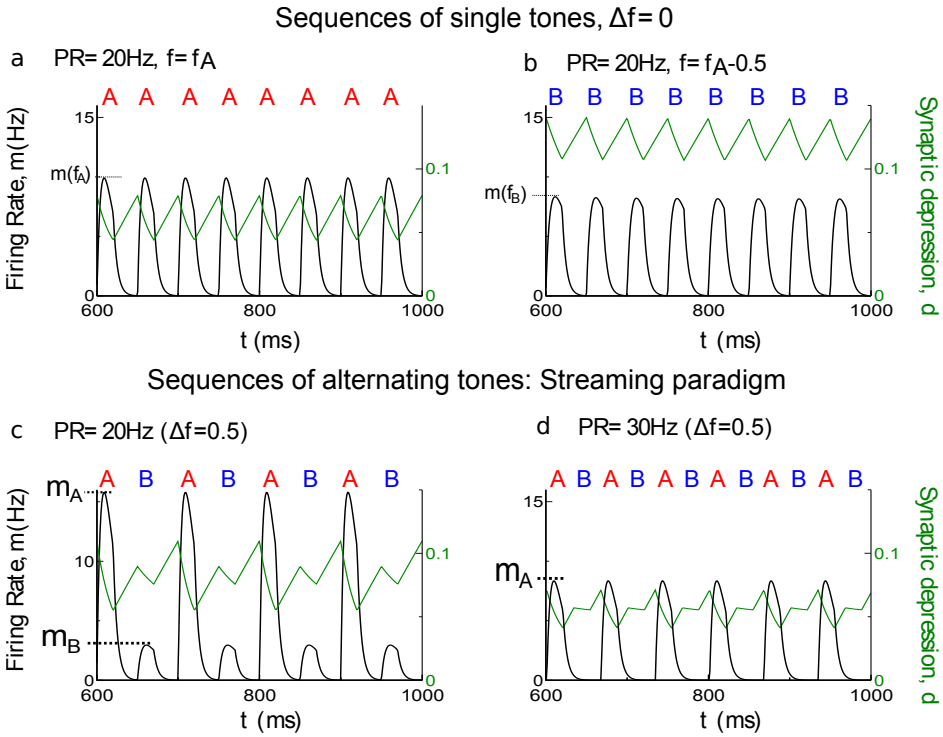


Figure 2.11: Time series of rate model's response to sequences of single (**a** and **b**) or alternating (**c** and **d**) tones. PR and Δf (octaves) given in figure, $U = 0.5$, $\tau_D = 800$ ms, rest of parameters standard. Note that 0.5 octaves corresponds approximately to 40% with respect to a tone's frequency in Hz.

(2.12) versus Δf. This can be interpreted as a generalized neuron's tuning curve for the streaming stimuli. The different curves correspond to the response ratios at different PR values. The curves clearly show an improvement of frequency selectivity as the stimulus is presented faster, in very good agreement with the experimental results of Fishman et al. (2004).

This mechanism of frequency selectivity is entirely due to the dynamic nature of synaptic transmission (henceforth, *dynamic frequency selectivity*), and requires PR to be fast compared with the characteristic recovery time constant of synaptic depression. Note that the streaming stimulus has $\frac{1}{PR} \sim [25, 200]$ ms, and synaptic depression has recovery time scales of hundreds of milliseconds (here we considered $\tau_D = 800$ ms) (Tsodyks, Pawelzik & Markram 1998, Abbott et al. 1997, Tsodyks & Markram 1997). Under these conditions, the amount of available neurotransmitter vesicles

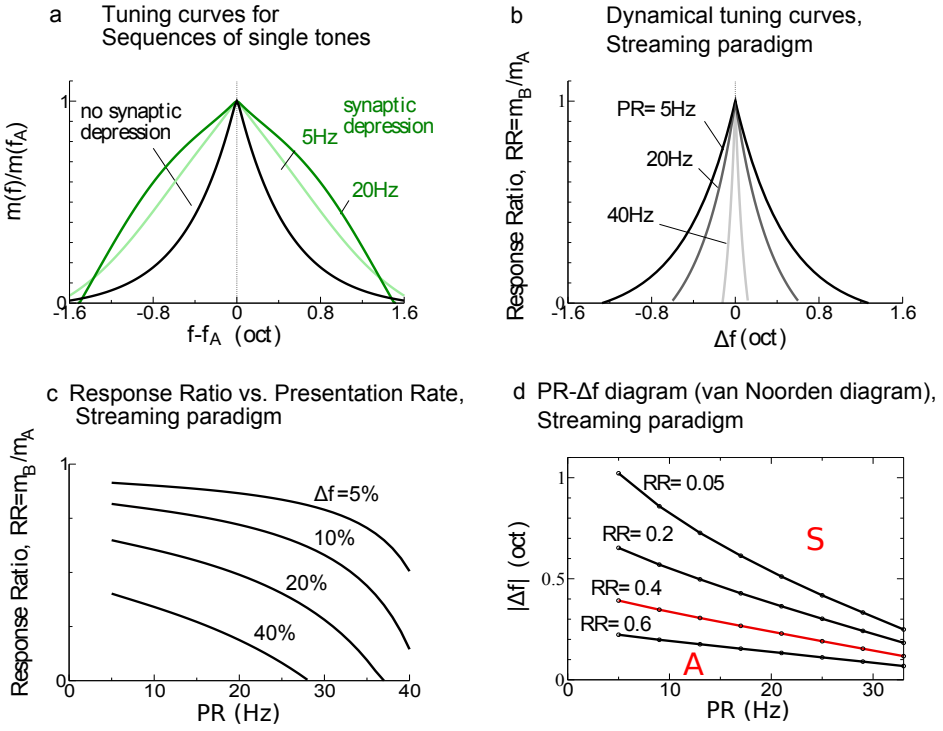


Figure 2.12: Synaptic depression degrades frequency tuning (see panel a) but improves dynamic frequency tuning (a.k.a. differential suppression; see panel b). In panel c, we see the response ratios plotted versus PR, with Δf changing across lines. Finally, panel d shows a contour plot of RR as a function of PR and Δf. Parameters as in figure 2.11.

when a BF tone is presented is similar to the amount available for non-BF tones. Therefore, in this case the presence of synaptic depression does not degrade the frequency tuning of A1 neurons, as occurs with stimuli composed of tones of only one frequency (figure 2.12a).

2.3.4.3. Model's transitory regime: Physiological forward masking and build-up of streaming

So far we have studied the model in its steady state. However, this steady state takes (with our parameters) some hundreds of milliseconds to be reached, as can be appreciated in figure 2.11c (notice time scale; the transient part of the underlying synaptic depression is illustrated in figure A.2 of appendix A). Once the steady state has been established, the

system shows the differential suppression phenomenon we have previously described. However, during the transitory regime the system shows a different phenomenology.

Specifically, considering only the two initial tones, the forward masking paradigm is recovered (Calford & Semple 1995, Brosch & Schreiner 1997, Wehr & Zador 2005). In this experimental paradigm a 'masker' tone of various frequencies and intensities is presented. After a certain time, the neuron's response to a BF tone (called the 'probe' tone) is measured.

However, as pointed out by Fishman et al. (2004), there are important differences between the stimulus conditions of the physiological forward masking studies and those focussing on auditory streaming. Indeed, the auditory streaming stimuli produce a transient response of the A1 neurons that, after at most a few seconds, is followed by a steady state regime, in which the time series of the neural activity is almost periodic. In contrast, in the forward masking paradigm tone pairs are presented in isolation after long periods of silence.

Figure 2.13 shows the recovery time of the response of a neural population with BF f_P to a probe tone of that same frequency, after the presentation of a masker tone of frequency f_M, as a function of $\Delta f = f_P - f_M$ (f_M varies). The shape and the time scale of the recovery time curves resemble the classical forward masking results found by Calford & Semple (1995) and Brosch & Schreiner (1997), indicating that thalamocortical synaptic depression may play an important role in explaining that data too (as found earlier by Denham 2001).

In the streaming paradigm, at the beginning of the stimulus, the sequence is generally perceived as an integrated percept. However, after up to a few seconds, this 'splits' into the two streams percept (Bregman 1978). This effect is known as the 'build-up' of auditory streaming, and its neurophysiological basis was recently studied by comparing the time evolution of the neural responses in A1 of monkeys with the reports of humans using the same stimuli (Micheyl et al. 2005). The study showed a similar trend of the A1 neural responses, on the one hand, and the build-up of streaming in human subjects, on the other.

2.3.4.4. Effects of context and memory

The fact that the streaming percept takes several seconds to build up suggests that this phenomenon is associated with a slow integration process. Therefore, auditory streaming is affected by the preceding context, so that the auditory events prior to the repeating AB sequence might influence the likelihood of perceiving streaming.

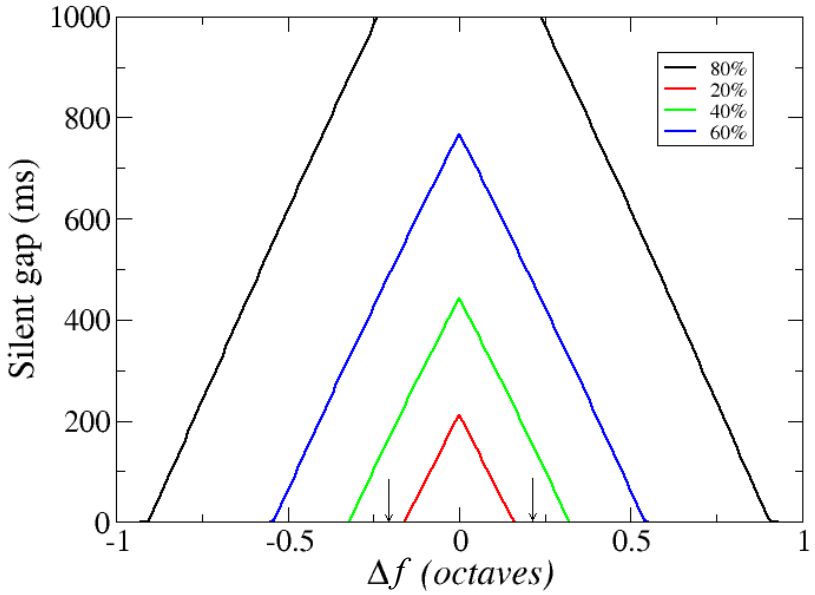


Figure 2.13: Time needed for the probe tone response to recover 20, 40, 60 and 80% of its maximum level (after being masked by a tone of frequency f_M), as a function of the frequency difference Δf between the masker and the probe tones (only the masker frequency f_M is varied). The curves have been obtained using equation (A.16). The region between the arrows indicate the values of Δf for which the neural model responds to the Masker tone. Note that there exists masking even if there is no response to the masker tone, i.e. when $|\Delta f| > 0.2$ oct. Parameters: $d_{start} = 1.0$, $U = 0.2$, $\tau_D = 800$ ms, $TD = 20$ ms, $\mu = 0.5$ oct., $W_0 = 0.35$, $\theta = 0.2$.

In particular, Cusack, Deeks, Aikman & Carlyon (2004) studied the effect of introducing a silent gap of various durations (1, 2, 5 and 10 s) in the middle of a streaming sequence that was perceived as segregated, in order to study the time needed to reset the streaming percept. Their psychophysical results showed that the one second silent gap resets the streaming percept, and mainly the integrated percept is heard. This trend did not show any significant difference for larger durations of the silent gap (2, 5 or 10 s) (Cusack et al. 2004).

The lower panel of figure 2.14 shows the model's response to the A and B tones as a function of time. At the beginning of the sequence, the system

has a transient regime of about one second, that strongly resembles the data found by Micheyl and collaborators (Micheyl et al. 2005, Micheyl et al. 2007). The upper panel shows the response ratios B/A, as a function of time.

Once the steady state regime has been reached, silent gaps of various durations are introduced. Note that the model is very sensitive to the temporal characteristics of the stimulus due to the slow time scales of the thalamocortical synaptic depression. This provides a simple mechanism of short memory (1-2 seconds), which is stored in the synaptic depression field (illustrated in figure 3.18, chapter 3) and which makes the current responses depend on the previous sound pattern, i.e., on the sound context.

After the silent gaps the sequence is presented again. The upper panel in figure 2.14 shows the progressive increase of the *RR* magnitude, that is correlated with the perception of streaming. Therefore, the model predicts a progressive reset of the streaming percept, that corresponds to low *RR* values. A complete reset of the percept occurs after 1-2 seconds of silence, in correspondence with the psychophysical data presented by Cusack et al. (2004). For a further illustration of this, we refer the reader to our scientific movies, based on simulations with our tonotopic HH model introduced in the next chapter (URL and description in section 3.3.4.2).

2.3.5. Hodgkin and Huxley model, approach one: Deterministic depression

The basic settings of parameter values for this network was described in section 2.2.3. For details, we refer to appendix B.
As explained earlier, in the feed-forward network there is no intracortical connectivity and no (active) inhibitory neurons. Given the lack of inhibition, the mechanism that can achieve the overall reduction of response in Fishman et al. (2004), as *PR* increases, is the synaptic depression. The left panel of figure 2.2 depicts the steady-state values of the synaptic depression variable, $d(t)$ as a function of both Δf and PR. The steady-state values of $d(t)$ can be used as initial values to save computing time for the feed-forward network, as also explained earlier.

2.3.5.1. The streaming paradigm: Differential suppression

In figure 2.15, we see the response of our feed-forward network once it has reached the steady-state (to save simulation time, the synaptic depression variables, $d(t)$, of the network were initialized at time $t = 0$ ms with the d_0 values of figure 2.2). Comparing with figure 3 in Fishman et al.

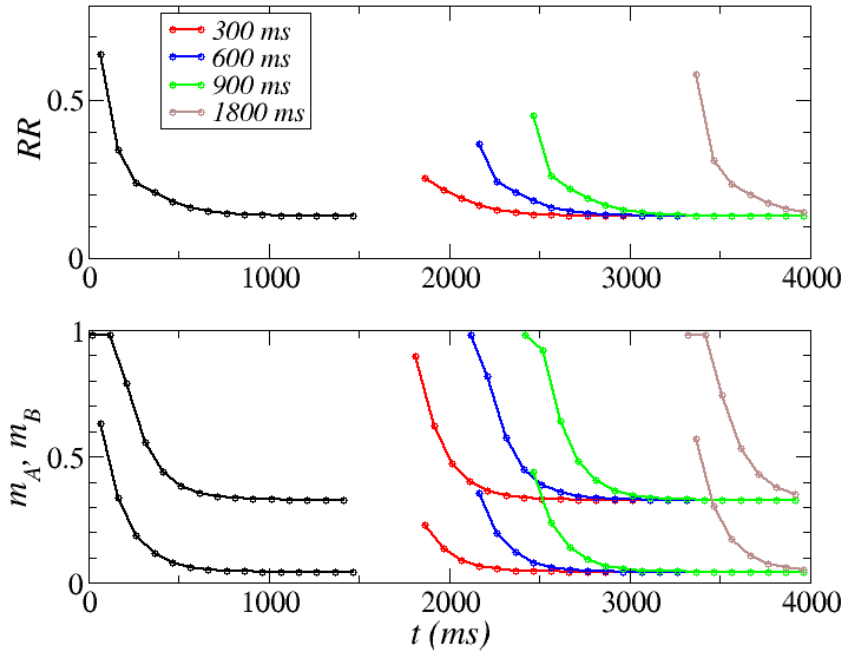


Figure 2.14: Initial transient and the effect of resets on the response ratios (parameters: $PR = 20$ Hz, $\Delta f = 0.4$ oct., $U = 0.2$, $\theta = 0.2$, $\mu = 0.5$ oct., $W_0 = 3.1$, $\tau = 5$ ms).

(2004) (reproduced in our figure 1.3), it is clear that the degree of accuracy achieved with a feed-forward network is impressive.

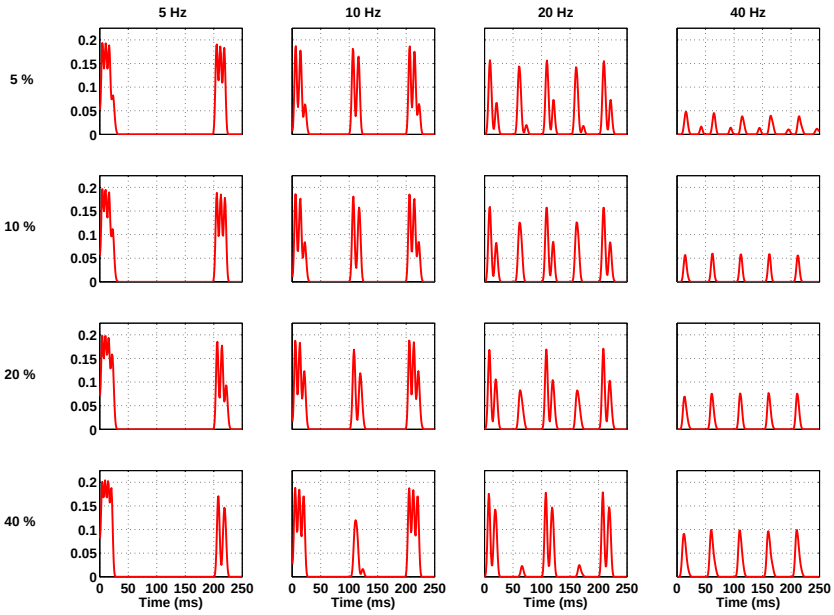


Figure 2.15: Figure shows a 1 ms bin histogram of average #spikes per neuron in response to an AB tone sequence (smoothed by a Gaussian filter) for different combinations of the two main parameters, Δf (ordinate) and PR (abscissa). Results without intracortical connectivity already show a good fit with experimental data (cf. fig. 3 in Fishman et al. (2004), reproduced in our figure 1.3). Here, $N = 1$, $M = 100$, $K = 100$, $g_{\mathrm{CTh}} = 0.02$, $\sigma_{g_{\mathrm{CTh}}} = 0.004$, $\mu = 0.4$ oct., $U = 0.4$ and $\tau_D = 0.5$ s (notably the same value we found in our analysis explained in figures 2.6 and 2.7).

Figure 2.16 depicts the B/A response ratios, as calculated using responses at the site φ_A (please see figure caption for comments).

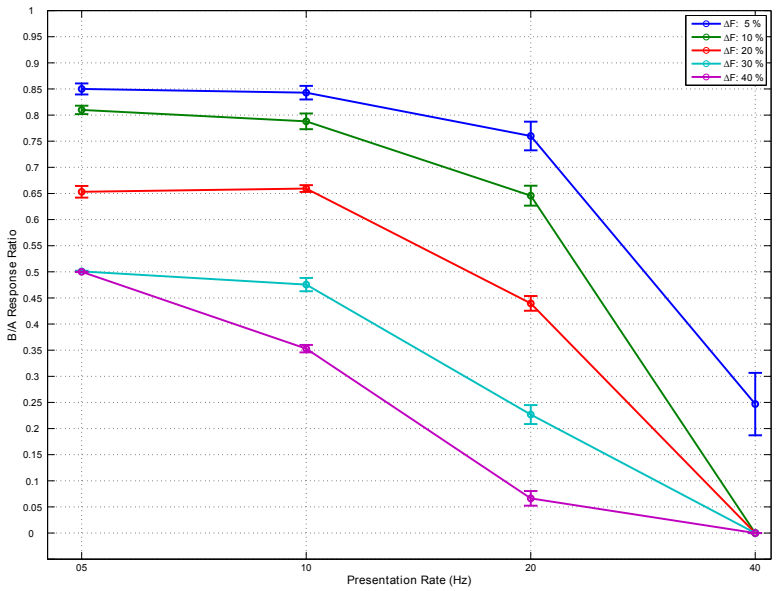


Figure 2.16: Figure depicts average B/A response ratios for the feed-forward network, as a function of PR (abscissa). Δf varies across plotted lines. Results without recurrent connectivity already show a good fit with experimental data (cf. left column of fig. 6(a) in Fishman et al. (2004), reproduced in our figure 1.4(b)). Parameter values as in figure 2.15. Average calculated over results of simulations with nine different random seeds. Error bars show standard deviation.

In figure 2.17, we may observe how the so-called build-up phase gives way to the steady state, during a total of 1.5 seconds of tone sequence presentation. The figure depicts 1 ms bin histograms, as in figure 2.15 (parameter values also as in that figure; histogram not smoothened by Gaussian filter). In fact, figure 2.15 essentially shows the last part of figure 2.17, in a blow-up.

We will now briefly discuss the effect of changes in parameters on our RR fit (both based on thought experiments and simulations; data not shown). First of all, increasing depression (higher U) lowers responses to A and B tones almost equally, but the response to B tones is lowered

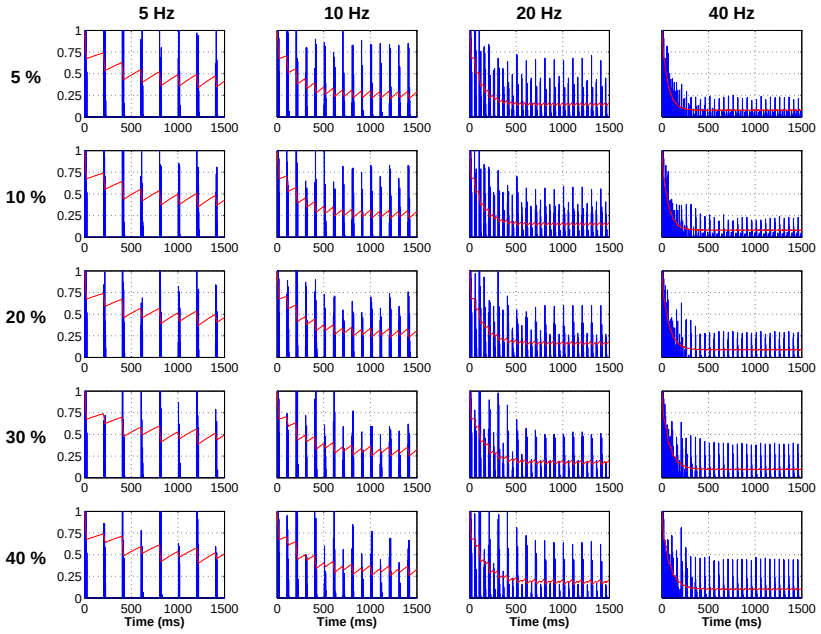


Figure 2.17: Figure shows a 1 ms bin histogram of average #spikes per neuron (normalized) in response to an AB tone sequence, for different combinations of the two main parameters, Δf (ordinate) and PR (abscissa). The build-up phase is included to illustrate how the steady-state is formed, both for neural activity and for the underlying synaptic resources (d) (thin red lines). Parameter values as in figure 2.15.

slightly less since the higher rate of the A tone has a greater impact on the value of d. Hence, RR is raised slightly by increasing U. Secondly, narrower input (lower μ) would cause B tone rates to decrease faster with Δf, which lowers RR. An added effect of depression is that lower B rates lead to less depression occurring during B tones, which in turn leads to $d(t)$ being higher right before an A tone and average steady-state d value being higher. This in turn lowers RR more. Thus, a narrower input yields a double lowering effect on RR. Finally, adding intracortical excitatory to excitatory connections would amplify differences in B and A tone responses, thus lowering RR . To counter this, one may add intracortical inhibitory to excitatory connections (as is in fact done in chapter 3, in the course of our fitting of neurophysiological data on conductances in response to short

tones).

In closing, we see that using a population of a hundred excitatory neurons with a certain variability in their intrinsic and synaptic parameter values (here only synaptic strength, as we use deterministic depression with U and τ_D values the same across synapses), we have no problem reproducing the data on the DS phenomenon shown by Fishman et al. (2004). Now we move on to approach two of our HH modelling, where we will see that the added variability of using stochastic synapses makes it possible to use a smaller number of neurons in the population (albeit averaging over a longer time interval when calculating the response ratio).

2.3.6. Hodgkin and Huxley model, approach two: Stochastic depression

2.3.6.1. Setting up the network

As mentioned in section 2.2.3, we first set out to find the parameter values we needed in our HH model in order to reproduce data in de la Rocha & Parga (2005), specifically those authors‘ comparison of the effects of stochastic and deterministic depression on response to a continuous input. To this end, we need to clarify some concepts. What is referred to by de la Rocha & Parga (2005) as the number of neurons of the pre-synaptic population, in our case translates to the number of thalamocortical synapses, K. Further, they state the strength of synapses in terms of mV (their J/C_m), whereas we give stochastic synaptic strengths in nC/cm^2 ($Q = U \times C = (J/C_m) \times c_m$, i.e., [mV]×[μF/cm^2]). Thus, integrated over time, the charges apported from each functional contact become currents. To distinguish this synaptic strength quantity from that when we use conductances, we will refer to it as q_{CTh}.

The procedure we followed to arrive at the final parameter values suitable for our model can be summarized thus:

- The first q_{CTh} values were 0.25 and 1.25 nC/cm^2, based on values for simulations underlying figure 3 in de la Rocha & Parga (2005).
- We then increased those values by the factor $1.613 \approx 1.6$, which was the quotient between our intrinsic threshold value and their Θ (we measured -55 - (-70) = 15 mV, whereas they used 9.3 mV). This yielded the values 0.4 and 2.0 nC/cm^2, respectively.
- Finally, we decreased the deterministic value to 1.5, because we found that with 2.0 we exceeded threshold too often, whereas with 1.25 we

were too far into the Fluctuation-Driven Regime (FDR), so that the standard deviation of the current didn‘t cause enough spikes.

- All other parameters were also set to their values in de la Rocha & Parga (2005), but their values did not need to be changed.

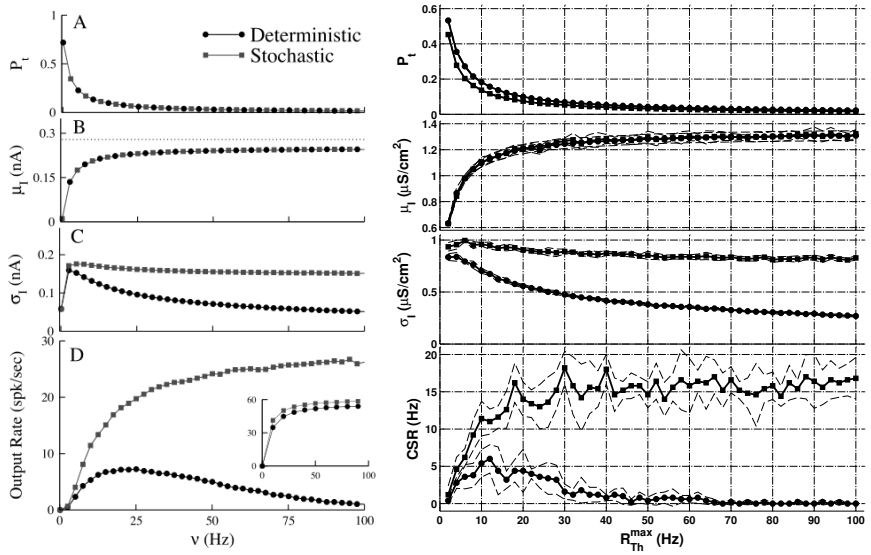
Thus, we set up a network with the parameters shown in table 2.1.

Name	Value	
$R_{\mathrm{Th}}^{\max}$	2-100	Hz
U	0.75	
τ_D	600	ms
q_{CTh}	1.5/0.4	$\mathrm{nC/cm^2}$ (D/S)
$\sigma_{q_{\mathrm{CTh}}}$	$0.4q_{\mathrm{CTh}}$	$\mathrm{nC/cm^2}$
K	400	
M	5	
N	1	
M_s	5	(S)

Table 2.1: Values of biophysical parameters for HH model, approach two: figure 2.18(b).

After this minor tweaking of parameters, we could reproduce figure 3 from de la Rocha & Parga (2005). Figure 2.18 shows the original (a) and the fit using our HH implementation (b) with the parameters of table 2.1.

While it is not surprising that we could reproduce the data of de la Rocha & Parga (2005) (apart from being more elaborate, the HH model should react similarly to similar input as the leaky Integrate-and-Fire model does, except exhibiting a higher ISI CV, in the absence of inhibition (Brown, Feng & Feerick 1999)), this step gave us the necessary groundwork needed to explore the response of such a network to sequences of tones of a single frequency (reproducing results of de la Rocha & Parga (2008), but in an auditory context) and, ultimately, the response to sequences of tones with two alternating frequencies. We worked with the hypothesis that with a minimum change of parameter values, not only should our HH model be able to reproduce the results mentioned, but it should be possible to get the approximate *RR* curves of Fishman et al. (2004), by simply changing the input stimulus to sequences of tones of alternating frequency.



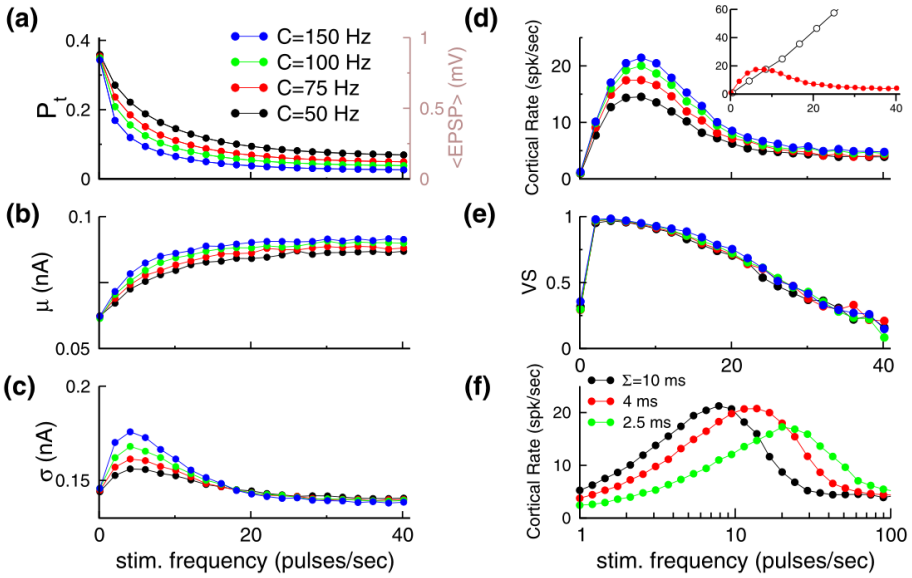
(a) Original figure (b) Reproduced figure

Figure 2.18: Original figure (adapted from de la Rocha & Parga 2005) in (a), our fit in (b). Both figures' panels show (from above to below): Transmission probability, current mean (μ_I) and s.d. (σ_I), Cortical Spike Rate (CSR) - equations (2.3) - (2.7), averaged across neurons ('trials' in data of (b)). Line with squares: stochastic depr. Line with circles: deterministic depr. Note how the greater current s.d. yields a higher CSR with stochastic depr., whereas both drop off towards zero with increasing input rates with deterministic depr. T_1=0 s, T_2=2 s, other parameters as in table 2.1.

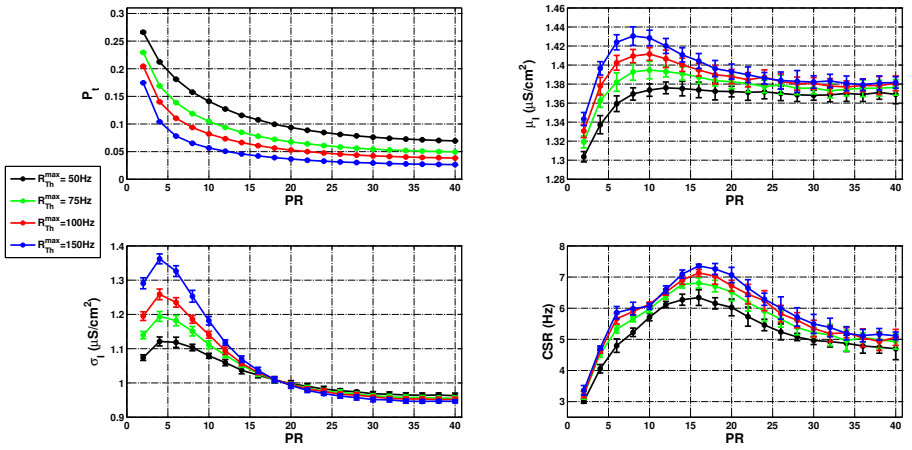
2.3.6.2. Model's response to sequences of a single tone

We thus set out to simulate the response of our model to sequences of tones of one single frequency, mimicking the settings of the study by de la Rocha & Parga (2008) (see table 2.2). Also here, we only changed the value of q_{CTh} and left all other parameter values untouched, and even so we were able to obtain the results found using the Integrate-and-Fire model in de la Rocha & Parga (2008).

Particularly, we wanted to make sure that we could obtain a similar dependency on PR of the key observables P_t, μ_I, σ_I and CSR (equations 2.4 - 2.7). The results of this fit can be appreciated by comparing the subfigures (a) (original) and (b) (our fit) of figure 2.19.



(a) Figure 3 from de la Rocha & Parga (2008).



(b) B only tone sequence. Transmission probability, current mean (μ_I) and s.d. (σ_I), Cortical Spike Rate (CSR) - equations (2.3) - (2.7), averaged across neurons (our 'trials'). All parameters as in table 2.2 (q_{CTh} value marked by ^{*1}), except cortical and thalamic background activity, which was set as in section 2.2.3.

Figure 2.19: The response of various observables to sequences of identical stimuli, in (a) somatosensory ones (de la Rocha & Parga 2008), in (b) our auditory B only tone sequence. Details in subfigures. By comparing the four subplots of our fit to the original figure's panels **(a)-(d)**, we see that we reproduce the trend in the data. VS (Vector Strength) was also a fit (not shown).

Name	Value	
$R_{\mathrm{Th}}^{\max}$	100	Hz (varies)
U	0.8	
τ_D	300	ms
q_{CTh}	$0.35^{*1}/0.5/1.05^{*2}$	$\mathrm{nC/cm^2}$
$\sigma_{q_{\mathrm{CTh}}}$	$0.25q_{\mathrm{CTh}}$	$\mathrm{nC/cm^2}$
K	85	
M	5	
N	1	
M_s	7	
μ	0.5	oct.

Table 2.2: Values of biophysical parameters for HH model, approach two: figures 2.19(b) (^{*1}deviant value), 2.22 (^{*2} deviant value), 2.23 and 2.24.

Having done that, we knew we had a functioning network and could concentrate on the issue at hand. First of all, we wanted to confirm what we had found already using the rate model, namely that when presenting only B tones we get a degradation of frequency tuning due to synaptic depression. First we corroborated this by using the same analytic methodology as when we analysed sequences of alternating tones in section 2.3.2. This yielded figures 2.20 and 2.21, clearly illustrating the effect (note that we don't use any threshold, Θ, here).

Finally, we simulated the responses to such a sequence of B tones in our second approach HH network. Figure 2.22 illustrates how the result is very similar to both that of the rate model (figure 2.12a) and of the analysis 2.21 (we raised q_{CTh} for this simulation).

2.3.6.3. The streaming paradigm: Differential suppression

Having thus reproduced the behaviour of the model of de la Rocha & Parga (2008) when stimulated with sequences of non-differing pulses, we finally had the opportunity to test our hypothesis that this very same network could be used to model auditory streaming correlates, with a minimum change in parameter values. First, we set about this using the same parameter values as for figure 2.22, but the result was discouraging (*RR* values far too high). Figuring the strength of the input had brought us out of FDR, we lowered the value of q_{CTh} successively until arriving at

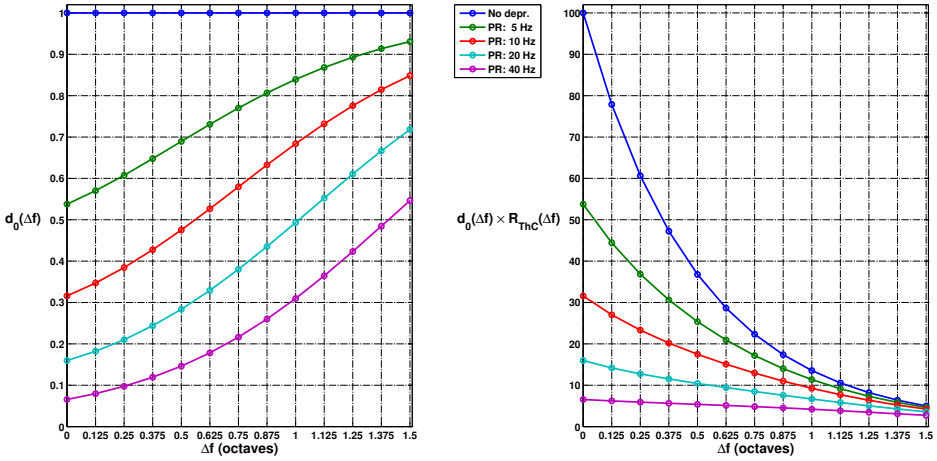


Figure 2.20: Making of BBBB... sequence 'tuning curve' - d_0 level multiplied by R_{Th} yields approximation for response at each Δf and PR . Parameters: $U = 0.8$, $\tau_D = 0.3$ s, $TD = 0.02$ s and $\mu = 0.5$ oct.

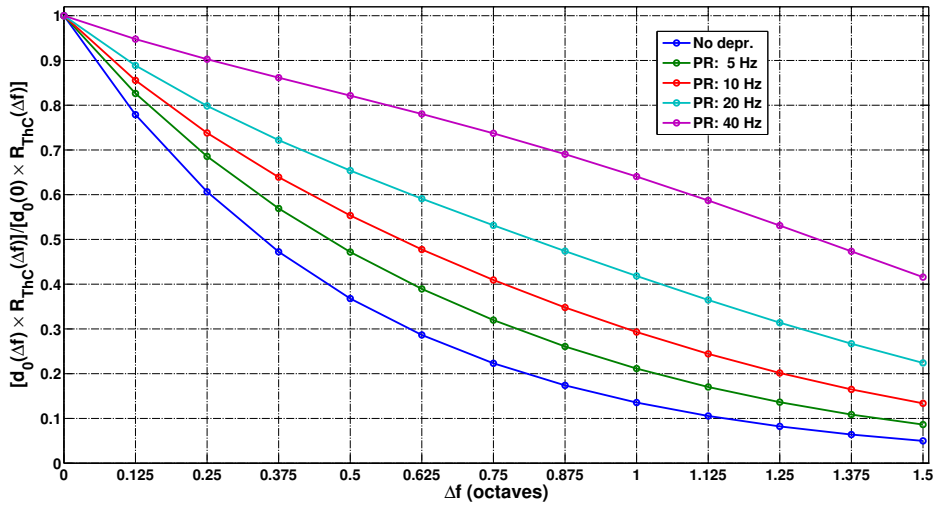


Figure 2.21: BBBB... sequence 'tuning curve' based on quotients calculated from curves in the right panel of figure 2.20 (for each curve, values at $\Delta f > 0$ divided by value at $\Delta f = 0$). Parameters as in figure 2.20.

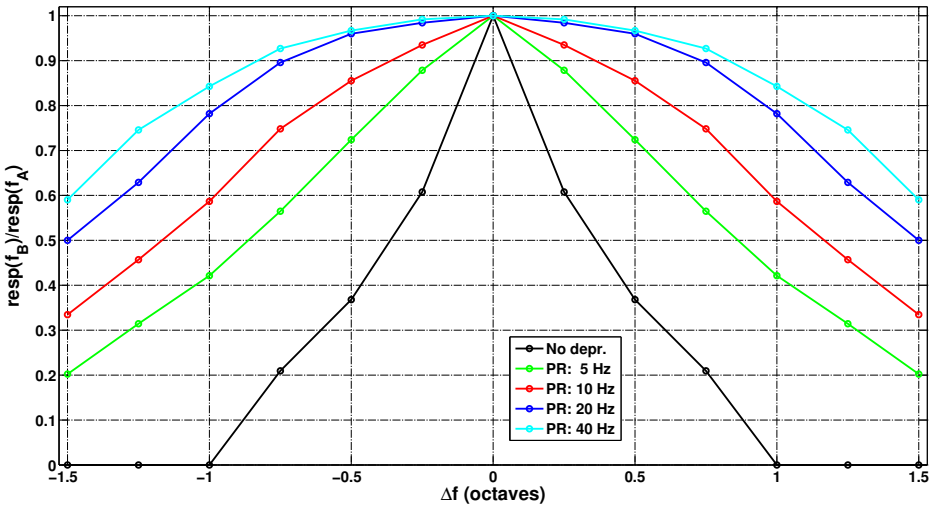


Figure 2.22: BBBB... sequence 'tuning curve'. Notice how the frequency resolution diminishes as *PR* increases. All parameters as in table 2.2 (q_{CTh} value marked by ^{*2})

$q_{\mathrm{CTh}} = 0.5$ nC/cm^2, value for which we obtained the curves shown in figure 2.23. This figure depicts how the results improve with a narrowing of the input (lower μ). Errorbars are omitted, since in spite of averaging over 100 seconds of simulation time, the coefficient of variation was consistently high (>1 at $\Delta f = 0$ %, gradually increasing with increasing Δf; data not shown). This was indicative of the important role of averaging for getting the results seen in Fishman et al. (2004), who in fact used averages across 75 stimulus presentations for their data, for each recording site, then further averaging across recording sites (cf. N of their figure 3, our figure 1.3). If we use a shorter data analysis interval, $T_2 - T_1$, we do not obtain the nice curves of figure 2.23.

We then proceeded to test the robustness of our result to the introduction of non-zero thalamocortical axonal spread, setting $\nu = 0.33$ octaves. As can be seen in figure 2.24, this dilutes what we refer to as the dynamic frequency selectivity substantially, for both values used for μ. This issue is a manifestation of a known dilemma of A1 (Wang 2007), namely that of a neuron maintaining frequency tuning while having a broad spectral perspective.

We shall return to a possible solution for this dilemma using intracortical

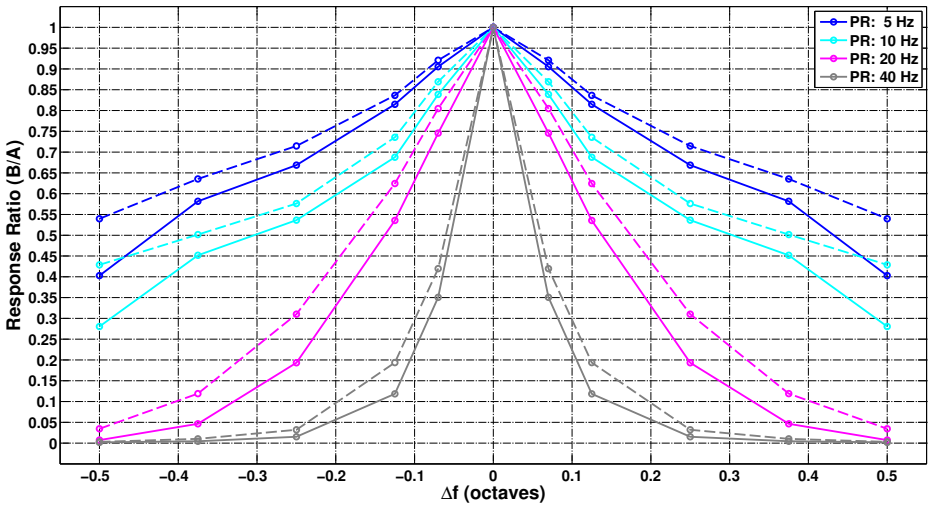


Figure 2.23: ABAB... sequence response ratio curves. Already with $\mu =$ 0.5 oct. (dashed lines) we get a good fit of Fishman et al. (2004) data, which improves further when narrowing the thalamic response profile by setting $\mu = 0.4$ oct. (solid lines). We simulated 200 s.; $T_1 = 100$ s., $T_2 =$ 200 s. $q_{\mathrm{CTh}} = 0.5$ nC/cm^2 , $\sigma_{q_{\mathrm{CTh}}} = 0.125$ nC/cm^2, $\nu = 0.0$ oct. All other parameters as in table 2.2.

recurrent excitation in chapter 3, inspired by data in Liu et al. (2007) that show that such excitation can act as a sharpening influence on the rather flat response to thalamic input seen when intracortical excitation is silenced.

2.4. Discussion

Several important aspects of the auditory stream segregation phenomenon, as manifested at the level of the primary auditory cortex (Fishman et al. 2004, Micheyl et al. 2005), are studied using a simple stimulus consisting of a sequence of pure tones that alternate between two frequencies (Bregman & Campbell 1971).

Although synaptic depression has been shown to degrade tuning, especially at high firing rates (Abbott et al. 1997), here we demonstrate that it still is able to produce an improvement of the neurons' frequency tuning via a synaptic mechanism that we designate *dynamic frequency selectivity*. This mechanism is implemented in three different neuronal models, all of which produce results having a good agreement with experimental data.

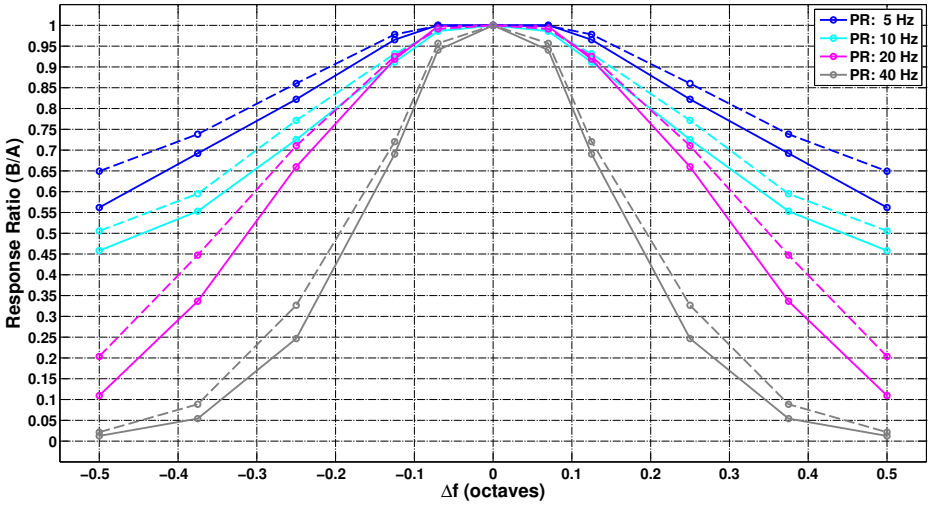


Figure 2.24: ABAB... sequence response ratio curves. When adding thalamocortical axonal spread ($\nu = 0.33$ oct.), the dynamic frequency tuning is obviously weakened by the integration of neighbouring inputs at φ_A (cf. figure 2.23). We simulated 200 s.; $T_1 = 100$ s., $T_2 = 200$ s. $q_{\mathrm{CTh}} = 0.5$ nC/cm^2 , $\sigma_{q_{\mathrm{CTh}}} = 0.125$ nC/cm^2. Solid lines: $\mu = 0.4$ oct. Dashed lines: $\mu = 0.5$ oct. All other parameters as in table 2.2.

We also show that the differential suppression phenomenon corresponds to the steady state, whereas the forward masking phenomenon (Calford & Semple 1995, Brosch & Schreiner 1997, Wehr & Zador 2005) is associated with the model's transient regime, a very important distinction considering it was earlier thought that these two phenomena shared the same mechanism (Fishman et al. 2001, Kanwal et al. 2003, Fishman et al. 2004, Denham & Winkler 2006).

Here, we've focussed on the few essential ingredients sufficient for our different models to be capable of reproducing the experimental data, relaxing the demands on neurophysiological accuracy of the models, both in terms of their constituent parts and of their parameter values. This required a process of trial and error, resulting in a feed-forward model design, representing an isofrequency patch of A1. This design included only excitatory neurons and thalamocortical synaptic depression, leaving out inhibitory neurons, the tonotopic extension of cortex and intracortical connections. The simplicity of design was a direct consequence of our findings during the work, namely that with the joint effect of the spiking threshold of neurons (providing non-linearity of responses), the thalamo-

cortical depression (establishing the steady state over time), the tonotopic organization of the thalamic response (making response in the modelled patch of A1 tonotopic as well, via thalamocortical connectivity), and a gradual increase or decrease of model response with experimental parameters (either by fiat or through cross-neuronal and/or cross-trial variability in responses, depending on the level of modelling detail), we are able to reproduce the data of Fishman et al. (2004) for short tones, as well as the dependence on time of the appearance of stream segregation (correlates) at the neural level (Micheyl et al. 2005) (for the latter issue, we also refer the reader to our scientific movies made with an intracortically connected conductance-based model network, detailed in chapter 3, including a link to the movies).

More specifically, the threshold explains the decline of response ratio with presentation rate; the depression explains the decline in absolute response to any tone with increasing presentation rate; the tonotopic organization of thalamic response explains the decline of response ratio with increasing frequency difference between a sequence's tones; finally, the gradual change of model response with experimental parameters was decisive in reproducing a smooth decline in both absolute responses and response ratios with increasing presentation rate and frequency difference. In the rate model this is simply imposed by design, through the use of a linear f-I characteristic (or transfer function; see equation (A.2) in appendix A) for the A1 population. However, when we simulate a whole population of conductance-based neurons of Hodgkin-Huxley type (see appendix B), this can not be imposed but relies instead upon a certain level of variability of neuronal responses, both across individual neurons and across trials. Such variability may be induced intrinsically through varying the neurons' resting potential, or through a spread in the values of thalamocortical synapses' parameters (e.g., U, τ_D, g_{CTh}) and/or making stochastic that very synaptic transmission, all of which can shape the effect of external inputs impinging on neurons. Applying various of the mentioned methods, we demonstrate that variability in the intrinsic qualities of single neurons and synapses is necessary for our solution to yield good averaged data fits in all models.

One caveat of our network is that when one takes into account the thalamocortical input spread seen in the literature (appr. 1/3 octave; see Miller et al. 2001), the difference of responses with frequency difference gets flattened (compare figure 2.23 to figure 2.24). In order to address this weakness, in chapter 3 we explore a tonotopically organized 'network of networks' ($N > 1$) of neurons not only in terms of its response properties in the auditory streaming paradigm, but also regarding how its structure can

be constrained using results in neurophysiological studies of responses to single tones (e.g., Wehr & Zador 2003, Liu et al. 2007, Wu et al. 2008) and two-tone forward masking stimuli (primarily based on a study using clicks; Wehr & Zador 2005). Regarding the dilution of frequency resolution with the spread in thalamocortical input mentioned here, we find that one needs the addition of intracortical recurrent excitation to the network in order to reasonably maintain the fit of the data found in the present chapter *without* such a spread.

In short, in this chapter we not only reproduced the data found by Fishman et al. (2004), but analysed the underlying possible causes of the differential suppression phenomenon at the circuit level and arrived at a simple yet powerful mechanism that explains *how* stream segregation correlates arise at the level of the primary auditory cortex. Our exhaustive analysis of the role of depression in our DS mechanism, both analytically and with minimal preliminary simulations, yielded many useful tools for assessing our model's response to sequences of tones, which will come in handy in future expansions of the work at hand. We further confirmed that our mechanism works as claimed in three different models, one rate model and two conductance-based models, without the need for intracortical connections. This already in and of itself is a considerable achievement in terms of robustness, made more remarkable by the fact that one of the conductance-based models (our second approach) was adopted with minimal parameter changes from an earlier modelling study (de la Rocha & Parga 2008) of sequences of sensory stimuli with *identical* attributes in a *haptic* context! Nevertheless, we further look into robustness in the next chapter, as already outlined above.

3 Mechanism robustness and beyond: tonotopic conductance-based model

The author of this thesis presented part of the work of this chapter at the Twentieth Annual Computational Neuroscience Meeting, held in Stockholm, Sweden in July 2011[1]

3.1. Introduction

In this chapter, we will complement our theory and model of differential suppression (possible correlate of the so-called *primitive* auditory stream segregation phenomenon; Bregman 1990) presented in chapter 2. Specifically, we further expand our neurophysiologically detailed modelling study using neurons of Hodgkin and Huxley type, by now coupling them via intracortical synapses, in order to form a network. Furthermore, we now add a physical extension to our network by representing up to six octaves of the tonotopic axis of the cortex, employing one such coupled network of neurons to represent each discrete point (or 'column') of the axis that we model (these columns are in turn mutually connected; see appendix B for details). Finally, we 'activate' our 'dormant' inhibitory neurons in order to explore the role of inhibition in the phenomena that interest us. Throughout this study we use an equal number of excitatory RS and inhibitory FS neurons, in order to limit the considerable complexity involved in our task. We discuss the repercussions of this and other limitations in chapter 5 (Discussion and Conclusions).

All this enables us to study in more detail which biophysical parameters are the most important in reproducing experimental results, both those parameters intrinsic to neurons and the ones pertaining to the synaptic

[1] Abstract published: Larsson, Montbrió & Deco (2011)

coupling of the network. We present here further results on both streaming correlates and responses to other stimuli, all of which nicely complement the previous chapter. We also identify a specific plausible intracortical connectivity architecture for layer IV of A1, the addition of which to our previous network does not perturb earlier results but rather improves upon them (see below).

Specifically, our model accounts for forward masking seen using noise stimuli (Wehr & Zador 2005), while showing approximately balanced excitation and inhibition (Wehr & Zador 2003, Wu et al. 2008, Tan & Wehr 2009) and specifically the sideband dominance of inhibition over excitation found by Wu et al. (2008). Regarding two-tone suppression (neurophysiological forward masking using pure tones), for a long time, studies argued for inhibition as the main mechanism responsible for seen in primary auditory cortex (A1) neurons (Calford & Semple 1995, Brosch & Schreiner 1997). However, both computational (Denham 2001) and experimental (Rose & Metherate 2005) studies afford a significant role to thalamocortical (ThC) synaptic depression in shaping the temporal response properties of A1. Also, the duration of inhibitory currents in A1 has been an issue of contention (Wehr & Zador 2003, Tan et al. 2004). The study by Wehr & Zador (2005) used noise click stimuli to show that while responses to the probe were not fully recovered even 512 ms after presentation of the masker, inhibitory currents evoked by the masker lasted at most 100 ms, coinciding in duration with the complete suppression of probe responses. The authors proposed that a longer-lasting mechanism such as ThC or intracortical (IC) synaptic depression could complement inhibition by accounting for the lingering effect seen. They also demonstrated that pentobarbital anesthesia significantly prolongs inhibition, thus calling into question earlier results (Calford & Semple 1995, Brosch & Schreiner 1997). All this led us to conclude that, using a version of our model with ThC depression *and* intracortical connectivity, we should be able to reproduce results by Wehr & Zador (2005) in considerable detail, which we do in this chapter.

In summary, we develop a tonotopically organized network model of layer IV of the primary auditory cortex (A1), capable of reproducing the main characteristics of important recent neurophysiological experimental results (mainly Fishman et al. 2004, Wehr & Zador 2005, Liu et al. 2007, Wu et al. 2008). We chose to focus on layer IV for simplicity, as it is the main thalamorecipient layer of A1, its pyramidal cells receiving afferent axons from the ventral Medial Geniculate Body (MGBv; Smith & Populin 2001, Winer & Lee 2007, Viaene et al. 2011). Inspired by Liu et al. (2007),

we here propose a plausible IC connectivity for layer IV of A1, which selectively amplifies the broad input from the thalamus to yield the sharp frequency tuning seen in many studies of A1 (e.g., Liu et al. 2007, Fishman & Steinschneider 2009). This amplification furthermore contributes to improving the frequency resolution of our model when we consider a non-zero spread of thalamocortical afferents, thus resolving the dilemma this provoked in chapter 2. We conclude that a combination of IC currents and ThC synaptic depression is imperative in any attempt to account for the wealth of data seen in the neurophysiological literature, including the phenomena we studied here.

3.2. Methods

When setting up this more elaborate model, neurophysiological constraints were of the essence in order for complexity not to run amok. In this sense, we took into consideration several important results from recent literature, for instance that about 60% of excitatory inputs to A1 neurons are thalamocortical in origin (Liu et al. 2007). In addition, the fact that Bruno & Sakmann (2006) found that somatosensory cortex LIV neurons receive a barrage of synchronous sensory-driven thalamic inputs strong enough to evoke cortical spikes, in conjunction with Wang et al. (2010) inferring that input synchrony is typically in the range of 20-60 synchronously active synapses, led us to strenghten our working hypothesis that thalamocortical synaptic depression likely contributes to, e.g., the forward masking phenomenon and, of course, differential suppression in A1. Moreover, Liu et al. (2007) showed a boost in excitatory currents of approximately 67% and subthreshold and spiking activity bandwidths (measured at zero % amplitude) of approximately 4 and 2 octaves, respectively. Later, Wu et al. (2008) showed that while ThC input is equally broad to excitatory and inhibitory cells in A1, the latter have broader spiking response than excitatory cells (77% vs. 54% of subthreshold response width, respectively). Also, in excitatory neurons, inhibitory current input range is 7.5 % narrower than excitatory ditto. These data led us to the conclusion that inhibitory cells are synaptically or intrinsically more excitable than excitatory cells. We further inferred an approximate spread of intracortical inhibition of $\lambda_I = 0.4 \pm 0.1$ oct. (when data was combined with Liu et al. (2007) results). Moreover, we take into consideration the result of Miller et al. (2001) that ThC projections are focal with a radius of at most $1/3$ of an octave. Finally, we introduce a baseline variability in cortical neurons' resting potential and their synaptic strengths, which in turn provides variability in cortical

responses across neurons in a population (Rose & Metherate 2005).

In this chapter we model three situations, namely the presentation of a single pure tone (25 ms duration, $\cos^2$(t) up- and down-ramps of 5 ms duration), the presentation of two noise clicks in succession and the presentation of sequences of pure tones (of the same frequency or of different frequencies, depending on parameter Δf). The first case is used when investigating which basic network setup is able to explain most of the available neurophysiological data (see sections 3.2.1 and 3.3). We present the tone at the center of the TT axis, i.e., at F_0. At the level of the cortex, the corresponding point is called φ_A, as in streaming simulations (since A tones are always presented at center). In the case of the noise clicks, used to model forward masking experiments (see section 3.3.3), we first present a masker and then a probe, with stimulus onset asynchrony SOA (= 1/(Presentation Rate), $1/PR$). We consider the values $SOA = 32, 64, 128, 256$ and 512 ms. For technical reasons we approximated these SOA values with $1/PR$, $PR \in \{32, 16, 8, 4, 2\}$, with no important effects on results (data not shown). As for the sequences of pure tones, they are used for modelling the neural correlates of auditory streaming with this recurrent network, in section 3.3.4.

3.2.1. The model: neurons, network and canonical settings

The primary auditory cortex, A1, is tonotopically organized according to the characteristic frequency (CF), i.e., that frequency which can evoke a spiking response at the lowest stimulus intensity. The tuning properties of the neurons are inherited from the ventral division of the Medial Geniculate Body (MGBv) of the thalamus, but are further refined in A1 (Escabí & Read 2005). Thalamic efferents terminate on excitatory and inhibitory neurons in A1 (primarily in layers II/III and IV; Cruikshank, Rose & Metherate 2002, Rose & Metherate 2005, Viaene et al. 2011), which constitute a subset of an extensive network of neurons. In consequence, the firing properties of neurons in A1 are determined by synaptic inputs from both the MGBv and the local cortical network of excitatory and inhibitory neurons. Based on recent experimental data from neurophysiological studies, a detailed neural model of layer IV in A1 has been developed (where the main, so-called driving thalamocortical input arrives; Viaene et al. 2011). The model consists of two types of neurons, regular-spiking (RS) excitatory neurons and fast-spiking (FS) inhibitory neurons, all arranged along an axis representing the tonotopic organization of the auditory cortex, and highly interconnected (see figures B.2 and B.3). This is the main difference

with the HH type models used in chapter 2 – here, we model the *extension* of A1 by setting $N > 1$. N takes on different values depending on how many octaves we model, but in general we use a density of 'columns' per octave of $\rho = 16$. Additionally, all neurons in the model receive feed-forward excitatory inputs via thalamocortical synapses, as before, but in this model we look more closely at the effects of setting $\nu > 0$, hinted at in chapter 2. As described in appendix B, we implemented this structurally rather than with a simple modulation of weights – a certain number K_n of axons is sent from each zone of the thalamus (not explicitly modelled) to all of the A1 'columns' within its reach. The advantage of this approach is that it renders each synapse *exclusive* to input originating at one particular point in the thalamus, as most probably occurs in the brain. In contrast, in our HH models of chapter 2, A and B tone input share synapses, something which mixes the effect of the different tones on a *particular* synapse. While this is not important when one looks at averages of network response, the realism introduced here is preferable and has the added advantage of enabling us to look at data from individual synapses, should the need occur. The details of all this, as well as all other aspects of the model may, as before, be found in appendix B. All implementations necessary to model the full extension of A1 were added to the earlier program written in C and C++.

To establish suitable canonical parameter values, we took the paper by Rose & Metherate (2005) as a starting point. In it, the authors experimentally establish values for a number of important biophysical parameters, including the standard deviations of the same. Fitting our network settings based on those values, helped with the task of constraining the model and lowering the number of free parameters. In the following, we describe the fitting procedure. Please note that we regard the fit values only as approximate, for the following reasons (amongst others):

- the Rose & Metherate (2005) paper is based on slice work
- data of other papers we attempt to reproduce sometimes use different species
- in general, experimental methods and setups may vary across studies

3.2.2. Data analysis

Although at first we only made an approximate fit of our network's response to single tones, using only feed-forward inhibition and recurrent excitation (sections 3.3.2.1 - 3.3.2.4), we found this to reproduce the

data insufficiently well. Thus, in order to fit as well as possible as much as possible of the neurophysiological data on the A1 microcircuit, we constructed a goal function to which we applied the minimum least squares method. Specifically, we focused on

1. the bandwidths (BWs) of spike responses of E and I neurons
2. the ratio of the BW of spike response to the BW of membrane potential response, for both E and I neurons (%),
3. the level of membrane potential at which an E neuron spikes (% of maximum), and
4. the average difference (in %) between inhibitory and excitatory conductance response BWs, at five different amplitude levels (0, 20, 40, 60, 80 % of maximum).

The goal function thus had ten constituents, which were weighted as follows (enumeration as above):

1. 15% each
2. 20% each
3. 10%
4. 4% for each level

The goal values we used were (see section 3.3.2 for reasoning, references and results of this fitting procedure)

1. 2.17 oct. (E), 3.09 oct. (I),
2. 54.3% (E), 77.3% (I),
3. 59.6%,
4. -7.8, -3.3, 15.0, 27.5 and 32.5 % difference, respectively.

The goal function value (to be minimized) will be referred to as GF_{val} in the following. The results of this more elaborate fitting procedure is presented in section 3.3.2.5.

Regarding the calculation of *RR* for streaming simulations, the tonotopic extension of our model adds another aspect which warrants comment.

Fishman et al. (2004) measure so called Multi-Unit Activity (MUA) extracellularly, comprising spikes from a sphere of about 100 µm diameter, which for us raises the issue of whether we should sum activity from 'columns' neighbouring the BF region or not. Basing our judgment on data saying an iso-frequency column has a diameter of 50-100 µm (Pickles 2008), we chose to only measure in the BF column at all times, as outlined in chapter 2.

3.3. Results

3.3.1. Setting up a canonical network

3.3.1.1. Resting potential variability

Rose & Metherate (2005) report a slight spread in values of measured membrane potentials in regular-spiking (RS) and fast-spiking (FS) neurons. As the membrane potential of a neuron will relax to its leak current resting potential, V_L, in the absence of any input, we decided to introduce a spread in V_L across neurons. We based our fit on the values in table 1 of the experimental paper, specifically fitting the resulting coefficient of variation, since we used the same V_L mean value in all our model neurons, inhibitory and excitatory alike. We arrived at $\sigma_{V_L,E} = 1.12$ mV (1.6 %) and $\sigma_{V_L,I} = 1.89$ mV (2.7 %), based on $V_L = -70$ mV, which we use throughout this work.

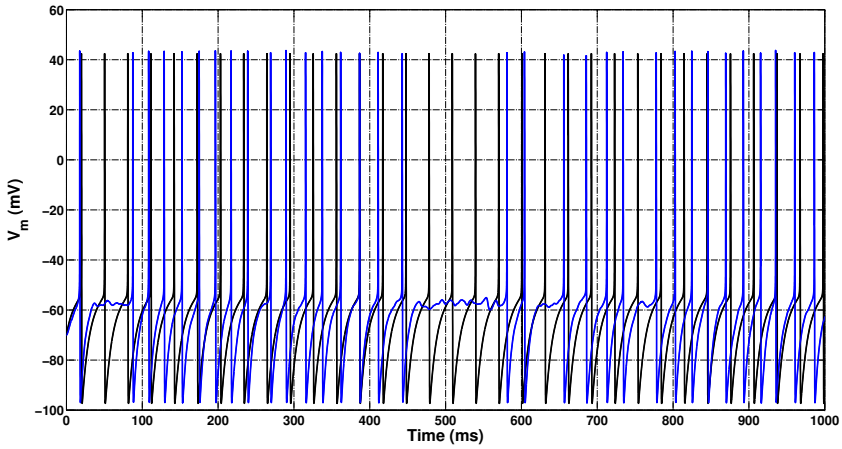
3.3.1.2. Spike Frequency Adaptation (SFA)

We performed a fit of SFA (see appendix B and specifically figure B.1). However, this is not further employed in the thesis, as results were not deemed to be negatively affected by its exclusion.

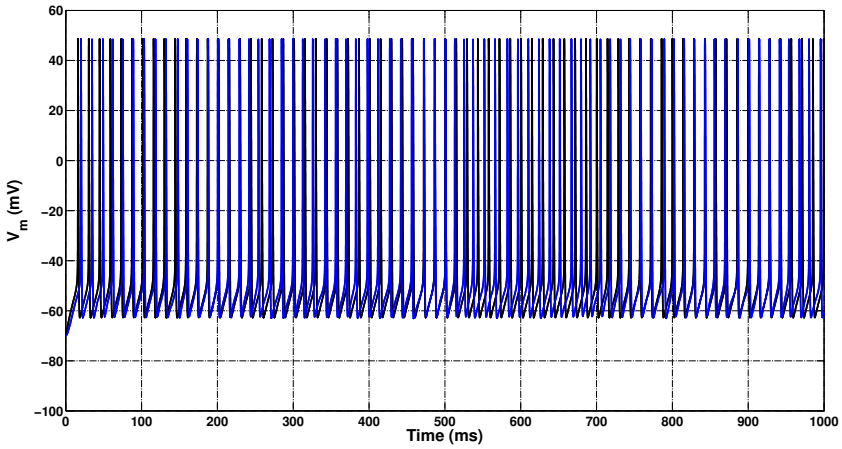
3.3.1.3. Spike response, FS and RS cells

In figure 1 in Rose & Metherate (2005), examples are given of RS and FS neuron traces in response to a minimal stimulating current. We have performed simulations to check what response we get to minimal stimulation with different parameter settings. First of all, we identified the (constant) current level at which both our RS and FS neurons showed a sustained response to a long-lasting input (900 ms in the experimental paper, we use 1000 ms). This level was identified as $I_{\mathrm{CTh}} = 1.66$ µA/cm^2. We proceeded to identify which value of g_{CTh} was needed to get the same level of I_{CTh} with Poisson input. The result was a value of $g_{\mathrm{CTh}} = 0.1$ mS/cm^2 (with $K = 1$) or $g_{\mathrm{CTh}} = 0.001$ mS/cm^2 (with $K = 100$, yielding more evenly spread out

response). In figures 3.1 and 3.2, we show the spike responses obtained with these settings, for the duration of the input and during the first 100 ms, respectively (for the conductance input simulations, we used $K = 100$). Subfigure (a) shows the regular-spiking cell response, whereas subfigure (b) shows the fast-spiking cell response.

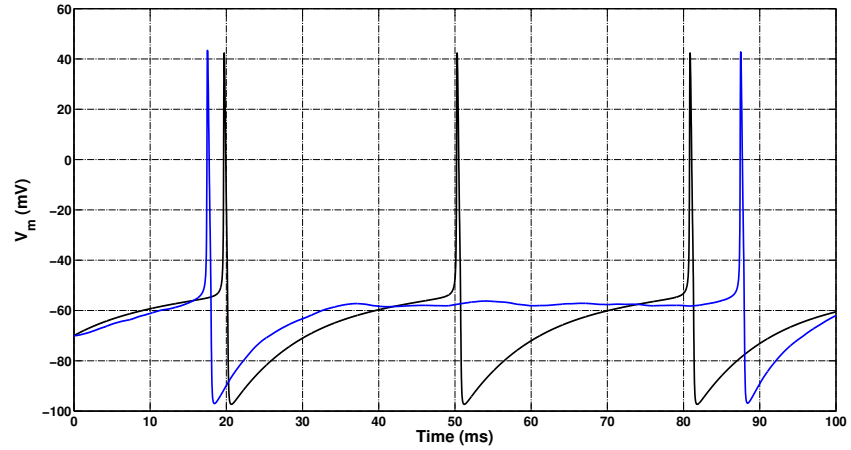


(a) RS Excitatory cell

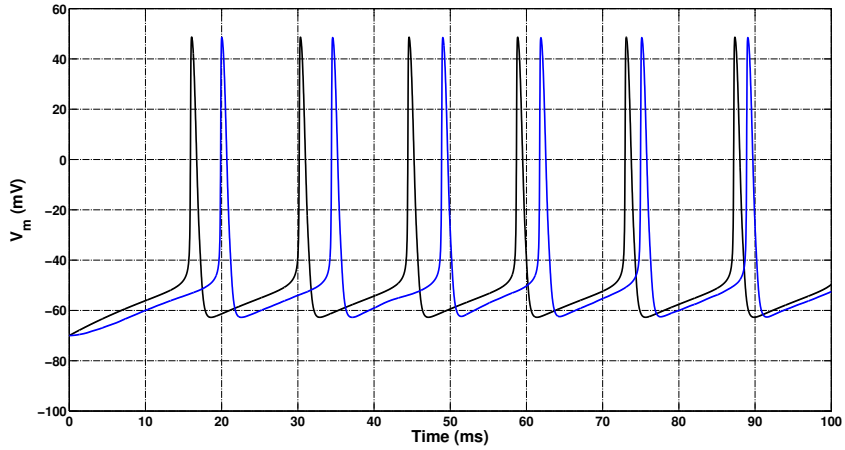


(b) FS Inhibitory cell

Figure 3.1: An excitatory (a) and an inhibitory (b) model neuron's response to minimal stimulation, delivered both as a constant current (black lines, $I_{CTh} = 1.66\mu\text{A/cm}^2$) and using a Poisson process to generate conductance-based input (blue lines, $g_{\text{CTh}} = 0.001\text{mS/cm}^2$, $K = 100$ and $R_{Th} = 100$ Hz). $TD = 1000$ ms.



(a) RS Excitatory cell



(b) FS Inhibitory cell

Figure 3.2: Detailed view showing the first 100 ms of the responses shown in figure 3.1.

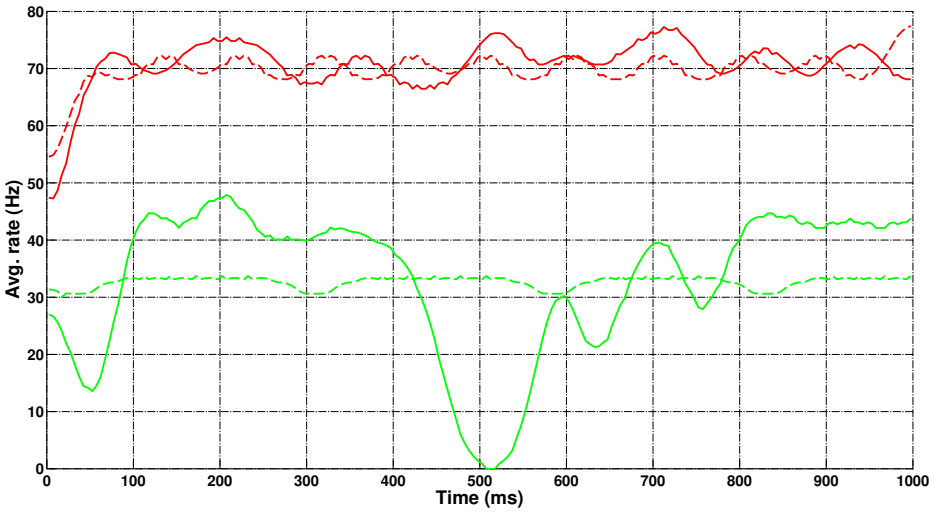


Figure 3.3: The RS and FS neurons' firing rate, smoothed with a Gaussian of width 5 ms, calculated from data shown in figure 3.1. Excitatory neuron's responses in green, inhibitory neuron's responses in red. Solid lines show responses to Poisson input, whereas dashed lines show responses to constant input. *Average* rates (ISI^{-1}) with constant current input: 32.7 Hz (RS), 70.2 Hz (FS); with Poisson input: 33.0 Hz (RS), 71.8 Hz (FS). Note how fluctuations in the Poisson input can affect response this close to the threshold of activation. Here, ISI means Inter-Spike Interval.

Figure 3.3 shows average rates for RS (green) and FS (red) neurons over 1000 ms, for both constant (dashed) and Poisson (solid) input. It is important to observe how fluctuations in the input makes transmission unreliable this close to the minimal stimulation needed to obtain a response. However, if we had situated the constant input just one hundredth of a $\mu\text{A}/\text{cm}^2$ lower, we would get no response at all in the RS neuron, while the fluctuations of a Poissonian input would then *increase* transmission probability when comparing to the response to constant input. This is important in obtaining responses when in a fluctuation-driven regime, whether it be set deliberately (de la Rocha & Parga 2005) or obtained due to continued stimulation causing a steady state (de la Rocha & Parga 2008), as we saw in our study of streaming correlates using our approach two HH model in chapter 2.

3.3.1.4. EPSP size, shape and variability

Using data from figure 2 and table 2 in Rose & Metherate (2005), we were able to preliminarily set values for g_{CTh} and $\sigma_{g_{\mathrm{CTh}}}$. Specifically, based on amplitude numbers in that paper, we aimed at an average EPSP value of about 1.5 mV (with depression, as fit in the next section). The value of 1.45 mV is the average of RS and FS means in the paper, and we use this since no statistically significant difference between their means was found. The standard deviation across trials for the amplitude was 0.3 mV, so we set $\sigma_{g_{\mathrm{CTh}}}$ with this value in mind. In particular, we set the values $g_{\mathrm{CTh}} = 0.02$ mS/cm^2 and $\sigma_{g_{\mathrm{CTh}}} = 0.004$ mS/cm^2, respectively. We ran simulations with 60 different seeds with these values. The results without (blue curve) and with (green curve) depression may be seen in figure 3.4. Dashed curves mark standard deviations for each mean (solid) curve.

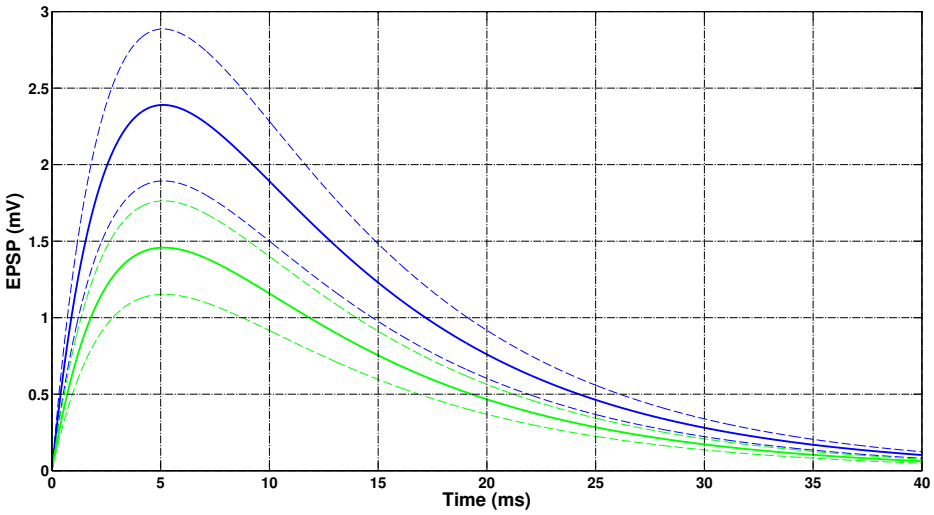


Figure 3.4: With $g_{\mathrm{CTh}} = 0.02$ mS/cm^2 and $\sigma_{g_{\mathrm{CTh}}} = 0.004$ mS/cm^2, we ran simulations with 60 different seeds. Shown are mean (solid) and standard deviations (dashed) without (blue) and with (green) ThC depression ($U = 0.4$). Peak values: 2.39 ± 0.50 mV and 1.46 ± 0.31 mV, respectively.

Finally, with regard to EPSP shape, we chose to model rise time as instantaneous for both RS and FS cells, for simplicity. While rise times found by Rose & Metherate (2005) were in fact different, in absolute terms they were negligable, in particular when compared with the remarkable difference in EPSP width between the two cell types. The primary parameter controlling the width is the decay time constant, τ_{AMPA}. However,

in our simulations we use the value of $\tau_{\mathrm{AMPA}} = 3.0$ ms for both synaptic connections, in order to keep the complexity of our model manageable. For completeness, it would be desirable to introduce two separate decay terms, $\tau_{\mathrm{AMPA,E}}$ and $\tau_{\mathrm{AMPA,I}}$, and look at how setting the latter higher than the former influences results, but we leave this for future work.

3.3.1.5. Fitting depression parameters to EPSPs in response to a 40 Hz tetanus

Based on the data in figure 3C of Rose & Metherate (2005), we made an orientative fit of the the values of U and τ_D. The word orientative is appropriate in this context, since depression has been shown to vary widely in strength and duration across studies (Thomson & Deuchars 1994, Abbott et al. 1997, Gil et al. 1999, Carandini, Heeger & Senn 2002, Chung et al. 2002). We simulated the EPSP responses to a single presynaptic spike presented at 40 Hz, using $g_{\mathrm{CTh}} = 0.017$ mS/cm^2 . Using least-squares fitting, we found that the values of $U = 0.4$ and $\tau_D = 250$ ms yielded the best fit to the data of Rose & Metherate (2005) (see figure 3.5). As can be appreciated in the figure, these values are not set in stone - particularly, the time constant may be allowed to vary and we still get a good fit. Nevertheless, these values helped us decide on suitable ranges for these parameters in our simulations.

3.3.1.6. EPSP statistics and input fan-out

As for the number of presynaptic spikes, we deal with two cases - with and without fan-out of ThC connections (regulated by value of ν). Thus, with $g_{\mathrm{CTh}} = 0.01$ mS/cm^2 and $R_{Th} = 100$ Hz I get the following results. With $\nu = 0$ and $K = 100$ *or* with $\nu = 1/3$ oct. and $K_n = 10$ (which yields a total K value of about 110 = $K_n \times$(∼ 5 columns each side of tone's frequency + center column)), we get 150-180 presynaptic spikes during a 25 ms $\cos^2(t)$-ramped tone, which yields a value of 30-36 presynaptic spikes in a 5 ms period. This value is in the range of 20-60 events per 5 ms of synchronous input, which was found by Wang et al. (2010). It is evident from looking at figure 3.6 that using fan-out makes the number of events at different points in the network more equal, as we foresaw in our input analysis in chapter 2.

3.3.2. Network response to short tones

In order to fix the network settings, with regard to ThC and IC connectivity, we used primarily two recent experimental papers to guide us,

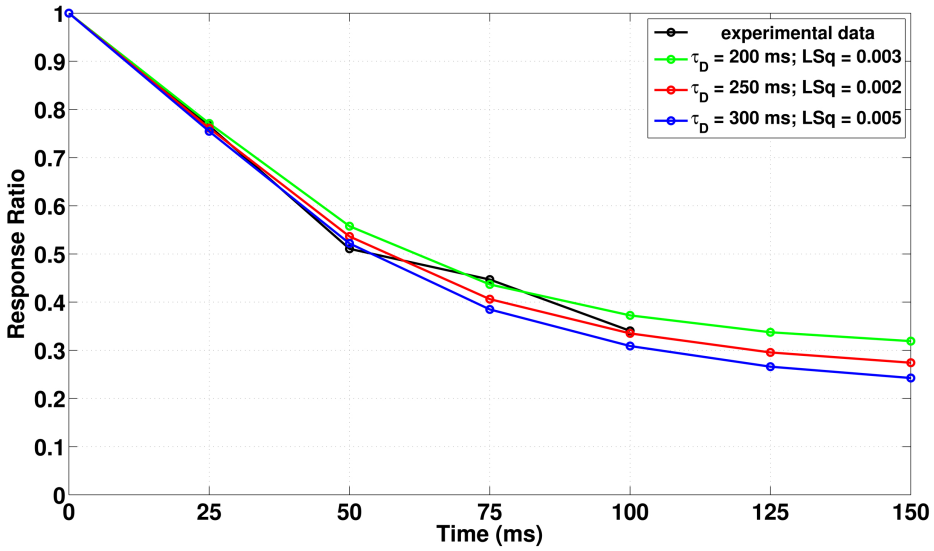


Figure 3.5: *RR* of successive EPSPs to first EPSP, evoked by a 40 Hz tetanus stimulus. Each EPSP is evoked by a single presynaptic spike presented every 25 ms. $g_{\mathrm{CTh}} = 0.017$ mS/cm^2. Black curve: data from Rose & Metherate (2005), extracted from their figure 3C. $U = 0.4$. τ_D values given in legends. While the best fit is achieved with $\tau_D = 250$ ms, good fits are achieved with that value set slightly differently.

namely Liu et al. (2007) and Wu et al. (2008). In summary, we tried to set the network up so that the following constraints, taken from those papers, were fulfilled as far as possible (BWX-Y means bandwidth at X dB and Y % amplitude):

- Liu et al. (2007)

  1. Average frequency range of spike response of RS cells, relative to membrane potential response, was 54 $\pm$7 %.
  2. RS cells show an approximate membrane potential BW60-0 of 5 octaves.
  3. About 61 $\pm$11 % of tone-evoked (TD=25ms) exc. input is ThC, 39 $\pm$11 % is IC.
  4. BW60-50 of ThC-only resp. is 25 % *wider* than ThC+IC resp.

- Wu et al. (2008)

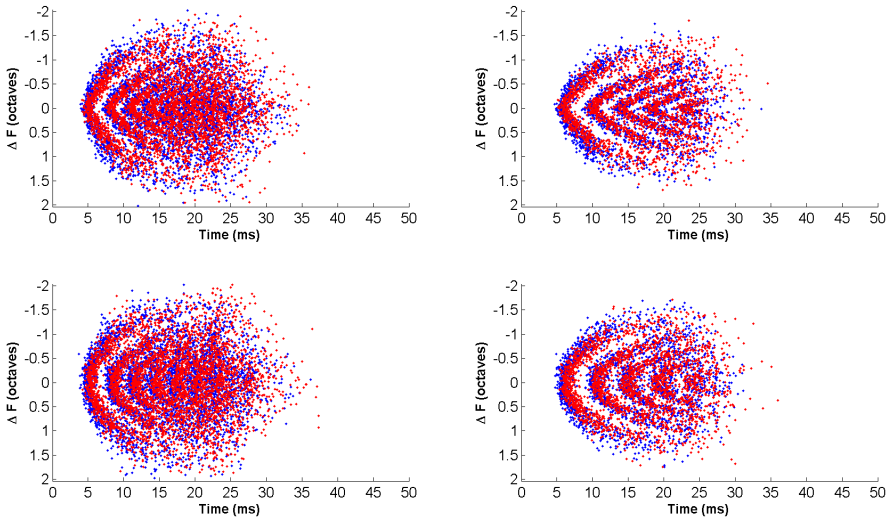


Figure 3.6: Spike rasters in response to input without and with depression (columns) and without and with fan-out of ThC input (rows), respectively. Parameters: $U = 0.0; 0.4$ (left to right), $\nu = 0; 0.33$ octaves (above to below). Red: inhibitory; blue: excitatory.

1. Average frequency range of spike response of RS cells, relative to membrane potential response, was 54.3 $\pm$7.6 % (confirming results of Liu et al. (2007)).
2. Average frequency range of spike response of FS cells, relative to membrane potential response, was 77.3 $\pm$8.5 %.
3. Both FS and RS cells' membrane potential responses show an approximate BW60-0 of 4 octaves.

The three points of Wu et al. (2008) allow us to calculate the avg. BW60-0 of spike responses, namely 2.17 octaves for RS cells and 3.09 octaves for FS cells, respectively. Since the BW60-0 of membrane potential responses are equal, this implies one of three things (or a combination of the three): 1. FS cells are intrinsically more excitable than RS cells. 2. λ_{IE} is wider than λ_{EE}. 3. g_{CTh} is stronger at FS cells' ThC synapses than at RS cells ditto. Alternative one is something we have included in our

model from the outset (see appendix B). Alternative two, while possible, is not contemplated in this study, for simplicity. While the data from Rose & Metherate (2005) refutes alternative three in saying that there was no statistically significant difference in mEPSP amplitudes between the two cell types, more recent results in other modalities (e.g., in somatosensory cortex, see Gabernet et al. 2005, Cruikshank et al. 2007, Daw et al. 2007, Hull et al. 2009) and even in A1 (Levy & Reyes 2012, Schiff & Reyes 2012), claim that minimal stimulation EPSPs are bigger in inhibitory neurons and attribute this to stronger synapses. This leaves also alternative three a possibility, i.e., a candidate solution mechanism. Our first priority in fitting this data was to be as faithful to the literature as possible. However, in the process of trying to satisfy all constraints, we found that compromise was necessary. Thus, in the end we opted for approximately fitting the frequency range of spike responses as found in both experimental papers, as well as fitting the BW60-0 of input conductances to between 4 and 5 octaves. By fitting the ranges we implicitly fit the spike response widths as well. Another important constraint we considered was the proportion of IC input of the total, which we fit to approximatly 40%. Fitting this with all IC currents active (voltage-clamp only blocks inhibitory currents in the *measured* neuron, not in the rest of the neurons in the network, which consequently integrate all their inputs to produces spikes that shape the afferent exc. currents throughout the network), we were able to also look at the relationship between the ThC-only and the ThC+IC responses. We found that also fitting the narrowing of the BW60-50 by IC input was impossible without negatively affecting the rest of the fit, including the IC portion of input and the spike response widths. Therefore, we compromised on this point, simply trying to ensure that the BW60-50 was maintained, i.e., at least not widened, when IC input was present (there is of course a sharpening of frequency tuning anyway, although closer to BF).

Below, we proceed as follows. First, we show the results of our initial approximate fit of the data outlined above, using only feed-forward inhibition and recurrent excitation, as mentioned in section 3.2.2 (sections 3.3.2.1 - 3.3.2.4). While those results were quite satisfactory, they did not meet our criteria and the fit was not sufficiently quantified, for which reason we employed a more exact methodology of minimizing error, explained in detail in that same section (3.2.2). The results of this second fit are explained in section 3.3.2.5.

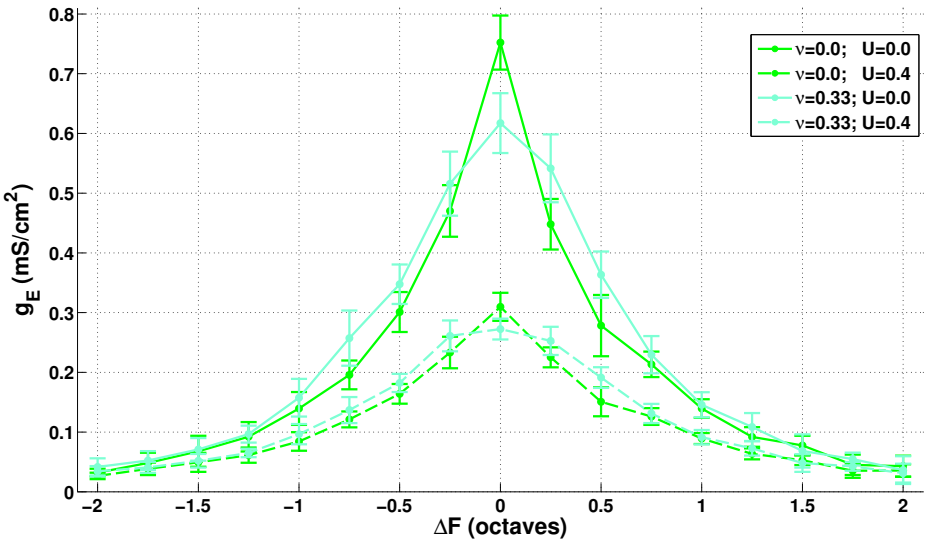


Figure 3.7: Setting $\nu > 0$ smoothes the conductance response evoked in cortical neurons across the TT axis by a single tone, as compared with when $\nu = 0$. The effect is robust when synaptic depression is activated ($U = 0.4$, $\tau_D = 250$ ms). Legend pairs lines with parameter values. $g_{kj} = 0$, $\forall k, j \in \{E, I\}$ (no intracortical connections active).

3.3.2.1. ThC input spread smoothes A1 conductance response profile

Figure 3.7 illustrates how the inclusion of a non-zero thalamocortical input spread smoothes the resulting conductance response profile evoked in the cortical neurons across the TT axis (cyan lines: conductances with $\nu = 0.33$, green lines with $\nu = 0.0$). Furthermore, dashed lines show the effect of depression in the two cases ($U = 0.4$, $\tau_D = 250$ ms). Here, $g_{kj} = 0$, $\forall k, j \in \{E, I\}$. This result (with $\nu = 0.33$) is in accordance with the fairly flat thalamocortical input profile across the tonotopic axis with intracortical connections silenced (Liu et al. 2007). A similar smoothing or flattening effect of a non-zero ν can be discerned in the contrasting rasters of figure 3.6.

3.3.2.2. IC excitation boosts excitatory conductance response by two-thirds

By setting $g_{\mathrm{EE}} = 3.9$ mS/cm^2, we got a boost in evoked exc. conductance by 67% (i.e., it adds 40% on top of ThC input), the figure found by

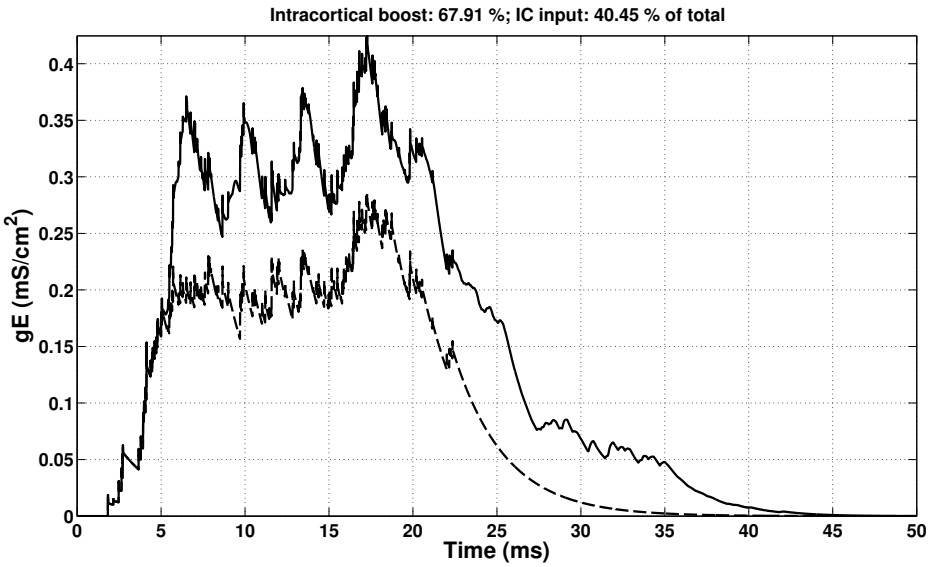


Figure 3.8: *Locally* (at φ_0) fitting data by Liu et al. (2007), we get a boost in cond. by 2/3 (i.e., it adds 40% on top of ThC input). Dashed line: no IC currents; solid line: $g_{\mathrm{EE}} = 3.9$ mS/cm^2 , $g_{\mathrm{IE}} = g_{\mathrm{EI}} = g_{\mathrm{II}} = 0.0$ mS/cm^2.

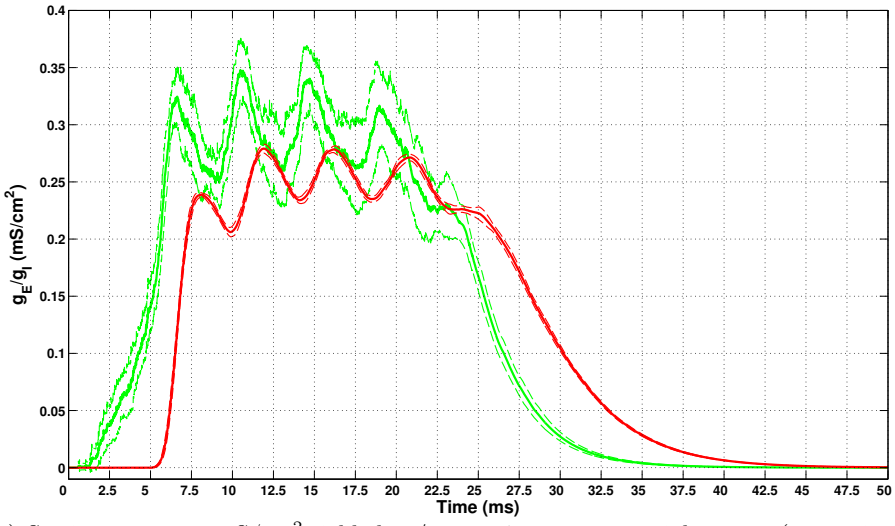
Liu et al. (2007). This is illustrated in figure 3.8, which shows an example trace in response to a single pure tone, with and without intracortical recurrent excitation. However, we later realized we should aim at fitting the IC boost value found by Liu et al. (2007) *globally*, i.e., calculated across the tonotopic axis. We elaborate on this in section 3.3.2.5.

3.3.2.3. Average excitatory and inhibitory conductance responses at φ_0 site and across A1

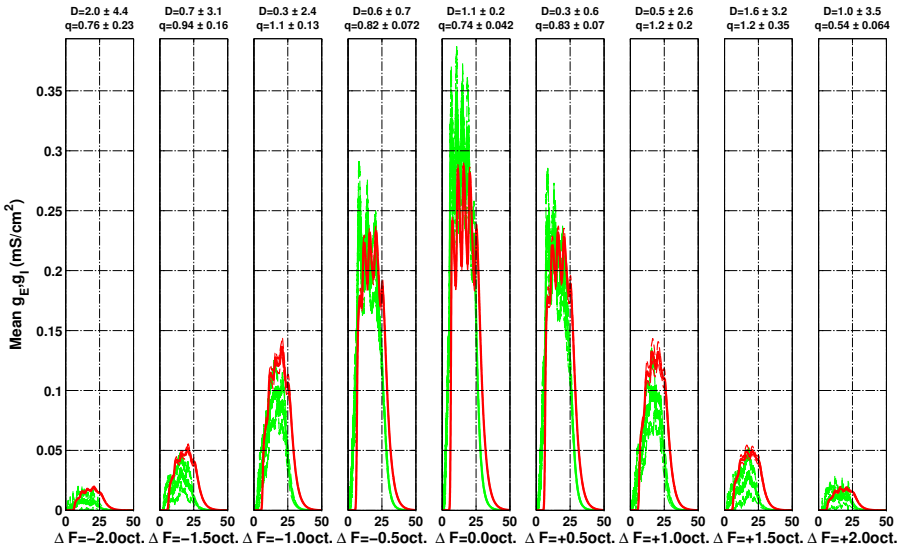
We set g_{EI} with the g_I/g_E relation found in Wehr & Zador (2003) in mind. The value $g_{EI} = 0.8$ mS/cm^2 yielded $g_I/g_E \approx 0.72$ across 10 seeds (cells), at φ_0 (figure 3.9(a)). Across A1, we see how inhibition gradually dominates excitation, as in Wu et al. (2008) (figure 3.9(b)). Subthreshold response width is 4 octaves, within the interval of 4-5 octaves obtained if one combines data found by Liu et al. (2007) and Wu et al. (2008).

3.3.2.4. IC excitation boosts spike response amplitude and IC inhibition narrows its width

We examined the effect of the conductance boost on the spike response of our network. We find that IC excitation slightly boosted response



(a) Setting $g_{\mathrm{EI}} = 0.8$ mS/cm^2 yielded $g_I/g_E \approx 0.72$ across 10 trials, at φ_0 (approximate fit of Wehr & Zador 2003). $g_{\mathrm{EE}} = 3.9$ mS/cm^{2}and $g_{\mathrm{IE}} = g_{\mathrm{II}} = 0.0$. mS/cm^2.



(b) Parameters as in (a). Across A1, we see how inhibition gradually dominates excitation toward the flanks, as in Wu et al. (2008). Subthreshold response width is 4 octaves, within the interval of 4-5 octaves obtained if one combines data found by Liu et al. (2007) and Wu et al. (2008). Above each subplot is shown the mean delay (D) and quotient (q) of inhibition relative to excitation at that point of the TT axis.

Figure 3.9: Conductance (red: inhibition; green: excitation) response at φ_0 (a) and across TT axis (b). Details in subcaptions.

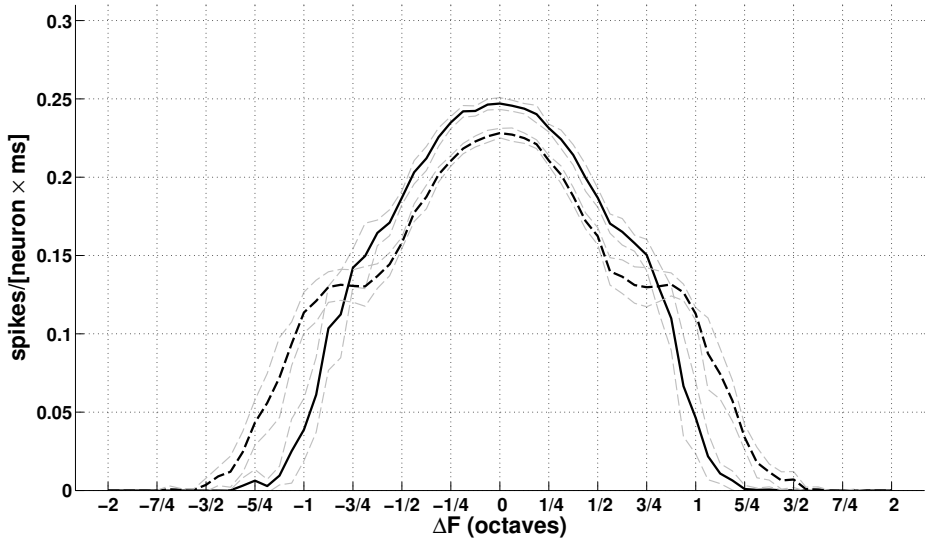


Figure 3.10: The figure shows the change in the spiking profile (histogram smoothened by Gaussian with $\sigma_G = 2.5$ ms) when contrasting $g_{\mathrm{EE}} = g_{\mathrm{EI}} = 0.0$ (dashed) with $g_{\mathrm{EE}} = 3.9$ and $g_{\mathrm{EI}} = 0.8$ (solid; units: mS/cm^2). IC excitation slightly boosted response amplitude, while IC inhibition narrowed the response profile width to two octaves, approximately corresponding to the data found by Liu et al. (2007).

amplitude, while IC inhibition narrowed the response profile width to two octaves (figure 3.10), as seen in Liu et al. (2007). Dashed bold line shows response with no IC currents, solid bold line shows response with $g_{\mathrm{EE}} = 3.9$ mS/cm^2 , $g_{\mathrm{EI}} = 0.8$ mS/cm^2 ($g_{\mathrm{IE}} = g_{\mathrm{II}} = 0.0$ mS/cm^2).

3.3.2.5. Elaborately reproducing data with full network setup

So, having achieved the approximate fits described up until now, although acceptable, we felt the lack of realism in not exploring what role g_{IE} and g_{II} conductances play in the intracortical response warranted a more rigourous approach to the issue of setting the connectivity. Therefore, we elaborated the goal function described in section 3.2.2 and proceeded to try to minimize its value, GF_{val}, while keeping an eye on individual values that were being fit in order to prioritize the most important characteristics (more than the goal function already does). While this was an arduous trial-and-error process, we will proceed directly to describe the similar fits of three different parameter setups. The different setups are defined in table 3.1, and the fits are described in detail with respect to important

characteristics of sub- and suprathreshold responses in tables 3.2 and 3.3.

Setup	μ	λ_E	λ_I	g_{EE}	g_{EI}	g_{IE}	g_{II}	$g_{\mathrm{CTh}}^{\mathrm{IEratio}}$
1	0.4	0.03	0.3	6.0	1.6	3.1	0.6	1.0
2	0.4	0.03	0.3	6.0	1.6	3.1	**1.2**	**1.5**
3	0.4	0.03	0.3	6.0	1.6	3.1	**1.3**	1.5

Table 3.1: The different parameter settings investigated. From one setup to the next, changing parameter values are in bold.

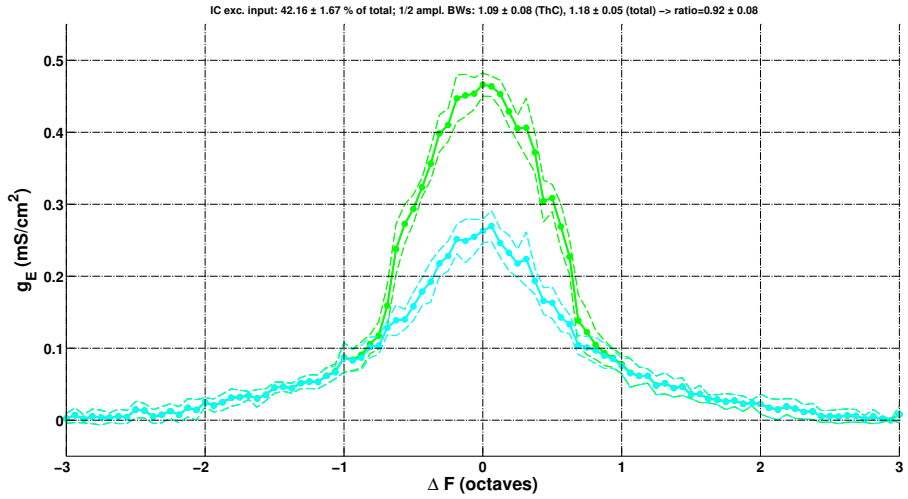
Setup	$\frac{100\times g_E^{\mathrm{IC}}(E)}{g_E^{\mathrm{ThC+IC}}(E)}$ (%)	$\frac{BW^{50}_{g_E^{\mathrm{ThC}}(E)}}{BW^{50}_{g_E^{\mathrm{ThC+IC}}(E)}}$
1	40.99 ± 1.50	0.85 ± 0.10
2	40.89 ± 1.70	0.93 ± 0.08
3	42.16 ± 1.67	0.92 ± 0.08

Table 3.2: IC portion of input (%) and ratio of BW of ThC to total (IC+ThC) input, for the different parameter settings.

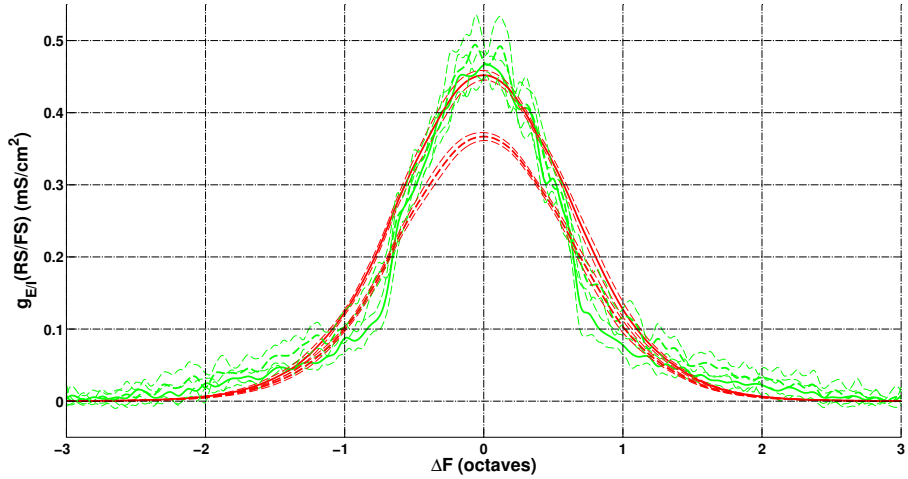
Setup	$BW_R^0(E)$	$BW_R^0(I)$	$BW_{V_m}^0(E)$	$BW_{V_m}^0(I)$	GF_{val}
1	2.23 ± 0.05	2.29 ± 0.15	4.61 ± 0.21	4.35 ± 0.33	310.60
2	2.11 ± 0.20	2.84 ± 0.15	4.97 ± 0.19	4.99 ± 0.18	288.98
3	2.11 ± 0.21	2.82 ± 0.15	4.97 ± 0.19	4.99 ± 0.18	283.96

Table 3.3: Bandwidth statistics for spike and membrane voltage responses for the different parameter settings.

As can be seen in the tables, the GF_{val} quantities are similar. In fact, most of the data we include in the statistics tables is similar, the only significant difference being that setup 1 has approximately equal bandwidths for suprathreshold responses for excitatory and inhibitory neurons, whereas for setups 2 and 3 these bandwidths differ substantially, as seen in Wu et al. (2008). However, in order to get that result, we needed to introduce $g_{\mathrm{CTh}}^{\mathrm{IEratio}} > 1.0$, i.e., set g_{CTh} for inhibitory neurons stronger by that relative factor (in this case, 1.5). This is in accordance with recent results in A1 (Levy & Reyes 2012, Schiff & Reyes 2012), as well as studies in the somatosensory cortex (Gabernet et al. 2005, Cruikshank et al. 2007, Daw et al. 2007, Hull et al. 2009) (though *not* in agreement, however, with results of Rose & Metherate 2005). Both setup 2 and setup 3 feature this stronger input to inhibitory neurons, but in the end we chose setup 3 as our best fit, due to its lower GF_{val} value. As so often in neuronal modelling,



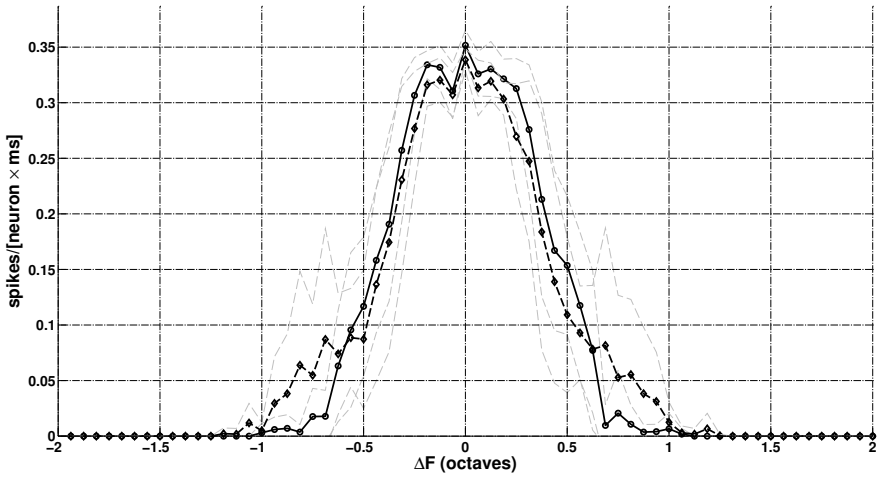
(a) Excitatory conductance change in an RS exc. neuron when adding IC currents (cyan: $g_{\mathrm{EE}} = g_{\mathrm{EI}} = g_{\mathrm{IE}} = g_{\mathrm{II}} = 0.0$ mS/cm^2 ; green: IC conn:s as in setup 3 of table 3.1).



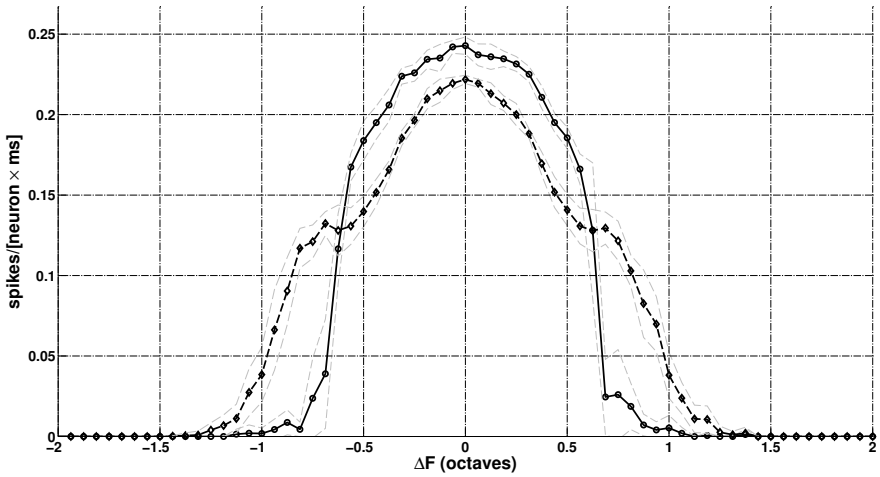
(b) All conductances (green: exc.; red: inh.) of RS exc. (solid) and FS inh. (dashed) neurons, respectively. IC conn:s as in setup 3 of table 3.1.

Figure 3.11: This figure shows the conductance response to a 25 ms tone. In (b), all conductances are shown (green: excitatory; red: inhibitory; dashed: of FS neuron; solid: of RS neuron), whereas in (a), we see how adding intracortical activity boosts RS neuron excitatory conductance response by approximately 2/3 (title shows IC portion of input is about 40 % of total), and bandwidth is slightly broadened at 50% amplitude (title data), but clearly narrowed if measured at lower amplitude levels.

this result is not to be seen as a unique solution, but rather one which fits neurophysiological data to a high degree and is chosen rigourously, lending it credibility and plausibility. We now illustrate the characteristics of the response to a single tone (using setup 3 for parameter values) in a series of figures. Please note that thinner lines tracing the thick, principal lines of the figures, always designate the standard deviation.



(a) Spiking response to 25 ms tone with $\cos^2$(t) ramps. Here, $\sigma_G = 1.0$ ms.



(b) Spiking response to 25 ms tone with $\cos^2$(t) ramps. Here, $\sigma_G = 2.5$ ms.

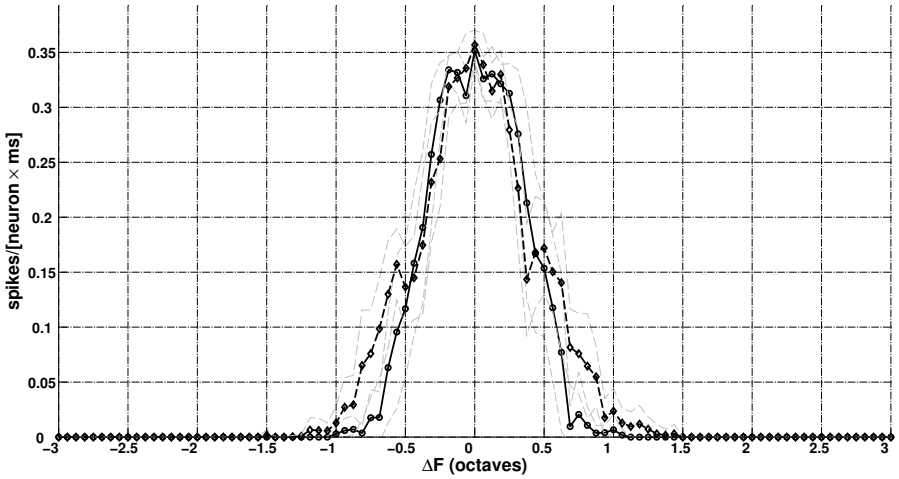
Figure 3.12: The figure shows the change in the exc. neuron spiking histogram profile when contrasting $g_{\mathrm{EE}} = g_{\mathrm{EI}} = g_{\mathrm{IE}} = g_{\mathrm{II}} = 0.0$ mS/cm^2 (dashed) with $g_{\mathrm{EE}} = 6.0$, $g_{\mathrm{EI}} = 1.6$, $g_{\mathrm{IE}} = 3.1$, $g_{\mathrm{II}} = 1.3$ (solid; units: mS/cm^2), for two different σ_G values: 1.0 ms (a) and 2.5 ms (b). As in figure 3.10, IC excitation slightly boosted response amplitude, while IC inhibition narrowed the response profile width to slightly less than two octaves, approximately corresponding to the data found by Liu et al. (2007).

Figure 3.11 is included to show how adding intracortical activity boosts excitatory response by about two thirds, as seen in Liu et al. (2007), this time *across* the tonotopic axis (in contrast with figure 3.8, only fit locally). Individual strengths and widths of evoked conductances in both neuron types are shown in subfigure (b), whereas in subfigure (a) the RS neuron excitatory conductance evoked using setup 3 (green line) is contrasted with the same response in the complete absence of intracortical currents (cyan line).

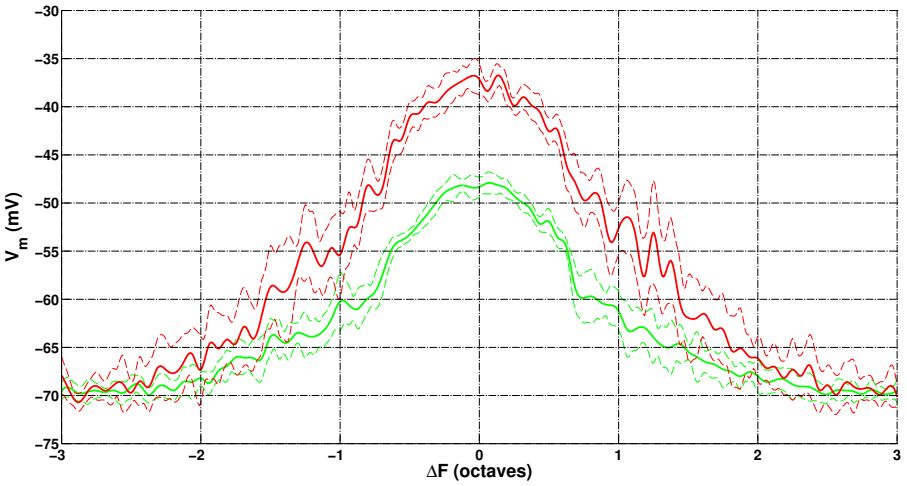
Figure 3.12 goes on to show how the spike response of excitatory neurons in the network is affected by adding intracortical connections (the two subfigures show the same data for two different widths of the Gaussian filter used on the histograms of spikes). We see how there is a slight boost also in suprathrehold response in the center of the TT axis, while on the flanks there is narrowing of the spiking profile due to lateral inhibition (especially clear in subfigure 3.12(b)).

Furthermore, in figure 3.13 we may observe the spiking response and the underlying membrane potential response (action potentials blocked) of both excitatory RS neurons and of inhibitory FS neurons. The stronger input to inhibitory neurons ($g_{\mathrm{CTh}}^{\mathrm{IEratio}} = 1.5$) causes a stronger subthreshold response (subfigure 3.13(b)), and a slightly wider suprathreshold response (subfigure 3.13(a)), in those neurons.

To illustrate the fit of a part of our goal function (section 3.2.2) not included in the statistics of tables 3.2 and 3.3, namely the relationship of inhibition and excitation across the tonotopic axis, we made figure 3.14, which shows our fit in (b) and the original data of Wu et al. (2008) in (a). As commented on in the subcaptions, we manage to fit two important characteristics of the data, namely the domination of inhibition on the flanks of BF, and inhibition being narrower than excitation, on average, at the base of the responses (at 0% amplitude).

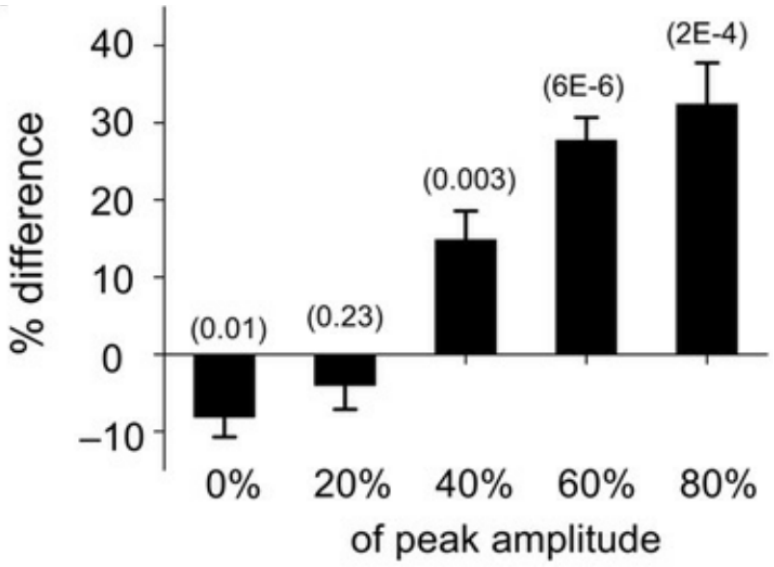


(a) Spiking response in the form of smoothed histograms (σ_G=1.0 ms; solid: RS exc.; dashed: FS inh.). BW at base is wider for inhibitory neurons, as seen in neurophysiological data (Liu et al. 2007, Wu et al. 2008).

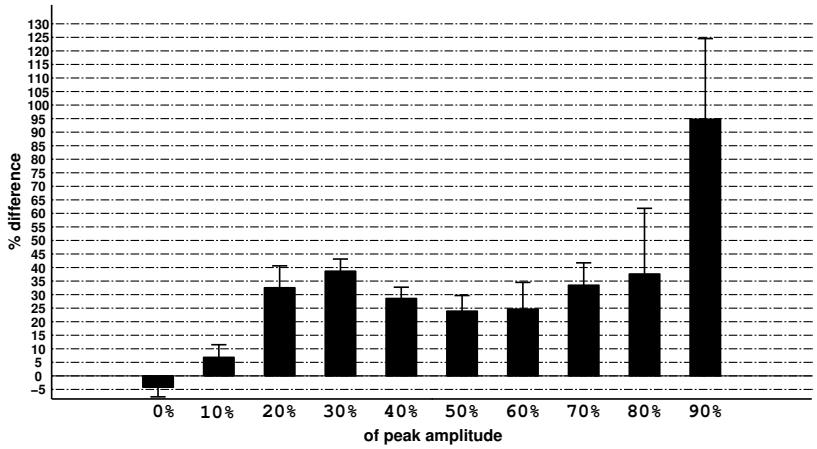


(b) Membrane potential response (spikes blocked) in inhibitory neuron (red) and excitatory neuron (green). Note the stronger response in inhibitory neurons.

Figure 3.13: In (a), we show the spike response of excitatory (solid) and inhibitory (dashed) neurons to a single 25 ms tone, whereas in (b), we show the membrane potential underlying this spike response (spikes blocked for this figure). We observe how the stronger input to inhibitory neurons renders their subthreshold response stronger, in (b), and their suprathreshold response slightly wider, in (a).



(a) The original data (adapted from figure 3A in Wu et al. (2008)) shows how inhibition is wider than excitation around BF, but actually narrower at 0-20% amplitude (average over 11 cells).

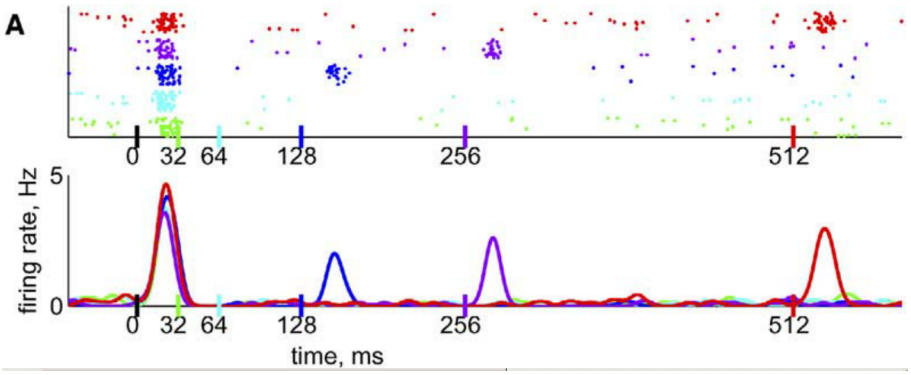


(b) Our fit displays the trend of the original data, showing inh. narrower than exc. at the base and increasingly wider the closer to BF one measures (average over nine seeds).

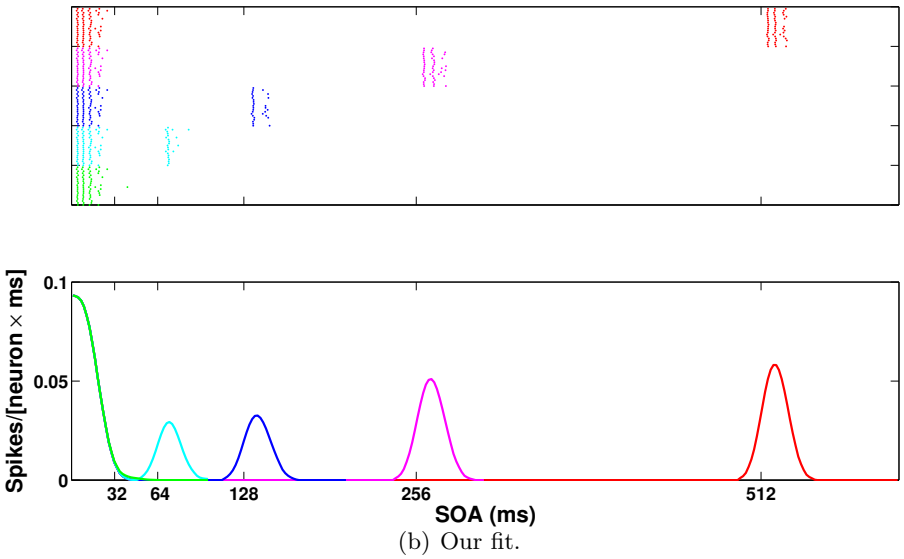
Figure 3.14: Average percentage bandwidth difference ((BW_I - BW_E)/BW_I), at different amplitude levels, between excitatory and inhibitory response in exc. (RS) neuron. In (a), we reproduce the original data (Wu et al. 2008), in (b) we show our fit, using setup 3 of table 3.1. Comments in subcaptions.

3.3.3. Simulating neurophysiological forward masking

We now proceed to attempt to reproduce the results by Wehr & Zador (2005), who investigated intracellularly the response of neurons in anaesthetized (ketamine) rat A1 to pairs of noise clicks with different *SOA* values (see Introduction). The clicks had an intensity of 102 dB, and for that reason and due to the stimulus being noise we increase the input by a factor of 2.4 (explained in appendix B).



(a) Original data, adapted from figure 1A of Wehr & Zador (2005).



(b) Our fit.

Figure 3.15: This figure shows the example raster and histogram from Wehr & Zador (2005) (a), and our fit (b).

For these simulations we used $g_{EE} = 3.9$ mS/cm^2 and $g_{EI} = 1.6$

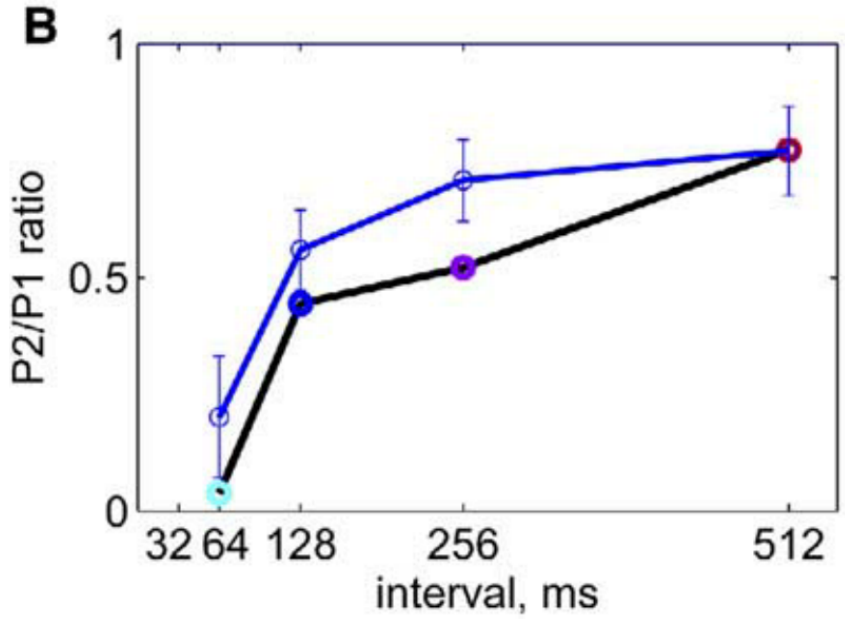
mS/cm^2 (inhibition double that of our approximate fit of before), whereas $g_{\mathrm{IE}} = 0.0$ mS/cm^2 and $g_{\mathrm{II}} = 0.0$ mS/cm^2. Depression parameters were $U = 0.6$ and $\tau_D = 500$ ms, also adjusted upwards from before. We obtain the results of Wehr & Zador (2005) as regards time course of recuperation from a masker tone. To illustrate this fact, we have prepared figures 3.15 and 3.16. Figure 3.15 shows (both original data in (a) and our fit in (b)) a sample raster plot and an averaged histogram of the responses to both masker and probe, for all *SOA* values explored. From comparing the subfigures it is clear that we reproduce the general trend in the data, although we have a bit more response for *SOA*= 64 ms than did Wehr & Zador (2005). However, this anomaly was within the error margin of their data, which we can appreciate when scrutinizing the average response ratios (Probe/Masker) depicted in figure 3.16.

Finally, we show evoked excitatory and inhibitory conductances in figure 3.17, where we can appreciate that both conductance types recuperate at about the same rate, as seen in Wehr & Zador (2005).

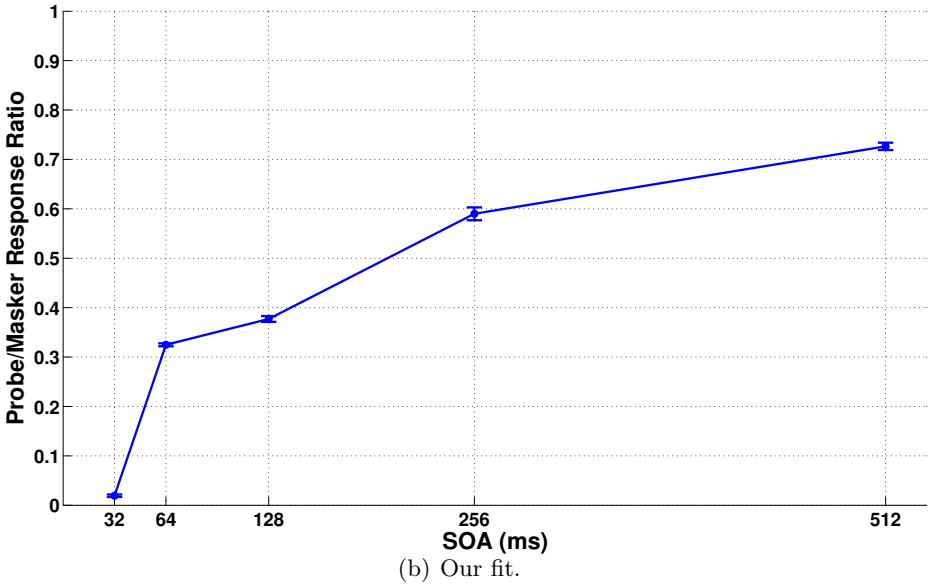
These results were obtained with slightly raised parameter values as compared to our first fitting of neurophysiological results with pure tones (section 3.3.2), as regards inhibition and depression. The need for this could be due to different preparations, different anaesthesia or other factors, and will need more work to be clarified. Furthermore, $g_{\mathrm{I}}/g_{\mathrm{E}}$ ratios are not a perfect fit to the data, which potentially raises issues (also, the suppression of the conductances is stronger for low *SOA* values in the neurophysiological data of subfigure 3.17(a), but we emphasize that this is data from an example cell). Finally, we would ideally like to reproduce the data using our full network fit as presented in this chapter, but it will have to be deferred to future work for lack of time. Nevertheless, we find that showing the gist of the data on time course of recuperation of both activity and conductances with a tonotopic network model is already a considerable achievement.

3.3.4. Auditory streaming correlates revisited

In chapter 2, we reproduced the main data of Fishman et al. (2004) using three different models. In particular, we showed that the same underlying mechanism to explain differential suppression worked in all three of them. There are however some shortcomings to the *local* approach we employed there (i.e., representing only a population of neurons with BF f_A). First of all, that solution only partly honours the fact that, as in every subcortical station of the auditory pathway, A1 displays a tonotopic organization.

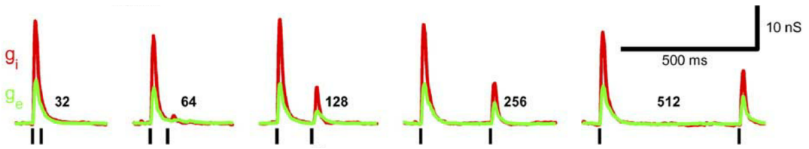


(a) Original data, adapted from figure 1B of Wehr & Zador (2005).

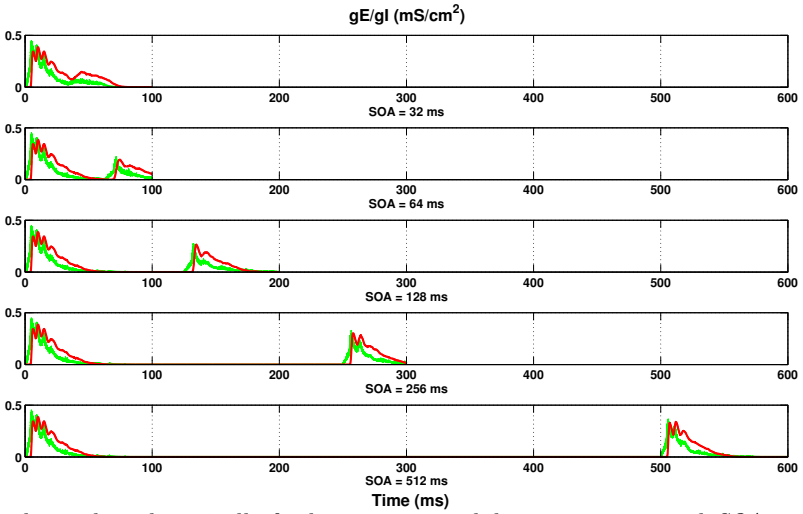


(b) Our fit.

Figure 3.16: This figure shows average P_2/P_1 ratio found by Wehr & Zador (2005) (blue line in (a)) and our fit of the same (b). Our average is taken only over two different seeds, hence the small s.d.



(a) Original, adapted from figure 2C of Wehr & Zador (2005).



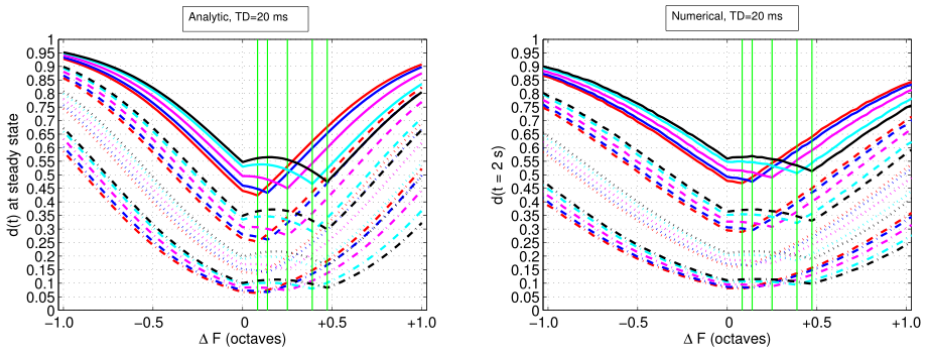
(b) Our data, plotted vertically for better comparability across time and *SOA*

Figure 3.17: The figures show the time course of excitatory and inhibitory conductances for all *SOA* values. (a): original. (b): our data

Secondly, we did not explore the role of introcortical connections (i.e., currents) in shaping the response to auditory streaming stimuli at this level.

Both these issues are addressed here, employing the full network setup (parameter setup 3) from above.

Figure 3.18 shows the so-called depression field to illustrate what was neglected in our local network. Particularly, the different level of depression across the TT axis means the response off-BF can be considerable and can have an influence on response at BF, especially when we set $\nu > 0$.



(**a**) (**b**)

Figure 3.18: Analytically and numerically obtained so-called synaptic depression fields (d_0, i.e., the steady state value of $d(t)$ right before an A tone is about to be presented, across the tonotopic axis). $PR = 5, 10, 20$ and 40 Hz are represented by solid, dashed, dotted and dotted-dashed lines, respectively, whereas $\Delta F = 5, 10, 20, 30$, and 40 % are represented by red, blue, magenta, cyan and black lines, respectively. The green vertical lines mark the corresponding F_B locations on the tonotopic axis. (a) Curves obtained using analytical results of appendix A. (b) The steady-state values of $d(t)$ (after a 1200 ms AB tone sequence), averaged over $K = 100$ synapses at one neuron, for all positions in an $N = 73$ column network. For both figures, $U = 0.4$ and $\tau_D = 250$ ms.

3.3.4.1. Full tonotopic axis of neural populations, with intracortical connections: Recurrent network

As mentioned before, excitation was taken to be very narrow, whereas inhibition is quite a bit broader (here by a factor of ten), as is derivable from data found by Liu et al. (2007) and Wu et al. (2008). With these settings (all parameter values in table 3.4), we got a good fit in the mean of the B/A response ratios seen in the study by Fishman et al. (2004). Figure 3.19 depicts the response ratios for 64 $\Delta f\,/PR$ parameter combinations. Due to the introduced variability, calculating an average is necessary (reflecting biological reality, where noise is ever-present; e.g., Shadlen & Newsome 1998). Here we used ten different seeds for each data point shown, meaning we ran a total of 640 simulations to get the data of figure 3.19.

Name	Value	
R_{max}^{Th}	100	Hz
μ	0.4	octaves
ν	0.33	octaves
U	0.4	
τ_D	500	ms
g_{CTh}	0.02 ± 0.004	$\mathrm{mS/cm^2}$
$K_n(K)$	10(110)	
M	15	
N	97	
g_{EE}	6.0	$\mathrm{mS/cm^2}$
g_{IE}	3.1	$\mathrm{mS/cm^2}$
α_E	2.0	$\mathrm{ms^{-1}}$
β_E	0.5	$\mathrm{ms^{-1}}$
λ_E	0.03	octaves
g_{EI}	1.6	$\mathrm{mS/cm^2}$
g_{II}	1.3	$\mathrm{mS/cm^2}$
α_I	0.5	$\mathrm{ms^{-1}}$
β_I	0.3	$\mathrm{ms^{-1}}$
λ_I	0.3	octaves

Table 3.4: Values of biophysical parameters for recurrent network

In figure 3.20, we show the same data as in figure 3.19, but superimposed on a so-called van Noorden diagram (cf. figure 1.4(d), reproducing figure 7(a) of Fishman et al. 2004). As can be appreciated, the trend in the model

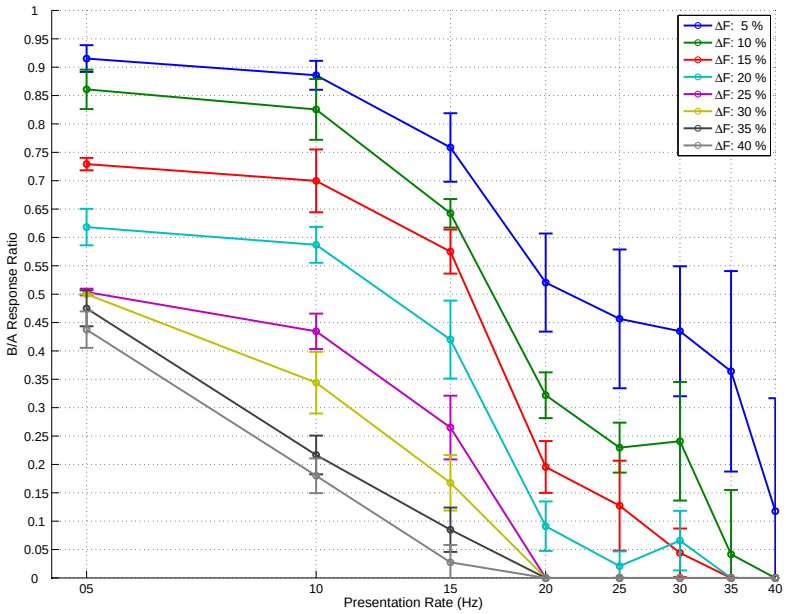


Figure 3.19: B/A response ratios plotted as a function of PR , for 8 different Δf values. Shown are the average values (error bars are standard deviations) across 10 different simulations, varying the random seed. Parameters as in table 3.4.

data compares well to that of the neurophysiological data obtained in Fishman et al. (2004). Most importantly, we used thalamocortical axonal spread in these simulation ($\nu = 0.33$ octaves) and managed to get the data fit as found again, in stark contrast with what we found with the feed-forward network (compare, respectively, figure 3.19 to figure 1.4(a), and figure 2.24 of chapter 2 to figure 1.4(b)).

3.3.4.2. Illustrating the build-up of streaming dynamically: Scientific movies

In order to further illustrate the results of our modelling of the underlying neural correlates of the auditory streaming phenomenon, we have prepared scientific movies of the evolution of neural activity over time. In the web page http://cns.upf.edu/johan/EmCAP/Movies/MoviePage.html we show a well-chosen sample of these. In particular, we chose five illus-

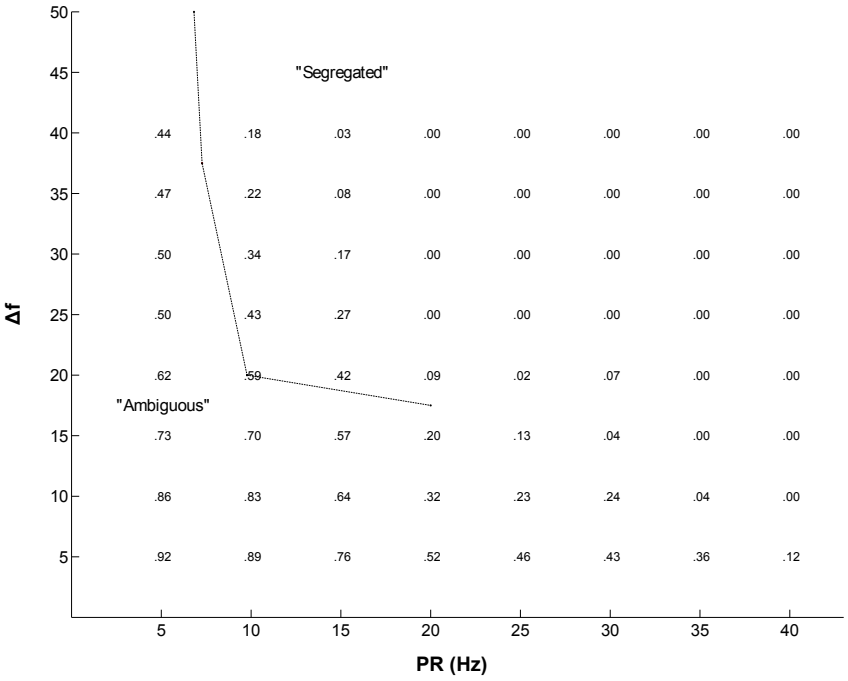


Figure 3.20: Plot showing average B/A response ratios (over 10 model 'trials'; same parameter values as in figure 3.19) superimposed on a van Noorden diagram (cf. figure 1.4(d), reproducing figure 7(a) of Fishman et al. 2004). The dotted line is the *temporal coherence boundary* (TCB), as defined in (McAdams & Bregman 1979, Fishman et al. 2004). Setting a suitable threshold value, based on the quotients' values close to the TCB, one can make a putative interpretation of whether the stimulus stream would yield a segregated or an ambiguous 'percept'. Quotation marks are to indicate the speculative nature of any interpretation of results as directly pertaining to actual perception.

trative points in the PR -Δf parameter space, three of which purportedly show the evolution of neural activity to a segregated 'percept' and two which show the same process for a so-called ambiguous 'percept' (the quotes merely illustrate the speculative nature of discussing percepts in a neural modelling context). All of these points are shown below (figure 3.21) in a diagram similar to the one by van Noorden (1975) (reproduced in our figure 1.1). The line is the temporal coherence boundary found by van Noorden. The fifth point ($PR = 40$, $\Delta f = 20\%$) is outside of the range of the boundary, but is nevertheless added as a clear illustration

of how a segregated 'percept' is formed in the A1 model. In all movies, first 1.25 seconds of the AB tone sequence is presented, followed by 0.8 seconds of silence, which in turn is followed by a further 0.45 seconds of AB tone sequence. During the first 1.25 seconds, one may observe the so-called build-up of the respective 'percepts'. After that, the silent gap is inserted to illustrate how the 'percept' arrived at can change as the network 'recuperates'. Then, when the sequence is turned on a second time, we show how the 'percept' reapproaches the state it was in before the silent gap. This is especially clear in the case of points 2 and 3, which illustrate segregation.

In the web page, there are links to two movies for each of the five points chosen. For the benifit of the reader, here we will reproduce snapshots of the final frames of the movies corresponding to points 3, 4 and 5 of figure 3.21, although we recommend a visit to the movie page for the added illustrative power of dynamics. We also reproduce here the relevant part of the explicative text which can be found by clicking on any 'Back' link on the movie page, or alternatively going directly to its mother page, http://emcap.iua.upf.edu/Emergent Cognition.html.

- The tone duration used was 20 milliseconds (ms). A tones were presented at the center of the tonotopic axis (at frequency F_A), whereas B tones were presented Δf octaves above the A tone (at F_B) The value of Δf and PR are given by the respective points in the van Noorden diagram. In both movies, quantities associated with A tones are marked in blue and those associated with B tones are marked in red.

- The first movie shows the so-called raster diagrams evolving over time, with the lower panel giving an overview of the whole 2.5 seconds, and the upper giving a 250 ms blow-up of the evolving activity. Each green dot represents one excitatory neuron emitting a spike.

- The second movie displays three things simultaneously, the main plot and two inset plots. The main plot shows how the the activity evolves over time, with the x-axis representing the tonotopic axis. The y-axis shows response in spikes per neuron each ms. During each tone presented, a vertical line with the tone colour (A: blue; B: red) is displayed at the tone's corresponding point on the tonotopic axis. The profile of the response to every tone is 'frozen' at its maximum, until the next tone with the same frequency is presented. When a tone is presented, the outline of the previous one becomes dashed. As for the insets,

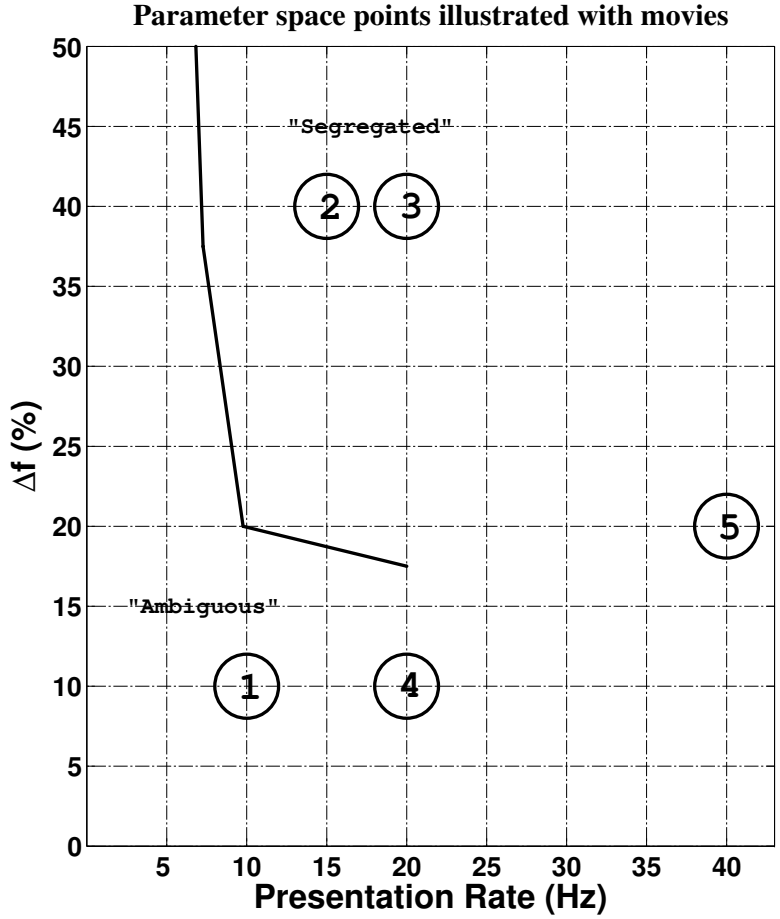


Figure 3.21: The diagram shows which five parameter settings were used for the movies in the web page http://cns.upf.edu/johan/EmCAP/Movies/MoviePage.html. For illustrative purposes, we include a reproduction of the temporal coherence boundary from McAdams & Bregman (1979), and the terms "Ambiguous" and "Segregated" in their corresponding parameter space, as per psychophysical results described in the Introduction (fission boundary not shown).

- the right inset plots how the time-averaged response to A and B tones at the A1 site with Best Frequency (BF) f_A (i.e., the center of the tonotopic axis) evolves over time.
- The left inset plots how the quotient of the response to a B tone to that of the preceding A tone (B response divided by A response, measured at φ_A) evolves over time. The horizontal dotted line is an estimated threshold value for determining whether the 'percept' is segregated or ambiguous (not for point 5).

In figures 3.22, 3.23 and 3.24 we show snapshots of the last frame of the movies corresponding to points 3, 4 and 5 in the diagram of figure 3.21, respectively.

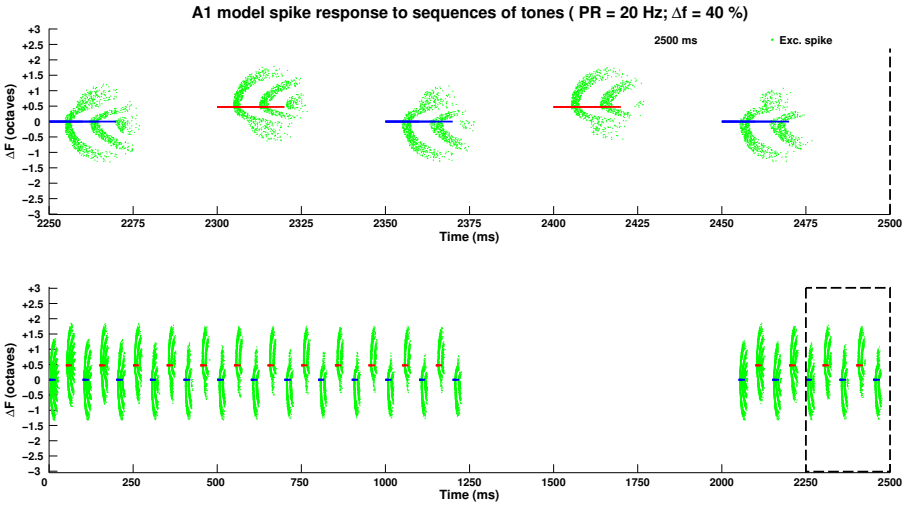
3.3.5. Robustness tests

3.3.5.1. Making inhibitory currents last longer has a transient effect

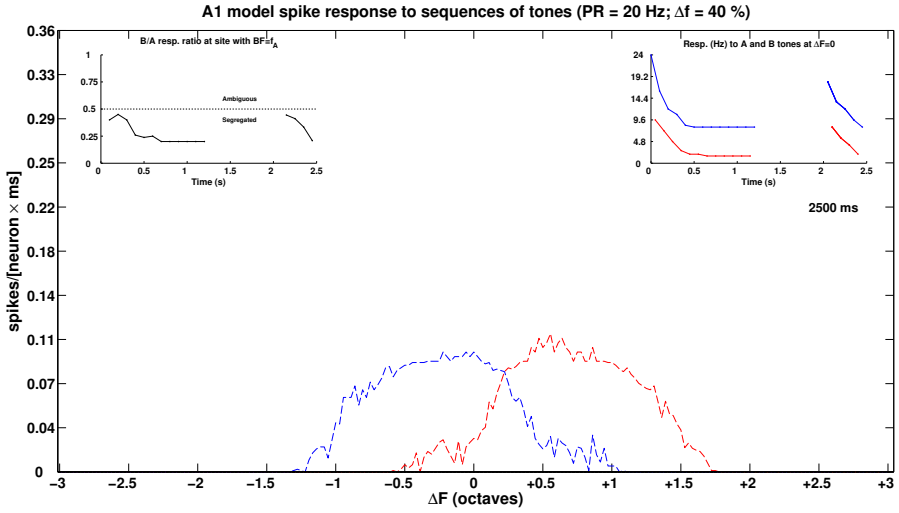
Looking at responses to long tones (figures 3.25(a) and 3.25(b)), one sees that the stronger the inhibition (increasing from top to bottom), the more response narrows as input progresses (top panel: RS *inh.* neurons; middle panel: FS inh. neurons, $\alpha_I = 2.0$ ms^{-1}, $\beta_I = 0.5$ ms^{-1}; bottom panel: FS inh. neurons, $\alpha_I = 1.0$ ms^{-1}, $\beta_I = 0.1$ ms^{-1}). However, most of the narrowing takes place during the first 20-30 milliseconds, which means using solely inhibition as a mechanism for differential suppression is unfeasible, contrary to claims in Kanwal et al. (2003), at least without some additional inhibitory or otherwise suppressing mechanism added to the network setup used to make these figures. This result of course speaks in favour of our use of synaptic depression in thalamocortical synapses as such an additional or even alternative (as shown in chapter 2) mechanism.

3.3.5.2. Modulating input with steady-state depression values can replace depression

Interestingly, one sees that making this change has a very small effect on responses of the network to sequences of tones. This shows that the term 'differential suppression' is possibly ill-chosen, since we could get that same effect by replacing specific depression field values just before A and B tones with their average, and using this average to down-modulate input rates while shutting down depression, before running the simulation. This lends further strength to our hypothesis that the key to getting these responses is

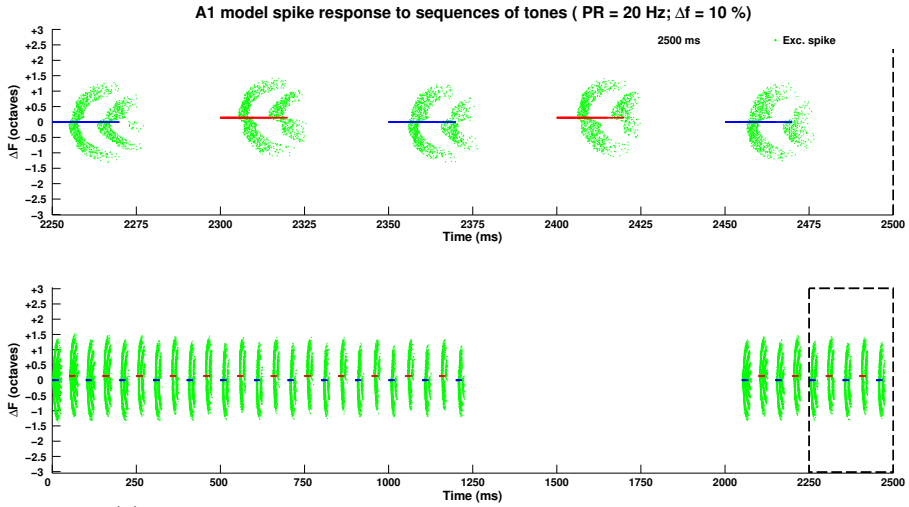


(a) Raster of excitatory neurons' spikes across the tonotopic axis.

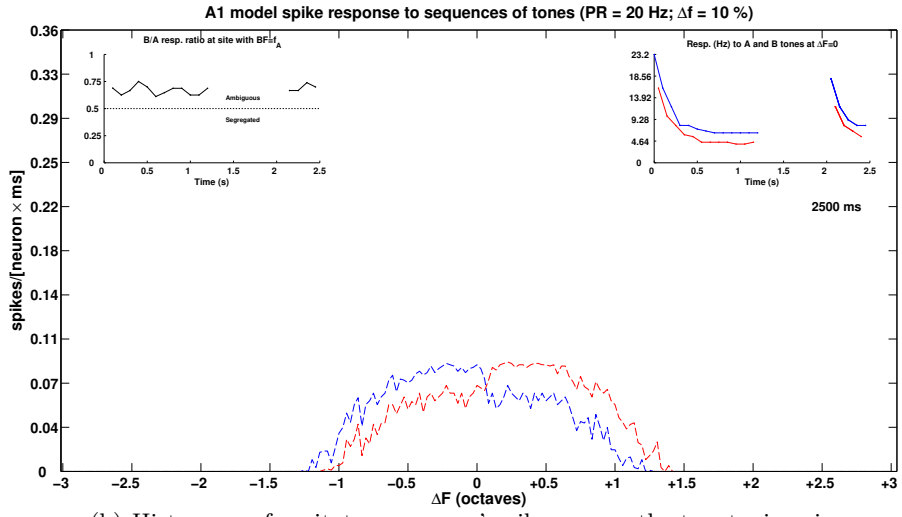


(b) Histogram of excitatory neurons' spikes across the tonotopic axis.

Figure 3.22: Illustration of point 3 in figure 3.21 (state of network *after* full simulation - for dynamics, see movie page). Insets in (b): Responses to A (blue) and B (red) tones at BF=F_A site (right) and the ratio of those responses (left). Scheme: 1250 ms sequence - 800 ms silence - 250 ms sequence.

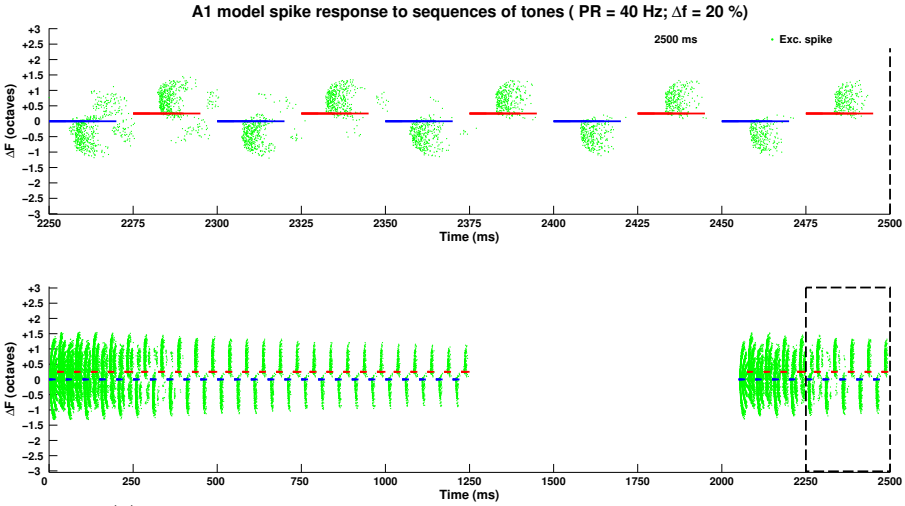


(a) Raster of excitatory neurons' spikes across the tonotopic axis.

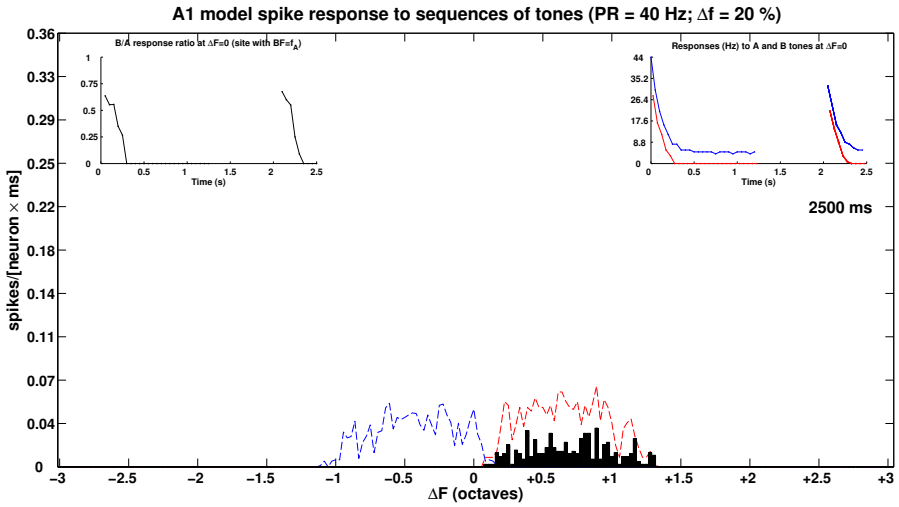


(b) Histogram of excitatory neurons' spikes across the tonotopic axis.

Figure 3.23: Illustration of point 4 in figure 3.21 (state of network *after* full simulation - for dynamics, see movie page). Insets in (b): Responses to A (blue) and B (red) tones at BF=F_A site (right) and the ratio of those responses (left). Scheme: 1250 ms sequence - 800 ms silence - 250 ms sequence.

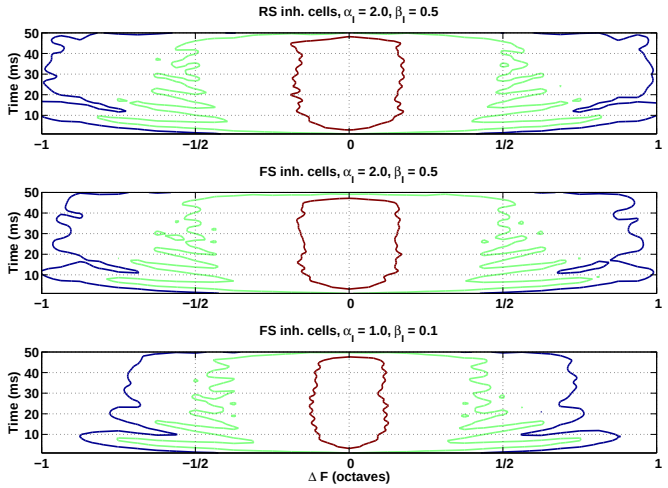


(a) Raster of excitatory neurons' spikes across the tonotopic axis.

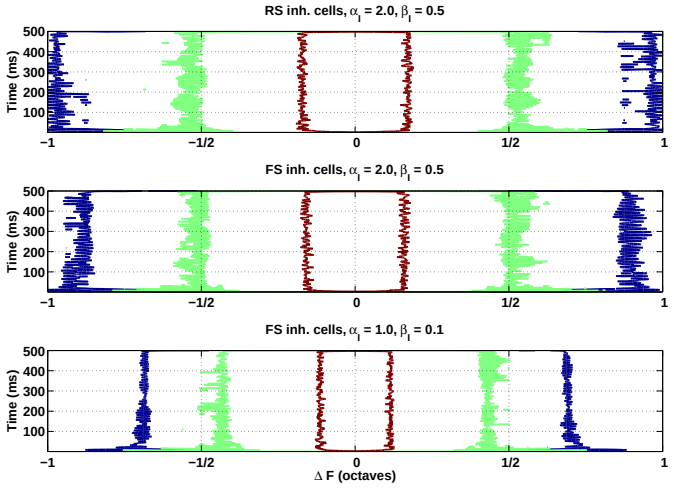


(b) Histogram of excitatory neurons' spikes across the tonotopic axis.

Figure 3.24: Illustration of point 5 in figure 3.21 (state of network *after* full simulation - for dynamics, see movie page). Insets in (b): Responses to A (blue) and B (red) tones at BF=F_A site (right) and the ratio of those responses (left). Scheme: 1250 ms sequence - 800 ms silence - 250 ms sequence.



(a) Contour plot of responses to centered tone of $TD = 500$ ms (first 50 ms shown).



(b) Contour plot of responses to centered tone of $TD = 500$ ms (whole duration shown).

Figure 3.25: Contour plots of responses to tone of $TD = 500$ ms (first 50 ms shown in (a), full duration in (b)). Colour code: red=0.25, green=0.15 and blue=0.05 spikes/[neuron $\times$ ms]. Here, $N = 33$, covering 2 oct., $U = 0.0$ (no depression), and inhibition is changed via α_I , β_I and type of neuron (subplot titles). All other parameters standard. Response narrows as inhibition is increased, but effect is concentrated to initial part of the stimulus (subfigure (a)) and is greatest for longer-lasting currents ($\beta_I = 0.1$ ms^{-1}).

in the response threshold of cortical neurons and synaptic depression, which is the only mechanism that can reduce the initial response of 4-6 spikes per neuron to only 0-1 spikes per neuron over a time-scale approximately corresponding to the build-up of auditory stream segregation (Bregman 1978, Anstis & Saida 1985, Cusack et al. 2004, Micheyl et al. 2005).

3.4. Discussion

In this chapter we find that a mexican-hat type IC connectivity can be used to explain approximate co-tuning of inhibition and excitation as well as flank dominance of inhibition over excitation (Wehr & Zador 2003, Wu et al. 2008). Furthermore, across the tonotopic axis we get a 67% boost of excitatory conductances (equivalent to currents), and also reproduce the spread of sub- and suprathreshold activity, as seen in data found by Liu et al. (2007). All this was achieved by a meticulous process of constraining our network setup *bottom-up* (section 3.3.1) and then examining in detail which network ingredients were necessary and/or sufficient to reproduce the response of A1 to the presentation of a single short (25 ms) tone (section 3.3.2). In the latter section we first found a basic fit using only feed-forward inhibition and recurrent excitation, and then proceeded to make a rigourous fit with all possible intracortical connections active, using for our guidance a goal function and a least-squares procedure (as defined in section 3.2.2).

Moreover, with some minor tweaking of parameters, we were then able to use the network fit to neurophysiologal parameters and results attained with a pure tone stimulus, to also simulate physiological forward masking with noise stimuli, following experiments by Wehr & Zador (2005). Specifically, parting from our basic network fit, by raising synaptic depression vesicle depletion rate and setting the recovery time constant to a higher value, as well as raising inhibition, we approximately reproduce results seen in that paper (section 3.3.3). Our interpretation of the role of various network ingredients is the following. ThC synaptic depression explains the overall time course of activity (quotient) recovery as SOA increases. Also, intracortical synapses contribute decisively; excitation gives the necessary boost to ThC-evoked activity for exc. neurons to spike before they are silenced by a combination of synaptic depression and IC inhibition; without inhibition the results can not be attained, if contraints are to be respected.

However, more work needs to be done to clarify why we need higher U and τ_D values to get results as in Wehr & Zador (2005), compared to the values fit to Rose & Metherate (2005); which exact mechanism is responsible for the two to three times higher inhibitory conductances seen in response to clicks as opposed to tones (seen in *example* traces in Wehr & Zador 2003, Wehr & Zador 2005); and finally, if we need to include currently left-out ingredients in the model (g_{IE} and/or g_{II} nonzero, IC depression, independent λ_{kj}, $\mathrm{GABA_B}$ currents, N-Methyl-D-Aspartate (NMDA) currents, etc.). For starters, we would like to explore how our rigourous network fit of chapter 3 would 'perform' in simulations of neurophysiological forward masking

(not included here for lack of time).

Furthermore, some parameter value choices need to be motivated. For instance, the issue of the width of the intracortical excitatory to excitatory connection profile (λ_{EE} in this work) is under debate. We set it to be quite narrow in our efforts to base our model as much as possible on the most recent (at the time) neurophysiological results, especially the studies by Liu et al. (2007) and Wu et al. (2008). There are however conflicting results which advocate a wider such connection profile, which would bring excitation to those parts of the tonotopic axis not receiving thalamic input (specifically, Kaur et al. 2004). Since Liu et al. (2007) showed that thalamocortical input is wide (but see Bartlett, Sadagopan & Wang 2010), we found that in order to get a stable system that fit as much of the data as possible, we needed to use a more narrow profile (something also supported by data in Liu et al. 2007). We stress however that this concerns horizontal connections in layer IV. We are fully aware that in other cortical layers, such as layer II/III, there exist (found in V1 by Holmgren, Harkany, Svennenfors & Zilberter 2003) or are postulated (as a possible explanation of de-/hypersensitation of auditory cortex neurons, see Noreña, Gourévitch, Aizawa & Eggermont 2006) long-range horizontal connections between pyramidal cells.

In this chapter we also sought to test the robustness of our mechanism for the differential suppression phenomenon seen in neurophysiological correlates of auditory streaming (see chapter 2), making even more realistic our model of A1 with Hodgkin and Huxley conductance-based neurons by representing the tonotopic axis explicitly as well as incorporating intracortical connectivity and thalamocortical axonal spread. Here, we have shown that our mechanism indeed not only holds in a more complex network setup, but is improved in the sense that the inclusion of intracortical currents (especially recurrent excitation) again sharpens the frequency tuning which is diluted by adding thalamocortical input spread to a mere feed-forward network, as seen in chapter 2.

As for the saturation of the transfer function used for the rate model (see equation A.2), we did not find that to be a necessary feature of the corresponding function for a HH network. While possible to attain (by driving the neurons so hard that they all respond at their maximum rate, as determined by their refractory periods), it yields unnaturally high firing rates in response to inputs which are supposed to represent perfectly normal auditory stimuli (data not shown).

An issue worth exploring would be how our network responds to longer stimuli, and how this response relates to the results using 50 ms and 100 ms tone durations in Fishman et al. (2004). Particularly the highly marked

differential suppression of responses to B tones at the A site, when the sequence is presented at the maximum possible rate (i.e., 20 Hz and 10 Hz, respectively), would be challenging to explain (see figures 4 and 5 in Fishman et al. 2004).

As to a possible change of results with NMDA receptors at the thalamocortical interface (*possibly* present in adults - see discussion in chapter 5), we have carried out some control simulations in order to look at that issue (data not shown). Since NMDA creates a slow current in response to each tone, which needs about 100-150 ms to reach zero, then if $PR \geq$ 10 Hz the current will consequently not have reached zero when the next tone arrives, which leads to a gradual increase in this component of the input current as the sequence progresses. If no inhibition is present (as in the feed-forward model of chapter 2), this creates a tonic response which is never zero between tones, but might be stronger at onset and weaker at the end of an SOA period. If inhibition is present in the form of $GABA_A$ receptors, even using slow decay times and fast-spiking inhibitory neurons, we see a domination of the NMDA currents over both AMPA and GABA.

Thus, the difference in amplitude of response to A and B tones 'rides' on top of a tonic firing caused by NMDA. This effectively raises responses to all tones, including to those which without NMDA elicited no response (B tones at high PR and Δf values). The RR values are consequently all raised, because the same absolute value is added to nominator and denominator alike. If instead we were to use a difference measure, it would not be affected, but responses would of course not correspond to those seen in Fishman et al. (2004). To deal with this excess in excitation, there are two alternatives:

1. lowering input ($R_{\mathrm{Th}}^{\max}$, K or g_{CTh}). It might also be necessary to lower the proportion of NMDA to AMPA in order for the former not to drown out the latter (in our simulations we used approximately three times stronger input to AMPA than to NMDA).

2. increasing inhibition by either increasing inh. neurons' thalamocortical conductance values (here, $g_{\mathrm{CTh}}^{\mathrm{IEratio}}$ was unity) or adding a form of $GABA_B$ receptor or some sort of extra-synaptic $GABA_A$ inhibition (see Belelli, Harrison, Maguire, Macdonald, Walker & Cope 2009).

We recognize that this issue needs more work in the future, especially if it were to be shown that NMDA inequivocally plays an important role in the experiments we attempt to reproduce.

Finally, while we set our parameters in a rigourous exercise following the literature at that particular moment in time, we are aware of recent

experimental results which are quite different in some respects (Schiff & Reyes 2012, Levy & Reyes 2012). There also exist other modelling studies using different settings (de la Rocha et al. 2008, Levy & Reyes 2011). These discrepancies should be addressed in a future continuation of the present work. It is however notable that Viaene et al. (2011) found that only LIV gets direct driving input from the MGBv, something which could be used to make a distinction between our study and that of Levy & Reyes (2011), especially as regards the detailed connection data used by those authors to model LII/III. Since we model LIV, our focus on other experimental papers when making the network setup is therefore not necessarily an issue.

4 An attractor network model of bilingual language perception data

The work presented in this chapter is published in the Journal of Cognitive Neuroscience[1]. *A follow-up study has also been published in that same journal*[2].

4.1. Introduction

What is the architecture of the speech processing system? How do the different types of lexical and sublexical information interact? For many years it has been accepted that the way the speech perception system deals with the problem of variability is that of 'filtering it out', thus representing information (phonemes, syllables, words) in an abstract code. One problem with such an approach is that of how the linguistic representations can be modified once they are established. Indeed, everyday evidence shows that, for instance, some time after having moved to a new region, where a different dialect is spoken, human beings modify their language production and their speech becomes accented with the new sounds. If the speech perception system filters out all the variability not conforming to the stored abstract representations, how can this occur?

Furthermore, in the last years a series of studies have shown that the way speech information is stored in the brain is highly dynamic. In a seminal paper, Norris, McQueen & Cutler (2003) showed that listeners can calibrate their phoneme representations to the particular characteristics of the speech they hear in a surprisingly short interval. In that study, participants were exposed either to some properly pronounced words (containing either /s/

[1]Larsson, Vera-Constán, Sebastián-Gallés & Deco (2008)
[2]Sebastián-Gallés, Vera-Constán, Larsson, Costa & Deco (2009)

or /ʃ/) or to words with ambiguous sounds (half way between /s/ or /ʃ/). The results showed that listeners shifted their phonetic boundaries to accommodate these 'variations'. The results of that study have been replicated several times in a number of studies. Interestingly, some of the studies indicate that these changes are not transient in nature, but relatively long-lasting. After 25 minute post-exposure periods, containing either silence or different types of intervening auditory stimuli, Kraljic & Samuel (2005) observed numerically larger shifts than those measured immediately after the main stimuli were presented. Eisner & McQueen (2006) have reported convergent results with 12 hour intervals (even after sleep). All these results indicate that exposure to variability may induce changes in the speech perception system that remain stable over time and despite new exposures. The changes studied in these investigations have mostly consisted of the modification of speech categories (phonemes) induced by exposure to non-standard utterances. However, an open question is to which extent the lexicon is modified and whether variation is also represented at this level.

As suggested, one way of studying this issue is analyzing word recognition in listeners exposed to dialectal variation. Most dialects are characterised by changes at the phonological (and, specifically, phonetic) level. The literature about adaptation of the speech perception system to dialectal variability has not been focused on how dialectal variability is encoded in the lexicon, but mostly on more broad aspects of perception of accented speech (see, for instance, Clopper & Pisoni 2005). Two studies are worth mentioning here, both addressing the core issue of how different pronunciations of the same words are stored in the mental lexicon. In the first one, Connine (2004) investigated the recognition of a frequently heard spoken word form variant in American English (flapping). She presented words and non-words including either the standard /t/ or the flap sound (e.g., /prɪdɪ/ - /brɪdɪ/ or /prɪtɪ/ - /brɪtɪ/). Participants were asked to identify the initial segment (/b/ or /p/). The results showed better performance in identifying the target sound when it occurred in the more frequently heard flap carrier. The author considered that the results showed that listeners stored in their mental lexicon the flap variant of the words together with the standard form. The second study where evidence about the representation of dialectal variation of spoken words was obtained was Sebastián-Gallés, Echeverría & Bosch (2005) (later replicated in Sebastián-Gallés, Rodríguez-Fornells, Diego-Balaguer & Diaz 2006). We will describe that work in some detail, because it constitutes the basis of the present work. Before that, however, a few words concerning the linguistic properties of Catalan and Spanish and of the sociolinguistic situation of the participants of that study are in

order.

Catalan and Spanish are two Romance languages, co-official in Catalonia (north-east of Spain, Barcelona being the largest city). While both languages are similar at different levels of linguistic description, critically they differ at the level of the phoneme repertoire. For instance, while Spanish has no voiced fricatives, Catalan does. Central to the present study, Catalan has eight vowels, and Spanish only five. In particular, Catalan has two mid-front vowels (/e/ and /ɛ/) while Spanish has only one, falling roughly midway between the two Catalan ones. This phoneme distribution is very similar to the one of the English liquid contrast /r/-/l/ and the Japanese liquid /l/. Like Japanese natives learning English with the /r/-/l/ contrast, Spanish natives have great difficulties in perceiving (and producing) the Catalan-specific /e/-/ɛ/ contrast.

Since Spanish and Catalan are co-official in Catalonia, they are both widely used, in formal as well as informal contexts. According to official statistics (www.idescat.net), about 95% of the population living in Catalonia understand Catalan and about 75% declare they can speak it. In the Barcelona area, more than 40% of the people declare that Catalan is their 'usual' language of use ('llengua habitual'). Illustratively, at the University of Barcelona more than 60% of the courses are taught in Catalan and students are free to use any of the languages to write their essays and exams. In fact, students entering Catalan universities must show a written and oral command of both Spanish and Catalan equivalent to that which students elsewhere in Spain must show for Spanish. So, both Catalan and Spanish are frequently used. Still, as a series of studies have shown (Bosch, Costa & Sebastián-Gallés 2000, Pallier, Bosch & Sebastián-Gallés 1997, Pallier, Colomé & Sebastián-Gallés 2001, Sebastián-Gallés & Soto-Faraco 1999, Sebastián-Gallés et al. 2006), most individuals (even those raised in Catalonia) born within monolingual Spanish families, fail to perceive (and produce) many Catalan-specific sounds; in particular, the Catalan /e/-/ɛ/ contrast. Instead, they use the Spanish five-vowel system when speaking in Catalan, thus producing Catalan with a Spanish accent. Because of the bilingual nature of Catalan society, it is very common to hear both native and accented Catalan. That way, listeners are frequently exposed to different varieties of the same words. For instance, the Catalan word GALLEDA (pronounced /gallɛda/[3]) is pronounced with the Spanish phonemes /galleda/. In the same way as American natives consider the liquid sound produced by Japanese speakers as an /l/, Catalan natives

[3]To simplify the reading, only the phonetic symbols of the critical phonemes (/e/,/ɛ/) will be used when transcribing the pronunciation of words.

assimilate the Spanish mid-front /e/ vowel to the Catalan /e/. Thus, for words like 'galleda' they hear two different acoustic varieties, the one produced by Catalan natives, including vowel /ɛ/ and another produced by Spanish natives, including vowel /e/. The exposure to different forms does not occur for Catalan words containing vowel /e/, like FINESTRA (pronounced /finestra/), since both Catalan and Spanish natives produce words that to Catalan natives sound equivalent.

In the most relevant study preceding this work (Sebastián-Gallés et al. 2005), Catalan participants were asked to do an auditory lexical decision task. Pseudowords were created by changing one vowel of real words. There were three types of pseudowords: Control pseudowords involved a vowel change existing in Spanish (for instance, the Catalan word CADIRA was transformed into CADURA, /i/-/u/ being a vowel contrast common to both Spanish and Catalan). ɛ-type pseudowords consisted of pseudowords created from Catalan words including vowel /ɛ/, which was replaced by vowel /e/ (like the GALLEDA example just described). e-type pseudowords were created from Catalan words including vowel /e/, being replaced with vowel /ɛ/ (like the FINESTRA example just described). As can be seen, ɛ-type pseudowords corresponded in fact to the way Spanish natives mispronounce some Catalan words. The results of the experiment showed that Catalan natives had great difficulties in rejecting the ɛ-type pseudowords as such (the percentage of correct responses in this category were less than 60%, while for the other two pseudoword categories they were close to 90%). Therefore, an asymmetry was observed in the percentage of correct rejections of both types of pseudowords. The authors interpreted these results by postulating that the repeated exposure to the dialect that Spanish natives use when speaking in Catalan created multiple representations for ɛ-type words, e.g., in the GALLEDA case, they had stored in their lexicon both forms, /gallɛda/ and /galleda/. Thus, when confronted with the 'pseudoword' /galleda/ they often failed to reject it as a pseudoword. However, this would not happen for e-type stimuli. As we have stated, Spanish natives never mispronounce these items, for which reason Catalan natives would never have stored /finɛstra/ in their lexicon. These results were considered to indicate that participants held in their mental lexicon both the correct (native) pronunciation of words and the incorrect (non-native accented) ones. The present work aims at elucidating possible mechanisms for and consequences of the asymmetry in the performance of Catalan natives, observed by Sebastián-Gallés et al. (2005), exploring the implications of such mechanisms for the architecture of models of spoken word recognition. To this end, a modelling approach is proposed, using a biologically plausible mathematical formulation of

Stimulus Type	*English Translation*	*Word*	*Pseudoword*	*Native Pronunciation*	*Dialectal Pronunciation*	*% Correct Pseudoword Responses*
/ɛ/-type	Bucket	Galleda	Galleda	/ɛ/	/e/	58
/e/-type	Window	Finestra	Finestra	/e/	/e/	89
Control	Chair	Cadira	Cadura	/i/	/i/	90

Table 4.1: Different types of stimuli, native and dialectal pronunciations, and percentage of correct responses to pseudoword stimuli used in Sebastián-Gallés et al. (2005).

the functioning of single neurons as a foundation and then constructing a network of these neurons, organized in functional pools, whose interactions can account for the behaviour observed. Specifically, two competing network architectures will be proposed, and consequently their behaviour will be systematically explored using the analytical tool of mean-field analysis as well as explicit simulations of the behaviour of the network of individual neurons. The pools of the model are interconnected and each represents one or more phonemes. Upon the presentation of a stimulus, the persistent co-activation of pools, which together represent a full word, is taken to be word recognition. The weights of the connections between pools are the parameters of the model, whose specific values will determine the behaviour of the network. A careful analysis of the network is performed to establish the appropriate parameter values. One of the proposed models supposes an interaction at the phonemic level as the main mechanism behind the observed results, while the other model explores whether the context of a complete word is necessary for the asymmetry to be accounted for, in which case it would rather be a lexical effect. In both cases, a particular parameter configuration of the networks will be considered to be the result of plasticity in the brain, having occurred on the long term due to exposure to ambient language. Such plasticity is commonly considered to be the result of Hebbian learning, in which the connections between co-activated neurons, such as those receiving the same stimulus, strengthen. Similarly, lack of common input can lead to the weakening of connections between neurons, so called anti-Hebbian learning. In a final modelling effort, we explore the behaviour of the models in response to a single-phoneme input, simulating a discrimination task, in order to further clarify which of the proposed models is more appropriate.

It is stressed here that this approach makes explicit use of the neurodynamics underlying any operation carried out by the human cortex, and

the stochastic nature of the fluctuations inherent in the cortex as well as in the model makes a decisive contribution to the solution of the problem. Specifically, with a proper description of the non-stationary transients, using a spiking neuron model and realistic synaptic dynamics, one can account for response times in psychophysical experiments, as well as make use of fluctuations to get a probabilistic response behaviour (basis for decision making). As models of this kind have already been successfully employed modelling working-memory (e.g., Brunel & Wang 2001) and its role in the discrimination between two sequential sensory stimuli (Machens, Romo & Brody 2005, Miller & Wang 2002, Deco & Rolls 2006) as well as in other decision making tasks (Wang 2002, Wong & Wang 2006), we believe they can make an important contribution to elucidating the workings of language processing as well.

Although there exist models of phoneme perception, notably McClelland & Elman (1986) and Norris, McQueen & Cutler (2000), to our knowledge an approach with such a solid basis in neuroscience as ours (time course and probabilistic behaviour) has not before been applied to this kind of problem. While it is acknowledged that other approaches include input with a closer correspondence to real sequential sound streams, this is not of crucial importance, as we see this work as a proof-of-principle of the applicability of such more neurodynamical modelling schemes to behavioural data from psychology.

Additionally, with the aim of exploring the prediction of perfect phoneme discrimination based on the model results, we design and conduct a psychophysical experiment to investigate whether phoneme categories are modified in the Catalan-dominant early bilinguals, or whether the effect is purely lexical, as predicted if indeed phoneme discrimination is perfect in these individuals.

4.2. Methods

We base our model on the single-neuron level of description, using *Integrate-and-Fire* neurons (see next section and appendix C). This allows for the use of realistic biophysical constants obtained in neurophysiological studies, such as synaptic conductances and delays, which enables us to study the time scales and neural firing rates involved in the evolution of the neural activity underlying cognitive processes. The main benefits of using such a detailed model are, as emphasized in the introduction, the possibility that opens up for studying the time course of the neural activity, including transient responses, as well as the inherent stochasticity in the model,

which helps explain variability in the experimental data. Specifically, the former can be used to obtain reaction times from the simulation data, whereas the latter can be used to obtain bistable response patterns, e.g., corresponding to a certain distribution of responses in data from a binary choice experimental task. Also, obtaining such elaborate model data allows us to compare at a detailed level with data from neurophysiological experiments, should it be available.

Using the above-mentioned model of single neurons as a building block, we construct networks of neurons, organized in interconnected functional pools. The neurons of such a pool share common input and are thus co-activated. This is assumed to have set the internal connections of that pool higher than network average through a Hebbian learning process (Hebb 1949). In the framework of *biased competition and cooperation* (Rolls & Deco 2002, Deco & Rolls 2004, Deco & Rolls 2005*a*, Deco & Rolls 2005*b*), such a pool typically represents a certain aspect of the external input, such as visual or auditory object identities. External input to a network of such pools can bias the internal activity of the network through intra-pool cooperation and inter-pool competition. With a sufficiently strong synaptic connection binding two or more pools, one can also get cooperation between pools. These connection strengths or weights describe relative deviations of the synaptic conductances from their average value across the network. Stronger intra-pool weights implement reverberation of a pool's neuronal activity, which can underlie the formation of working memory, i.e., appearance of sustained high activity coding for a previously presented stimulus, which is absent at present. This concept of an *attractor* in the system's dynamics is used to encode presented stimuli, which are then seen as having been recognized by the system.

The modelling scheme using integrate-and-fire neurons grouped in such functional pools, has the advantage that it can easily be analyzed using the *mean-field* approach, which is a theoretical tool that uses the pool organization of the network to calculate approximate solutions of the stationary state of the dynamics. This will be elaborated on in the Analysis section.

4.2.1. The Integrate-and-Fire Model

The model we use is based on non-linear leaky integrate-and-fire neurons, which are coupled together to form a network of neurons. This neural network basis is adapted from Brunel & Wang (2001) and we refer to that paper for a more detailed description of the neurodynamical properties. In appendix C, the mathematical model for a single neuron is defined.

Essentially, it consists of an equation governing the evolution of the neuron's sub-threshold membrane potential. Via synapses, excitatory and inhibitory input currents effect an increase or decrease of the membrane potential, respectively. In the absence of such input, the membrane potential decays exponentially over time. When the membrane potential reaches the threshold potential, an action potential (a spike) is emitted. Spikes in the membrane potential are in the model treated as unitary events of equal magnitude and duration. An emitted spike propagates to connected neurons via explicitly modelled synaptic mechanisms, namely AMPA, NMDA and GABA_A receptors on the postsynaptic neurons.

More specifically, the membrane potential of a single neuron is modelled by an electrical circuit, which consists of a capacitor C_m connected in parallel with a resistor R_m. This circuit describes how the membrane potential $V(t)$ evolves in time depending on external currents entering from other neurons. When the threshold membrane potential V_{thr} is reached, an action potential is emitted and propagates on to other neurons, while the membrane potential of the neuron that spiked is set to the reset potential V_{reset}, at which it is kept for a refractory period τ_{ref}. Both excitatory and inhibitory neurons have a resting potential $V_L = -70$ mV, a firing threshold $V_{thr} = -50$ mV and a reset potential $V_{reset} = -55$ mV. The membrane parameters are different for excitatory and inhibitory neurons. The former type of neuron is modelled with membrane capacitance $C_m = 0.5nF$, leak conductance $g_m = 25$ nS, membrane time constant $\tau_m = 20$ ms, and refractory period $t_{ref} = 2$ ms, whereas the inhibitory neurons have the corresponding values $C_m = 0.2nF$, $g_m = 20$ nS, $\tau_m = 10$ ms, and $t_{ref} = 1$ ms. Values are taken from McCormick, Connors, Lighthall & Prince (1985).

The incoming synaptic influences are both excitatory and inhibitory. The excitatory neurons transmit their action potentials via the glutamatergic receptors AMPA and NMDA, which are both modelled with exponential terms. We neglect the rise time of the current mediated by the AMPA channel, because it is typically very short ($\ll 1$ ms), and just model the decay period with $\tau_{AMPA} = 2$ ms (Hestrin, Sah & Nicoll 1990, Spruston, Jonas & Sakmann 1995). The NMDA channel is modelled with a rise term, $\tau_{NMDA,rise} = 2$ ms, a decay term, $\tau_{NMDA,decay} = 100$ ms (Hestrin et al. 1990, Spruston et al. 1995), and an extra voltage dependence controlled by the extracellular magnesium concentration, $C_{Mg^{2+}} = 1mM$ (Jahr & Stevens 1990). The inhibitory postsynaptic potential is mediated by the GABA_A receptor and is described by a decay term with time constant $\tau_{GABA} = 10$ ms (Salin & Prince 1996, Xiang, Huguenard & Prince 1998).

4.2.2. Network setup and architecture

Our network contains a total of 2000 neurons, all connected to each other. Of these, $N_E = 1600$ are excitatory and $N_I = 400$ are inhibitory . This proportion of inhibitory to excitatory neurons is consistent with the observed proportion of interneurons to pyramidal neurons in the cerebral cortex (Braitenberg & Schütz 1991, Abeles 1991). These neurons are subdivided into six pools of three types. Four of the pools are selective pools, each consisting of 150 excitatory neurons. These pools are activated in association with the input, in this case one or many phonemes. In addition, there is a non-selective pool consisting of 1000 excitatory neurons, which are not activated in association with the input. Finally, there is a pool of 400 inhibitory neurons, which provide global inhibition, thereby helping control network activity. Figure 4.1 schematically illustrates the basic network structure. However, before creating this pool structure by differentially setting connection strengths through the modification of synaptic *weights*, all connections' synaptic *conductances* are initially given values, which ensure realistic spontaneous activity of 3Hz and 9Hz for the excitatory and inhibitory neurons, respectively (for this procedure, see Brunel & Wang 2001), values which approximately correspond to those obtained in neurophysiological studies (Koch & Fuster 1989, Wilson, Scalaidhe & Goldman-Rakic 1994).

Each selective population is then characterized by a higher internal connection strength, designated by the weight w_+. This is in accordance with the assumption that shared input leads to correlated activity, which in turn strengthens the connections between these neurons according to the concept of Hebbian learning (Hebb 1949). Here, it is assumed that weights have achieved the constant values they have in a certain simulation through such a Hebbian learning process, possibly over a very large time span as in long-term plasticity.

The connection strength between different selective populations, w_-, has a value of less than one and mediates synaptic depression, which compensates for the effect of the high value of w_+ on the network level. In specific, the value of these connections is given by $w_- = 1-f(w_+-1)/(1-f)$, where f is the fraction of excitatory neurons present in each selective pool (in relation to *all* excitatory neurons, i.e., $f = 150/1600$), which ensures that the overall recurrent excitatory drive in the spontaneous state remains constant even when w_+ varies (Brunel & Wang 2001). This recurrent excitation, mediated by the AMPA and NMDA receptors, is assumed to be dominated by the NMDA current to provide a more robust behaviour during the period after stimulus offset.

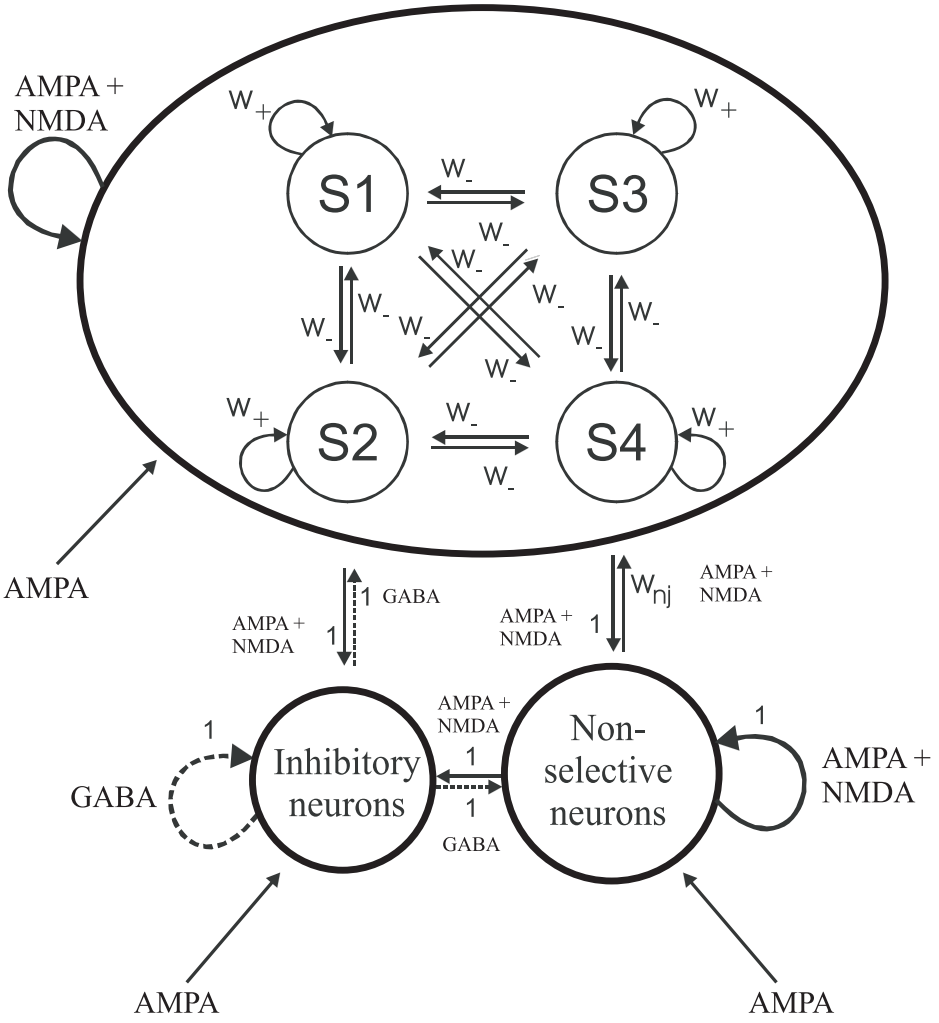


Figure 4.1: Network architecture. The network has five pools of excitatory neurons, four selective (S1-S4) pools and one nonselective pool, as well as one pool of inhibitory neurons. Dashed lines indicate inhibitory synaptic connections, solid lines indicate excitatory synaptic connections. For each connection, the type of neurotransmitter(s) employed is (are) indicated. See main text for explanation of synaptic weights w_+, w_-, and w_{nj}.

Connections from the inhibitory pool (GABA-ergic connections) to all other pools, including itself, are set to 1. The connections from the non-selective pool to itself are set to 1, since its neurons do not share selectivity. The connections from the non-selective pool to the selective pool j is given by $w_{nj} = \frac{(1-\sum_{\forall i \setminus n} f w_{ij})}{f_n}$, where f is as before, f_n is the fraction of excitatory neurons in the non-selective pool, w_{ij} is the weight value from pool i to pool j and the sum goes over all selective pools. This setting of the weights ensures that the average input to each selective pool is 1.

This brings us to the definition of the two specific architectures we have considered (figures 4.2 and 4.3). In both models, there are two types of

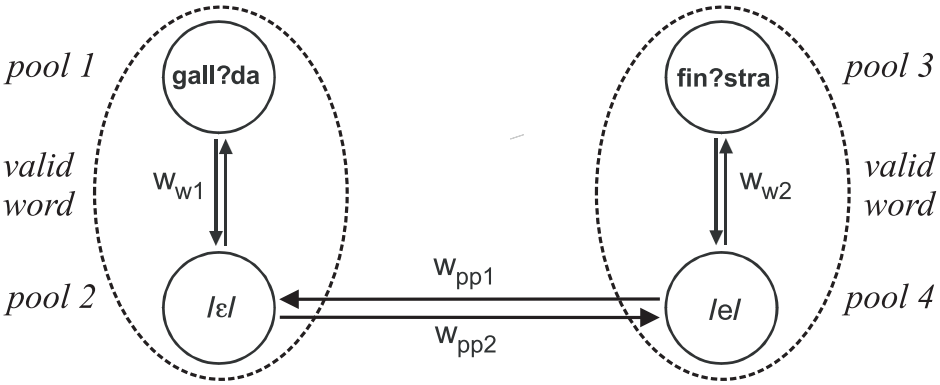


Figure 4.2: Model 1. The figure shows the parameters which, apart from w_+, are varied in the search for specific phenomena in the network behaviour. In the figure, connections with the value w_- are omitted. This model was designed to look for an explanation of the experimental data at the phonemic level, represented by the interconnected Pools 2 and 4. The four pools illustrated here correspond to the selective pools in Figure 4.1.

selective pools, phoneme pools and word context pools, which together can represent a word. The phoneme pools simply code for a specific phoneme, while the word context pools can be imagined as a conglomeration of single phoneme pools which together create the context. Alternatively, one could model each phoneme of the ones constituting the whole word by its own selective pool, but this is not computationally feasible, for which reason we have chosen the intermediate representation of a context pool, while explicitly modelling the crucial critical contrast phonemes of the stimuli.

In model 1 (figure 4.2), parameters w_{pp1} and w_{pp2} interconnect the two phoneme (p) pools, which opens for the possibility of a change having

occurred in the representation of phoneme categories. The parameters w_{w1} and w_{w2} each help define a complete word (w), by binding a context pool with a phoneme pool to enable the representation and recognition of a *valid* (see figure) word, as opposed to a non-word. Each of these two parameters controls this interaction bidirectionally, the same value being given to connections in both directions. Figure 4.3 is a schematic drawing

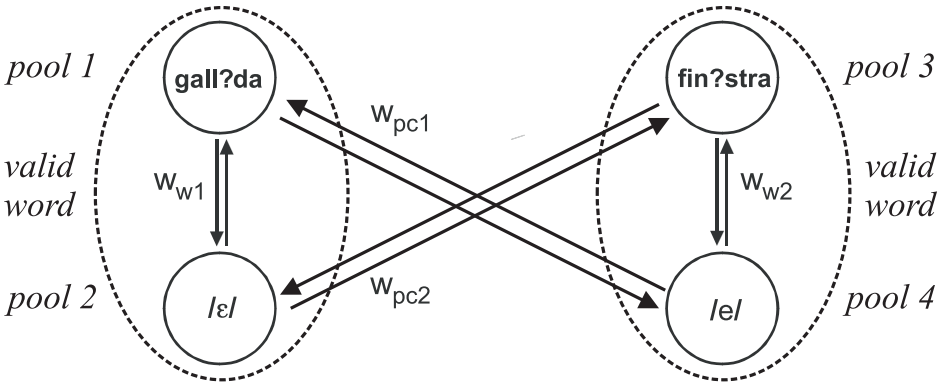


Figure 4.3: Model 2. The figure shows the parameters which, apart from w_+, are allowed to vary in the search for specific phenomena in the network behaviour. In the figure, connections with the value w_- are omitted. This model was designed to look for an explanation of the experimental data at the lexical level, interconnecting a context pool with a different phoneme pool than its normalcritical phoneme constituent. The four pools illustrated here correspond to the selective pools in Figure 4.1.

of model 2. Here, w_{w1} and w_{w2} have the same function as in model 1. However, instead of parameters connecting the phoneme pools, we have the bidirectional w_{pc1} and w_{pc2}, which each interconnect a phoneme (p) pool with a context (c) pool. These connections open for the possibility that a new lexical representation may have evolved over time, thus enabling the recognition of a non-word, such as /galleda/ (enabled by w_{pc1}), as a word.

It is stressed here that, although the pools in figures 4.2 and 4.3 are marked by the critical phonemes and their corresponding word contexts, which together constitute the two stimuli which were used as representatives in the discussion of the behavioural experiment we are modelling, they might equally well have been assigned other names, because the results obtained should be equally valid for other stimuli used in the underlying experimental study. Henceforth, the selective pools will be referred to by their respective numbers (see figures 4.2 and 4.3). The networks have no

spatial structure or extension.

All neurons in the network are exposed to an external Poisson spike train input through 800 excitatory connections, mediated by AMPA receptors at a rate of $\nu_{ext} = 3$ Hz per synapse, corresponding to a typical value for spontaneous activity in the cerebral cortex (Wilson et al. 1994, Rolls & Treves 1998). This results in a total external input of 2.4 kHz to each neuron in the network, representing noise input from surrounding areas in the cortex not explicitly modelled.

The network is presented with a stimulus by increasing the external input of the pools corresponding to the stimulus with $\lambda_{stim} = 0.1$ Hz, yielding $\nu_{stim} = \nu_{ext} + \lambda_{stim} = 3.1$ Hz. For example, to present the input /galleda/, this higher external input is applied to pools 1 and 4, adding a total of 80 Hz to the total external input to each neuron in those pools. Input is always maintained for 500 ms, regardless of the stimulus presented, a value which approximately corresponds to the durations of the four three-syllable stimuli used as representatives in our discussion, the Catalan words *galleda* and *finestra* and their corresponding non-words.

The stimulus input influences the inherent competition and cooperation between areas in the network and has the capacity to drive the activity of the pools given the input into an attractor. Specifically, the setting of the inter-pool parameters of each model higher than the surrounding inter-pool connections, combined with the higher value of w_+ (creating so-called pool cohesion), can give rise to a natural tendency to competition between distinct groups of pools, which internally display cooperation between the pools constituting such a group. In our case, an example of this would be the tendency to co-activate in the group consisting of pools 3 and 4, representing /finestra/, because of setting w_{w2} higher than w_-.

4.2.3. Simulations and Analysis

We use spiking simulations to analyze the time course of network activity and the influence of fluctuations. However, they are time consuming and can therefore not be run for all parameter configurations of the network, especially not with five free parameters as in our case. It is here that the mean-field approximation is a highly useful tool for assessing network behaviour. It assures that the dynamics of the network will converge to a stationary attractor which is consistent with the asymptotic behaviour of an asynchronous firing network of integrate-and-fire neurons (Brunel & Wang 2001). In the standard mean-field approach, the network is partitioned into populations of neurons which share the same statistical properties of the afferent currents, and fire spikes independently at the

same rate. The essence of the mean-field approximation is to simplify the integrate-and-fire equations by replacing, in accordance with the diffusion approximation (Tuckwell 1988), the sum of the synaptic currents by the average direct current and a fluctuation term. The stationary dynamics of each population can be described by the population transfer function, $F()$, which provides the average population rate as a function of the average input current. For further details, see appendix D.

We used mean-field simulations to calculate the attractor states of the network in a large number of points in parameter space, thereby assessing which points displayed the behaviour of interest. Once this was determined, we ran spiking simulations, because although the mean-field attractor states reflect the true behaviour of the network, the fluctuations and external changes added in the spiking simulations can yield a significantly different behaviour. Apart from this, we need the temporal dimension to compare simulation results with experimental data, so spiking simulations are of the essence.

At a first stage, we ran extensive mean-field scans for both models over their respective five network parameters, i.e., the internal pool cohesion parameter, w_+, and the four inter-pool weights (lateral, word-cohesion weights w_{w1} and w_{w2} for both models, plus w_{pp1} and w_{pp2} for model 1 or w_{pc1} and w_{pc2} for model 2). In the mean-field scans we looked for parameter space points in which the behaviour of the network be the following. Upon the presentation of *words*, e.g., /gallɛda/ and /finestra/, the average firing rates of their corresponding pools should be above threshold, while the other two pools of the network should display below-threshold activity in the mean; upon presentation of unlikely *non-words*, e.g., /finɛstra/, all pools should show below-threshold firing-rates; however, upon presentation of *non-words* more likely to occur, e.g., /galleda/, the pools receiving input (pools 1 and 4) should show an above-threshold mean firing rate, indicating that in this point we might find the bistable behaviour we are looking for in the network's response to this kind of stimulus. That is, we expected this mean behaviour to fluctuate when adding the noise of the full spiking simulations, yielding different results in different trials, which would reflect the differing behaviour of subjects on the task using this kind of non-word stimulus.

In specific, typical values for the mean-field scans of both models were: $w_+ \in [1.9, 2.5]$, while w_{w1}, w_{w2}, w_{pc1}, w_{pc2}, w_{pp1} and w_{pp2} typically varied between 1.0 and 1.2. Step sizes varied depending on the interval being searched for each parameter, with the smallest used being 0.01 and the largest 0.1. Very large scans can take a long time, although they save immense amounts of time compared to spiking simulations, and they can

also generate huge files which have to be stored and analyzed. Since the step size and the interval determine file size and time consumption, we used early results from course-grained scans to identify regions in parameter space of particular interest, for which we could then run more fine-grained mean-field scans, and eventually spiking trials, when the region got small enough. One important constraint in finding interesting points were the inter-relations between different weights. Specifically, the value of w_+ should always be greater than all other weights in the network, for individual pool reverberatory activity to be maintained. This constraint only aided in the initial choice of parameter intervals which were scanned for the other parameters. Two further constraints were however used to discard points which otherwise displayed wanted behaviour in the mean-field simulations. One of them, $w_{w1} < w_{w2}$ reflected the fact that the frequency of occurrence of /gallɛda/ is lower in the environment than ditto for /finestra/, which through Hebbian learning should have led to the connection binding the constituent pools for the latter word being stronger than the one for the former word. The second additional constraint was $w_{pc1} > w_{pc2}$, which was added to increase the likelihood of bi-stability in spiking trials following the presentation of /galleda/, while decreasing the likelihood of /finɛstra/ being recognized as a word. By varying these weights, we predicted one could modify overall network behaviour in these two crucial, non-word, cases.

We started the spiking simulations with a pre-stimulus period of 300 ms, corresponding to the time during which the asterisk is shown to participants in the LDT experiment in Sebastián-Gallés et al. (2005). We present the stimulus for 500 ms to the corresponding context and phoneme pools, together constituting the stimulus word or non-word, by increasing the external input as described before. Afterwards, a period of another 1000 ms is simulated, a value corresponding approximately to a typical reaction time in the LDT task. This yields a total of 1800 ms for each trial simulated. In this context, a point in parameter space corresponds to an individual participant and a trial corresponds to one trial for that participant. We typically ran 100 trials per point. The analysis of the data thus obtained (an 1800 ms time course of the firing rate of each pool, averaged over its neurons and sampled every 20 ms) was analyzed in the following manner: first, an average of the last 500 ms simulated was calculated for each trial (leaving time for network activity to stabilize somewhat after stimulus offset). Then, it was checked if this average firing rate exceeded a threshold set to 10 Hz, and if so, this pool was taken to be in an attractor state, meaning it had retained a representation of the stimulus it codes for. Thus, we obtained either attractor states or spontaneous states for all the 100

trials, for each pool in the network. This data is obtained for each stimulus given to the network, for a total of four different such data sets to analyze (one for each word and one for each non-word), for *each* point/individual we chose to simulate.

For each point, once we had the outcome of the 100 trials with respect to the threshold, for each of the four stimuli, we checked if the outcome corresponded to a 'yes' or a 'no' response, i.e., if the stimulus was taken to be a word or not. The definitions of these responses in terms of network outcome pattern differed, according to which stimulus had been presented, to be further illustrated in the next section. For each stimulus, thus, using the 100 different trials we obtained a percentage of *correct* trials using a binary test based on the particular definition of a 'yes' (or 'no') response. Then we used this percentage in the comparison of model data with experimental data, either directly or computing the A' statistic for comparison with the corresponding values in the experimental paper. The A' values were calculated using the following formula (Macmillan & Creelman 1991):

$$A' = 1/2 + \frac{(Hit - FA) \cdot (1 + Hit - FA)}{4 \cdot Hit \cdot (1 - FA)}, Hit \geq FA \tag{4.1}$$

where FA is the fraction of false alarms, i.e., the response 'yes' to the presentation of a non-word, and Hit is the fraction of correct responses, i.e, the response 'yes' when the stimulus presented was indeed a word. For the case when $Hit < FA$, Hit and FA change places in equation (4.1).

4.2.4. Experimental Methods

In the Sebastián-Gallés et al. (2005) study, participants were only tested with a lexical decision task. In the present work, inspired by our modelling results, the very same participants are tested in a discrimination task and in a lexical decision task. The methods of the lexical decision task were exactly the same as those employed in the study just mentioned.

4.2.4.1. Lexical Decision Task

For a full description of this experiment, see Sebastián-Gallés et al. (2005). Thirty two native Catalan participants (8 male) took part in this experiment (age average: 22.7 years, s.d. 9.4). They were students from the University of Barcelona and they were born in Barcelona or its metropolitan area. They participated in exchange for course credits. No auditory or language learning problems were reported. Materials were

exactly the same as in the Sebastián-Gallés et al. (2005) experiment. One pair of stimuli was not included in the final analysis because it elicited a very high error rate. As in the previous study, experimental stimuli were words and non-words containing either vowel /e/ or vowel /ɛ/. To create non-words, both vowels were exchanged (see table 4.1). Participants were seated in a sound-proof booth in front of a personal computer screen where instructions were displayed. Instructions emphasized that pseudo-words could be very similar to real words, and that the participants should pay attention to vowels because pseudo-words had been created by replacing a single vowel. It was stressed that in many cases the replacement involved the exchange of vowels /ɛ/ and /e/. Responses were made by pressing one of two labeled buttons ('yes' for words, 'no' for non-words) with their dominant hand. Participants were encouraged to keep their response fingers over the response buttons to respond as quickly as possible. Stimuli were presented in two lists. Each member of each word-pseudoword pair appeared in only one list. Half of the participants were tested with one list, and the other half with the other list. Presentation order was fully randomized for each subject. Reaction times were measured from stimulus onset. Total duration of this experiment was approximately 15 minutes.

4.2.4.2. Discrimination Task

Three different female Catalan native speakers recorded in a single session several tokens of syllables /de/ and /dɛ/ in a soundproof room. Individual files were created for each token using the Cool Edit © software, with the sampling rate set to 16000 Hz. To select stimuli, each speaker discarded from their own utterances the ones they considered poor exemplars. Then, these three speakers plus a fourth native Catalan speaker selected at least three tokens as the most representative and prototypical from each category. The final selection was done on those stimuli selected as the best for at least two speakers. The average length for e-type stimuli was 390 ms (s.d. 49.3) and 412 ms for ɛ-type stimuli (s.d. 38.5). There were no statistical differences between the two types of stimuli. Amplitude was controlled with Praat software (average e-type: 69.8 dB; ɛ-type: 69.7 dB).

Participants were tested in individual sound proof booths, seated in front of a personal computer screen where instructions were displayed. Stimuli were delivered binaurally through Sennheiser HD 435 headphones. Participants were asked to press a designated key whenever they detected a change in category, not in voice. Stimuli were presented in a pseudorandom order with at least three tokens of the same category and a maximum of eight tokens before a change was produced. As there were two categories,

two changes of direction were possible: from /dɛ/ to /de/ or vice versa (ɛ → e and e → ɛ, respectively). There was a training phase in which participants were asked to do the same task with the same procedure, but only two tokens - one from each category- from one speaker were played. In this phase, visual cues were used to indicate that there were two categories (numbers 1 and 2). Also, when a change took place another cue appeared indicating that they should press the key. These cues were not presented in the experimental phase, during which only a fixed asterisk was shown as a fixation point. The training phase consisted of 25 trials. The experimental phase consisted of 400 trials, divided into two blocks. Subjects could rest between these two blocks.

After extensive piloting, ISI was fixed to 800 ms. The experiment lasted for about 15 minutes. Stimulus presentation was controlled by the Presentation 0.60 software.

4.3. Results

The model results are presented in two stages. Firstly, we will focus on the main modelling study, the lexical decision task (LDT) described in the introduction. Secondly, we describe the discrimination task modelling and the consequences its results had as to the interpretation of results obtained in the LDT simulations, as well as regarding the implications for model structure and the problem studied.

Looking for the desired behaviour of the network using mean field scans (see methods), yielded many feasible points in model 2 in a fairly contiguous parameter space, which were candidates for full spiking simulations. However, in model 1 only a single parameter configuration was found which displayed this behaviour in the mean-field simulations, after *very* extensive scans, indicating the intractability of that model for explaining the experiments, already at an early stage of simulations.

The single feasible parameter configuration found in mean-field simulations for model 1, shown in table 4.2, was used in a spiking simulation of 100 trials. These spiking simulations did not show satisfactory results. Specifically, this network configuration showed a breakdown in identifying /finɛstra/ as a non-word, with an almost chance performance of 54% correct trials. In figure 4.4, a spiking trial representative of the 46% erroneous trials is shown, with activity high in pools 2 and 3, signaling word recognition. This is not the data we wish to reproduce, the real Catalan-dominant bilinguals having a far higher score on these stimuli. These results indicated that model 1 was not feasible and all our analytical efforts thereafter

	w_+	w_{w1}	w_{w2}	w_{pp1}	w_{pp2}
Point 1	2.10	1.01	1.02	1.11	1.08

Table 4.2: The only feasible parameter configuration, used in spiking trials run for model 1.

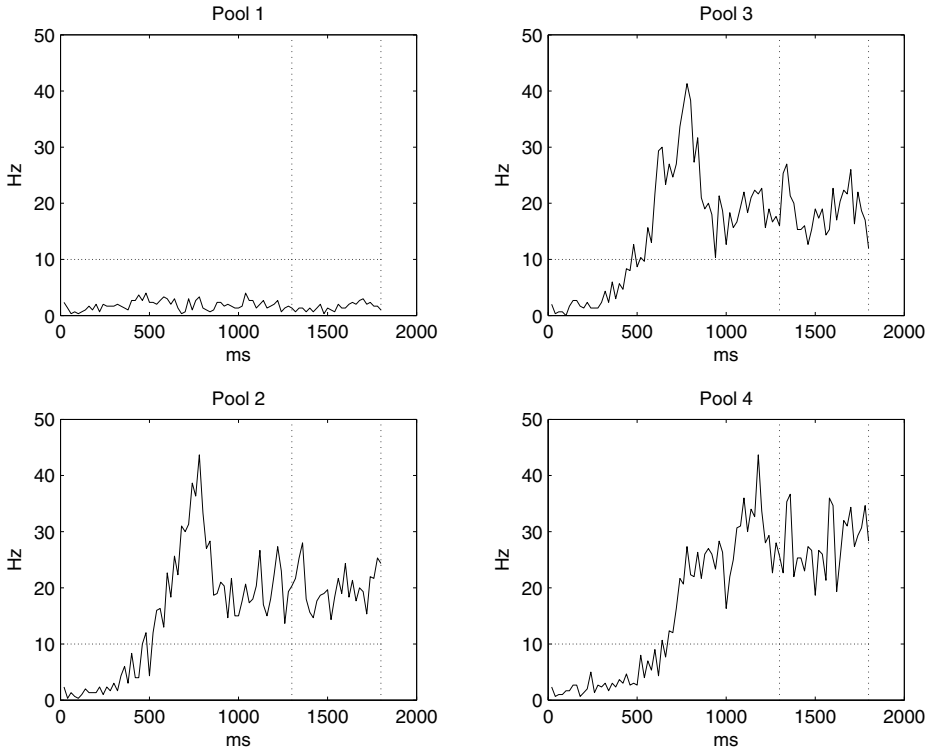


Figure 4.4: Model 1 spiking simulation for the input /finɛstra/. The figure shows the activity of the four pools, organized in relation to each other in the same way as in the model architecture. As can be seen in this trial, the pseudoword input drives the corresponding pools into an attractor and is thus recognized as a word, something which only very rarely would happen in reality. The horizontal line indicates the threshold; the two vertical lines enclose the time interval used for averaging the signal to obtain a response.

focused on model 2.

As concerns model 2, having decided on the points which should be interesting behaviourally, we consequently ran full spiking simulations with 100 trials per point, for 11 parameter configurations in total. These configurations can be seen in table 4.3 (a *point* means a point in the 5-D parameter space, i.e., a specific parameter configuration). The results

	w_+	w_{w1}	w_{w2}	w_{pc1}	w_{pc2}
Point 1	2.02	1.06	1.09	1.06	1.00
Point 2	2.02	1.07	1.09	1.06	1.00
Point 3	2.02	1.07	1.09	1.05	1.00
Point 4	2.02	1.08	1.09	1.06	1.00
Point 5	2.02	1.09	1.09	1.06	1.00
Point 6	2.03	1.06	1.09	1.06	1.00
Point 7	2.03	1.07	1.09	1.06	1.00
Point 8	2.03	1.08	1.09	1.06	1.00
Point 9	2.02	1.08	1.09	1.07	1.00
Point 10	2.02	1.09	1.09	1.07	1.00
Point 11	2.02	1.09	1.09	1.04	1.00

Table 4.3: Parameter configurations used in spiking trials run for model 2.

of the spiking simulations of model 2 looked very promising when going through the pool activity time courses for each stimulus and point. As an illustration, we here include a spiking simulation with the input /galleda/, presented to pools 1 and 4, showing the sought behaviour, i.e. taking the non-word stimulus to be a word (figure 4.5). However, as it is hard to get an overview here of all the data obtained in these simulations by presenting the exact time course, we had to find a good measure which, at the same time as having a clear-cut correspondence in the experimental data we model, was easily presentable in a written form. To this end we calculated the percentages 'yes' and 'no' responses over trials to each stimulus presentation (using the same response definitions as in the mean-field scans recounted before) and thereafter calculated the A' statistic for each network configuration simulation, using the different outcomes of the distinct trials to obtain the percentages. Using this approach, it was

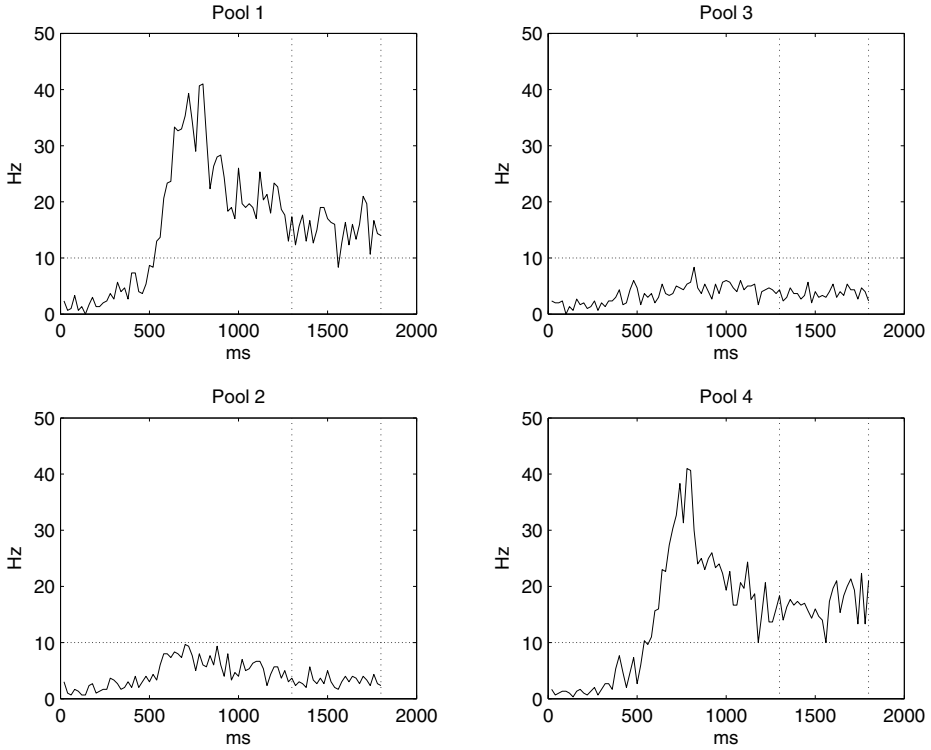


Figure 4.5: Model 2 spiking simulation for the input /galleda/. The figure shows the activity of the four pools, organized in relation to each other in the same way as in the model architecture. As can be seen in this trial, the pseudoword input drives the corresponding pools into an attractor and is thus recognized as a word, which also happens frequently in reality for individuals showing the asymmetry in performance. The horizontal line indicates the threshold; the two vertical lines enclose the time interval used for averaging the signal to obtain a response.

found that the experimental data was readily reproduced in the mean, with a variability of responses corresponding roughly to the distribution of behaviour in the experimental paper (Sebastián-Gallés et al. 2005). In table 4.4, you can see the percentages *correct* responses ('yes' response to a word presentation, 'no' response to a non-word presentation) to the presentation of the four stimuli we have considered in our modelling, together with the corresponding A' scores. The average values are in the bottom of the table, and correspond well to the average values obtained for Catalan-dominant early bilinguals in the experimental study (see table 4.1). To facilitate

	% Correct Responses					
	/ɛ/-Type		/e/-Type		A'	
Point	*W*	*PW*	*W*	*PW*	*/ɛ/-Type*	*/e/-Type*
1	57	70	92	94	0.715	0.962
2	76	75	91	97	0.838	0.969
3	81	87	91	94	0.905	0.960
4	89	77	92	93	0.900	0.960
5	94	77	92	90	0.919	0.951
6	78	54	99	87	0.751	0.964
7	89	55	98	88	0.824	0.964
8	94	60	96	90	0.869	0.963
9	86	62	89	93	0.833	0.960
10	93	64	89	90	0.876	0.941
11	100	85	100	81	0.963	0.953
Mean	85.18	69.64	93.55	90.64	0.854	0.959

Table 4.4: Individual and mean percent correct reponses and corresponding A' values for all stimuli, obtained in model 2 spiking simulations (W = word; PW = pseudoword).

comparison of average values, we have generated figures. Shown in figure 4.6 is a comparative graph of average percentages correct responses, from experiment and model. Furthermore, in figure 4.7 there is a comparative graph of the corresponding average A' values, from experiment and model.

Based on the results obtained in the modelling of the lexical decision task, we predicted that in order for the effect to manifest itself in our simulations, the connections between the two phoneme pools (pools 2 and 4) needed to be 'symmetric', i.e., their weights' values needed to be equal. In model 2, whose strong results indicate its suitability for the LDT task modelling, these weights are indeed equal, since they are both set to w_{-}. In the only point found in model 1, these weights (w_{pp1} and w_{pp2}) were 'asymmetric', i.e. differed in value. This led us to run further simulations, this time of a 'discrimination task', in which the network response must reflect a change

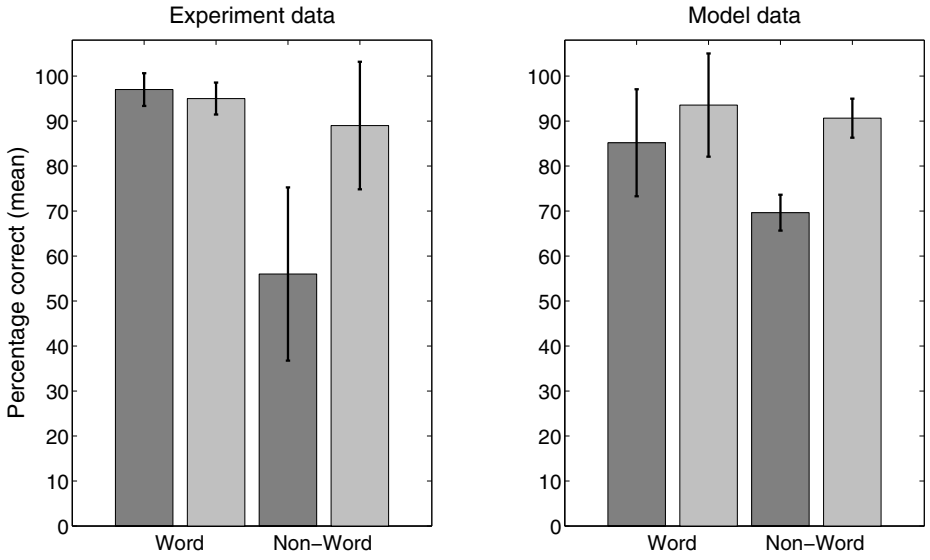


Figure 4.6: Shown are the average percentages of correct responses obtained in the experiment (left), for all four stimuli. The experimental values are for Catalan-dominant early bilinguals, taken from Sebastián-Gallés et al. (2005). Dark grey: /ɛ/-type stimuli. Light gray: /e/-type stimuli. Error bars represent standard deviations.

in phoneme stimuli. The motivation for this was our prediction that in a discrimination task using the critical phonemes as stimuli, the Catalan-dominant bilinguals should show equally perfect performance regardless of the direction of the change in stimuli in a sequential presentation. That is, we suspected they would not show the asymmetry displayed in the lexical decision task, i.e., in this case they would *not* perform worse on /ɛ/ stimuli than on /e/ stimuli. Linguistically this would be due to the lexical context not being present in such a task, in which case performance would not be affected by long-term changes in the lexicon. What we hoped to show with the modelling of this task, was that using the same points which showed the asymmetry in the LDT, we would obtain equal or nearly equal, very high A' values on the discrimination task, but *only* if there was not a significant asymmetry in weight values between the phoneme pools.

In this task, we simulated 100 trials for each of the 11 points already run for the LDT task, but with a different stimulus setup and duration. The initial period of zero input was maintained at 300 ms. Thereafter, the stimulus /ɛ/ or /e/ was presented to its corresponding pool, for a duration

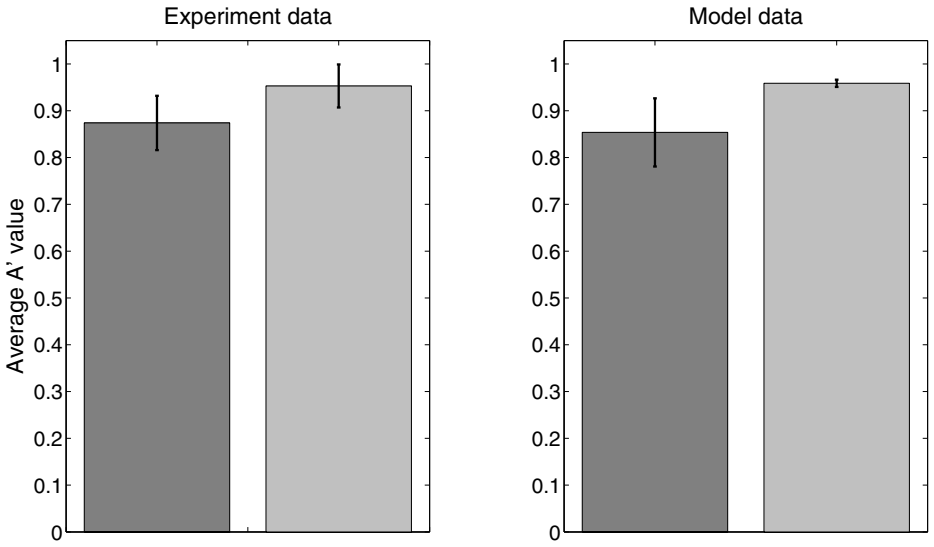


Figure 4.7: Shown are the average A' values obtained in the LDT experiment (left) and with model 2 (right). The experimental values are for Catalan-dominant early bilinguals, taken from Sebastián-Gallés et al. (2005). Dark grey: /ɛ/-type stimuli. Light gray: /e/-type stimuli. Error bars represent standard deviations.

of 400 ms in both stimulus cases. This was followed by an interstimulus interval of 400 ms, whereafter either the same phoneme was presented again, or the other phoneme was presented, representing a switch in stimuli. Finally, the same period of 1000 ms as in the LDT task was simulated after the last stimulus offset, for a total of 2500 ms simulation for the discrimination task.

In all four cases (two switches and two non-switches of stimuli), if activation showed that the last phoneme input was still retained by the network, the corresponding trial was deemed a correct response. That is, for the critical changes to be detected, the last stimulus presented to the network in a 'change' trial had to be the winner in the competition induced by the switch. Using this paradigm, we again calculated both percentages and A' values for all 11 parameter configurations, but not only for these exact points. We also did the same for points in which an asymmetry in the phoneme-to-phoneme weights, w_{pp1} and w_{pp2} (model 1 terminology), was introduced, one of the weights set to a greater-than-w_- value and the other retained at w_-. Specifically, the values were $w_{pp1} = w_- = 0.89$ and $w_{pp2} = 0.97$, or vice versa, in order to investigate the effect of an

asymmetry in both directions.

The results largely corroborated our prediction. In table 4.5, we show the mean A' values obtained from the spiking simulations for the non-asymmetric case and for the two asymmetric cases, calculated from the individual A' values of all 11 points (data not shown). As can be seen in

	Direction of Change	
Weight (A)symmetry	/ɛ/→/e/	/e/→/ɛ/
$w_{pp1} = w_{pp2}$	0.958	0.934
$w_{pp1} < w_{pp2}$	0.986	0.866
$w_{pp1} > w_{pp2}$	0.908	0.982

Table 4.5: Mean A' values obtained for the different spiking simulations of the discrimination task.

table 4.5, in the symmetric weight case, A' values differ by a mere 2.5 %, while in the asymmetric cases, they differ by 14 % and 8 %, respectively, in the order of presentation of the table. This indicates that there is a sensitive balance to be maintained, in which these inter-phoneme pool weights should be kept equal or may at least not differ by a significant amount, for discrimination performance not to show asymmetry.

The results of our simulations of the discrimination task strenghtened our prediction that plasticity at the lexical level does not alter phoneme categories' representations and that the performance on a phoneme categorization task using the critical phonemes here discussed, should approach perfect discrimination. Subsequently, as described in the Methods section, we conducted behavioural experiments to investigate the prediction, including a syllable discrimination task.

4.3.1. Experimental Results

4.3.1.1. Lexical Decision Task

As in Sebastián-Gallés et al. (2005), ɛ-type pseudowords yielded high error rates. The percentage of correct responses for each type of stimulus is shown in table 4.6.

	Word	*Pseudoword*	A'
/ɛ/-stim	95.1	56.9	0.865
/e/-stim	95.5	85.2	0.946
Control	98.2	95.1	0.982

Table 4.6: Percentage of correct responses and A' scores for each condition (Lexical Decision Task).

4.3.1.2. Discrimination Task

To obtain the A' statistics[4], the percentages of correct responses (*Hit*s) and false alarms (*FA*s) were calculated for each change of direction. Participants could give their response whenever during the entire trial duration. A trial in which the presented stimulus belonged to a different category than the previous one, was considered a 'change trial'. A *Hit* was scored when the response was given during change trials. Any response given at any other time was considered a false alarm. Participants showed very high performance, independently of the direction of change ($A'_{\varepsilon \to e} = 0.945$; $A'_{e \to \varepsilon} = 0.936$). In table 4.7, the percentages of *Hit*s and *FA*s are given.

	Correct Responses		*False Alarms*	
	/ɛ/→/e/	/e/→/ɛ/	/ɛ/→/e/	/e/→/ɛ/
Mean	86.4	81.8	5.3	4.4
SD	12.8	15.7	4.2	3.0

Table 4.7: Percentage of correct responses (Hits) and false alarms (FAs) for each direction of change (Discrimination Task).

[4]The use of A' statistics is due to the fact that this statistic was already employed in the Sebastián-Gallés et al. (2005) study. We have run parallel analyses with the d' statistic and the same pattern of results was obtained.

4.3.2. Comparison of Results

As the modelling predicted, the Catalan-dominant early bilinguals in the discrimination task indeed showed virtually equal performance on both directions of change of critical vowel. In figure 8 we show the average A' values attained by participants in the task, together with the average A' values obtained in the simulation of a phoneme discrimination task using the model. Like the results of the modelling (right part of figure 4.8), the

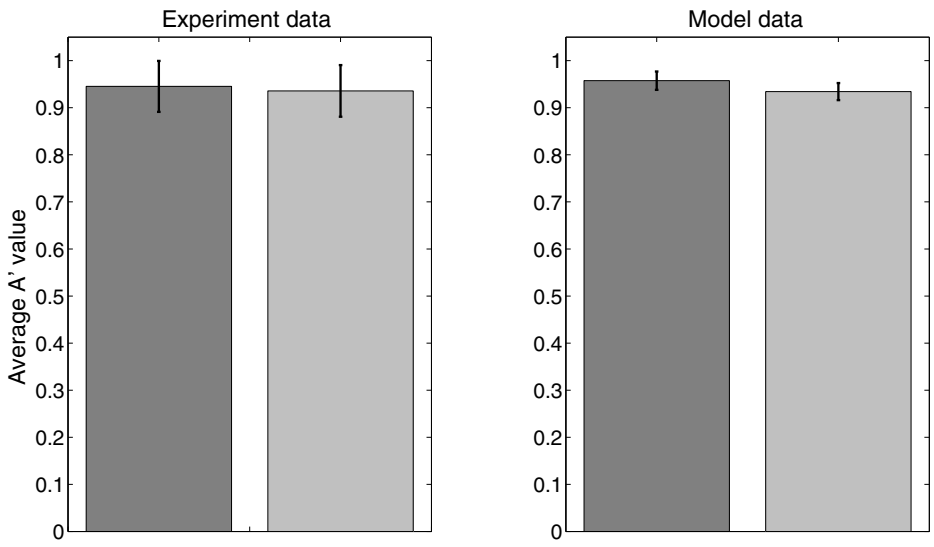


Figure 4.8: Shown are the average A' values obtained in the discrimination experiment (left) and with the model (right). The experimental values are for Catalan-dominant early bilinguals. Dark grey: /ɛ/→/e/ change. Light gray: /e/→/ɛ/ change. Error bars represent standard deviations.

performance in the experiment (left part of figure 4.8) shows practically equal values on both directions of change, indicating that the performance is indeed equal at the phoneme level for the same kind of bilinguals who displayed asymmetry across stimuli in the lexical decision task.

4.4. Discussion

We take the modelling results to suggest that model 2 (depicted in figure 4.3) is the appropriate architecture for showing the asymmetry effect found in Sebastián-Gallés et al. (2005). The failure of model 1 combined with the results of our discrimination task simulations substantiate the

appropriateness of model 2 by showing that as soon as an asymmetry occurs, neither an LDT nor a discrimination task shows the appropriate behaviour. Furthermore, the results of the discrimination task simulations strengthens our belief that the asymmetry in the LDT task is a lexical effect and thus does not reflect an alteration of representations at the phoneme level, since such an alteration (weight asymmetry) deteriorated performance in the simulation of the LDT.

Based on the results obtained in our simulations, we maintain that to explain the effect observed in Sebastián-Gallés et al. (2005), the context of the word is needed, i.e., it is not merely a modification of phoneme categories that underlies this behaviour (as indicated in Norris et al. (2003)), but a specific modification in the connections between phoneme pools which together represent a new word in the lexicon. Hence, the acceptance of the mispronounced word as a word is a lexical level phenomenon. Furthermore, based on our modelling of a discrimination task, we expected the performance on a real discrimination task of subjects with the same background as those showing the asymmetry in the LDT in Sebastián-Gallés et al. (2005), to tend to be perfect for both directions of change.

In the subsequent psychophysical experiments, we showed that the prediction from the modelling results is indeed reasonable, as the very same kind of subjects as in Sebastián-Gallés et al. (2005) showed the asymmetry found earlier in a lexical decision task, but practically equally high performance on a syllable discrimination task involving the same critical phonemes as in the lexical task. This indicated that a new word form stored in the lexicon could account for the asymmetry in the results, without the information having affected phoneme category representations, since performance at the phonemic level was virtually equally high for both categories.

In a recent experimental paper, we found evidence in this direction (Connine 2004). The author of that work used a phoneme identification experiment to examine the representation of phonological variants, in specific the flap variant of spoken words which according to spelling are pronounced with a /t/, such as /prɪdɪ/ instead of /prɪtɪ/. She found that her results provided "strong evidence for the claim that representation of auditory form includes explicit representations of the frequently heard variant" (p. 1088). She claimed that "words with highly frequent auditory forms (e.g., containing a flap) that differ from their orthographic forms may develop parallel representations in the auditory lexicon" (p. 1088) and her results indicated that listeners did not recode the flap variant into an underlying /t/ version, but rather recognized it via a pre-existing lexical representation of the word containing this variant. Thus, those results

are strongly supportive of our claim that the frequently heard variant /galleda/ could have induced in the subjects an explicit new lexical entry through long-term plasticity (Hebbian learning), rather than a change in phoneme categories. Consistent with these results, Cutler, Weber & Otake (2006) recently used an eye-tracking paradigm to demonstrate an asymmetric mapping from phonetic to lexical representations in Japanese listening to English (their L2). Specifically, the authors found that even though such subjects fail to make a distinction between the isolated English phonemes /r/ and /l/, they still have lexical entries for words containing both phoneme categories. In the paper, the purported explanation for these seemingly paradoxical results was that the lexical entries had been stored through direct language instruction, although it was recognized that this would not explain if similar phenomena were to be found in subjects who had not acquired their L2 through explicit teaching. While the situation in Cutler et al. (2006) is distinct from the one we have investigated, seeing as we focus on L1 listening in bilinguals, it can nevertheless be argued that the false-alarm recognition of, e.g., the non-word /galleda/ in our case has similar root causes, namely two minimal-pair lexical entries /gallɛda/ and /galleda/ having been stored, although through environmental exposure instead of explicit language instruction, especially in the case of /galleda/, which is never taught as an actual word at school in Catalonia.

While we in this work have shown a 'proof of principle' of the feasibility of using detailed neurodynamical modelling for higher cognitive phenomena, such as language perception, a possible extension of this model would be to incorporate lower-level auditory processing, thus connecting our more representative firing rate input with the underlying stages, such as frequency mapping and spectrotemporal response fields in auditory cortex (Kowalski et al. 1996) or even models of the basilar membrane distributed frequency mapping in the cochlea. One good model of peripheral auditory processing is the Development System for Auditory Modelling (DSAM, O'Mard & Meddis 1997), whose output is the actual spike trains of auditory nerve fibers. This kind of lower-level processing output could then be fed to higher-level networks, such as ours. One interesting possibility would be to use the approach found in Drew & Abbott (2003), where the authors also use an integrate-and-fire model and show that in order to get the model to process more than two sequential (birdsong) syllables, all one needs to do is to expand the number of pools in the network from the original four by an additional two per syllable. This would be interesting to explore in combination with a more realistic input, e.g., a filter bank or the DSAM model, presented with the real sound files of three-syllable words, which are then processed by a network able to handle three syllables. The

principle on which the work by Drew & Abbott (2003) is based is a general one, so although the spectral complexity of human language syllables is higher than more harmonic birdsong syllables, this should be feasible. One could also look at more technical or mathematical extensions of the model, including explicit incorporation of plasticity in order to study how the resultant changes we here hypothesize come about through experience.

For the representation of a single phoneme used in our model, we find support in an interesting study by Houde & Jordan (1998), which looks at adaptation in vowel production following a training phase with real-time feedback of the subject's own transformed speech. The transformation corresponded to a shift in the vowel formant space, and the adaptation in the production of test words went in the other direction, so as to counter the transformation. The adaptation of the vowel /ɛ/ generalized to different word contexts than the ones presented during training, something taken by the authors as evidence for a common representation of the production of the vowel /ɛ/ in the brain, shared by the production representations of the training words. Although the Houde & Jordan (1998) study focused on speech production, we argue that their results indicate that it is reasonable to assume a representation of words composed of representations of its constituent phonemes, also for speech perception. As for the neural substrate of this representation, the results of a study by Obleser, Lahiri & Eulitz (2004), using magnetoencephalography to measure the neural activity evoked by tokens of German vowels, point to the auditory association cortices, i.e., areas of the auditory cortex peripheral to the primary auditory cortex, A1.

Concerning the discrimination experiment, it could be possible that the lack of differences between the two directions of change be due to a ceiling effect (although extensive piloting was done to avoid it). While this explanation cannot be totally discarded, further experimental results, both with behavioural and electrophysiological measurements, speak against such a possibility. For instance, the results of Sebastián-Gallés et al. (2009), measuring the MMN in an odd-ball paradigm, show no differences between the two directions of change, thus confirming the present behavioural study and supporting the modelling.

5 Discussion and Conclusions

5.1. Discussion

This thesis has a wide scope and its results therefore have wide-ranging potential implications. The largest part of the thesis presented a novel biophysically plausible mechanism of sensory attribute selectivity improvement, here used to explain the phenomenon of differential suppression seen in neurophysiological experiments exploring neural correlates of auditory streaming (Fishman et al. 2001, Fishman et al. 2004). In chapter 2 we explain this mechanism, *dynamic frequency selectivity*, Essentially, with the three ingredients of the non-linearity of processing in neurons as manifested by their threshold, the approximate linearity of their input-output function (above threshold), and the existence of synaptic depression (input causing the depletion of finite synaptic resources over time) at the thalamocortical synapses, one can with a feed-forward model reproduce differential suppression of non-BF tones in a sequence of alternating BF and non-BF tones. We present three different implementations of such a model, one being a rate model (Wilson & Cowan 1972) and the other two being conductance-based (Hodgkin-Huxley, HH) models (Hodgkin & Huxley 1952, Soto et al. 2006). The rate model and approach one to the HH model both use deterministic synaptic depression (see sections A.1.2 and B.1.1; Tsodyks et al. 1998), whereas HH approach two uses stochastic synaptic depression (section B.1.2), as employed in modelling studies of the nature of synaptic transmission (de la Rocha & Parga 2005) and of responses to sequences of (non-differing) haptic stimuli in the primary somatosensory cortex (de la Rocha & Parga 2008).

In chapter 3 we then go on to find that a mexican-hat type IC connectivity can be used to explain approximate co-tuning of inhibition and excitation as well as flank dominance of inhibition over excitation (Wehr & Zador 2003, Wu et al. 2008). Furthermore, across the tonotopic axis we get a 67% boost of excitatory conductances (equivalent to currents),

and also reproduce the spread of sub- and suprathreshold activity, as seen in data found by Liu et al. (2007). We first found a basic fit using only feed-forward inhibition and recurrent excitation, and then proceeded to make a more rigourous fit with all possible intracortical connections active, both described in the Results section of chapter 3. After that, using our basic network fit, we were able to approximately reproduce results on physiological forward masking seen in Wehr & Zador (2005) (section 3.3.3). Finally, and *importantly*, we managed to use the rigourous network fit to reproduce auditory streaming correlate data, this time *with* thalamocortical axonal spread (something which diluted the feed-forward network fit due to the absence of intracortical currents, particularly recurrent excitation).

We now go on to emphasize the main results found in this thesis.

1. We show that the underlying mechanisms responsible for neurophysiological forward masking and auditory streaming data, while sharing the ingredient of (thalamocortical) synaptic depression, are essentially distinct, finding that the former corresponds to the transient regime of the dynamics of synaptic depression, while the latter is readily explained at steady-state (in contrast with earlier explanations; Fishman et al. 2001, Fishman et al. 2004).

2. We find a plausible connectivity scheme for layer IV of the primary auditory cortex, rigourously constrained by data as well as able to explain several important results from the neurophysiology literature (see, e.g., Wehr & Zador 2003, Wehr & Zador 2005, Liu et al. 2007, Wu et al. 2008, Tan & Wehr 2009, Wang et al. 2010). Using the resulting network setup, we then went on to show that intracortical currents make a crucial contribution to fitting the DS of Fishman et al. (2004) when thalamocortical afferent spread is non-zero ($\nu > 0$).

3. We have to our knowledge constructed the first neurophysiologically well-founded model to explain the differential suppression phenomenon exhaustively. We expect this model may serve as an input layer for the presumably forthcoming more complicated models aiming at explaining the complete phenomenon of auditory streaming, including switching between 'perceptual' states.

4. Finally, we have proposed an attractor network model (model 2 of chapter 4) which not only reproduced data from the lexical decision task by Sebastián-Gallés et al. (2005), thus strengthening the hypothesis that the acceptance of the mispronounced word as a word is a lexical level phenomenon, but also made the prediction (modelling a

phoneme discrimination task) that those same subjects displaying the acceptance/asymmetry would perform near-perfectly on a real syllable discrimination task. This was then experimentally confirmed by our collaborators. We have published these and other results (Larsson et al. 2008, Sebastián-Gallés et al. 2009).

Finally, while there is some overlap between the work in chapter 3 of this thesis and that by de la Rocha et al. (2008), we make a further contribution by studying the tuning of inhibition and excitation using a more detailed conductance-based model (as opposed to their rate model). Above all though, our overarching goal was to explain auditory streaming and forward masking correlates in the auditory cortex, and in so doing we covered subjacent phenomena such as (approximate; Wu et al. 2008) co-tuning of inhibition and excitation in order to have a firm foundation for our model's parameter values in the literature.

Some of the results we have obtained were also qualitatively found by Loebel et al. (2007), using a model with intracortical depression in excitatory A1 neurons. However, in their model inhibitory neurons have no depression (or facilitation) in their incoming intracortical synapses and thalamocortical depression is also absent. Most significantly, however, the solution of Loebel et al. (2007) depends strongly on the specific imposed tonotopic organization of the spontaneous activity of their network (we have confirmed this in our own implementation (data not shown), and it is also explained in the first study using this population spike model — see Tsodyks, Uziel & Markram 2000). For these reasons we believe that, while theirs is a legitimate alternative approach to modelling A1, our model finds a more minimal mechanism for explaining neurophysiological correlates not only of forward masking, but also of auditory streaming (something not addressed by them). Further, we incorporate a greater neurophysiological detail (mainly chapter 3), based on the recent literature (e.g., Liu et al. 2007, Wu et al. 2008, Wang et al. 2010).

Bartlett et al. (2010), performing *in vivo* experiments on awake marmosets, found a shape of responses in MGBv (exponential across the tonotopic axis) very similar to those employed in our study. However, in a majority of neurons they find the BW at half maximum amplitude to be about 0.2 octaves, which is less than the width employed by us. We based our values on the earlier study by Liu et al. (2007) and found no reason to alter these once the new study came out, both because of our considerable data fits using the bigger value and because, in any case, this issue still seems to be contentious and assessments have varied over the years and

species (e.g., see Aitkin 1973, Calford 1983, Ma & Suga 2009). Bartlett et al. (2010) summarize this issue in table 1 of their paper.

As for the shape of the *cortical* response to tones, Fishman & Steinschneider (2009) lends further support to the exponential shape of responses in A1, in particular their 'on' responses (10-30 ms after stimulus onset) resemble both that seen by Liu et al. (2007) and consequently that which we obtain in our simulation results. Also, Bartlett et al. (2010) show an exponential shape of the response profile across A1, with about 25 % of responses less than 0.2 octaves wide, the rest wider (we note that this figure also is much lower than in Liu et al. (2007), which may be due to species or methodology differences).

Regarding the broad inhibition seen in the study by Liu et al. (2007), recently the same group of researchers have published another study showing a very similar organization of inhibition, this time in the visual cortex of mice (Liu, Li, Ma, Pan, Zhang & Tao 2011). However, in the latter study, excitation in simple cells was also quite broad, unlike what we deduced from the former study's A1 data. In fact, cases where excitation is about as broad as inhibition have also been found in A1 (very recently measured in *in vitro* samples from mouse; Levy & Reyes 2012). This is definitely something that might need a revision in our modelling of intracortical connectivity, as presented in this thesis, although as we noted in the discussion of chapter 3, Levy & Reyes (2012) focus on LII/III and we have modelled layer IV.

5.1.1. Limitations

We have made several simplifications in the present work, for several reasons. The foremost among them was to limit the number of free parameters and thus simulation time. We have performed robustness tests in those cases where we considered the simplifications to warrant it, confirming that obtained results were reproducible in key cases with more realistic settings (data only partially shown). That said, our conductance-based models are very complex and we retained a considerable number of degrees of freedom with the setup we used. The brain (and, for that matter, any neuron or synapse) is very complex and all models are necessarily simplifications of reality (Sejnowski, Koch & Churchland 1988, Abbott 2008). Our models are no exception to this rule. For completeness, however, we will now proceed to listing some things that are candidates for inclusion in a future, more complex model (all the while emphasizing that each parameter's inclusion would need to be very well motivated first). Thus,

straightforward model 'ingredients' which may potentially have an impact on our results, that we did *not* include, were

1. intracortical NMDA receptors,

2. thalamocortical NMDA receptors, and

3. $GABA_B$ receptors.

Regarding the absence of NMDA receptors, this was a conscious limitation imposed on our model in order to keep complexity down. Their presence in cortex is undisputed (e.g., Thomson & Deuchars 1994), and we recognize that working out the implications of their inclusion in our model is of the utmost importance for future work. However, whether NMDA receptors are part of the thalamocortical interface of the auditory modality is unclear, especially in the adult organism. While there have been studies showing such receptors indeed participate in thalamocortical transmission to somatosensory (Gil & Amitai 1996, Armstrong-James, Welker & Callahan 1993) and visual (Fox, Sato & Daw 1989, Miller, Chapman & Stryker 1989, Krukowski & Miller 2001) cortices, we have only found two studies directly addressing this issue in the auditory modality, in our view inconclusively. While Cruikshank et al. (2002) showed NMDA receptors were putatively present in P13-P19 mice (*in vitro*), researchers from the same group later found that the signature long-lasting intracortical depolarization, taken in their previous study as an indication of thalamocortical NMDA, was highly likely to be polysynaptic in nature. In fact, scrutinizing the data of Cruikshank et al. (2002), one finds that also here the later onset which signals polysynaptic inputs was present. We emphasize that influential studies have concluded that the presence of NMDA is particular to the developmental phase of the animal and its nervous system (Thomson & Deuchars 1994, Crair & Malenka 1995), casting some doubt on its role in the adult thalamocortical transmission mechanism. Certainly, this issue will need to be followed closely in the years ahead.

The issue of $GABA_B$ receptors was briefly touched upon by Wehr & Zador (2005), who found that the $GABA_B$ component constituted only 12% of the total inhibitory conductance amplitude evoked by an isolated tone. However, it was of a long duration (200-300 ms), rendering it possible that it have an effect in spite of its low amplitude. Curiously, Wehr & Zador (2005) found that blocking $GABA_B$ actually *enhanced* forward suppression (masking), which further adds to the motivation of clarifying its role in future work.

Another thing we have not addressed in this work is using a certain probability of connection between cortical neurons (e.g., probability of monosynaptic connection between two neurons, less than 200 μm apart, was found to be 0.1 by Markram, Lübke, Frotscher & Sakmann 1997). Instead, we used all-to-all connectivity (with probability 1.0), but let conductance strength of synaptic connections fall off with distance (see appendix B). We have not investigated the difference between these two approaches and such a comparison would be interesting to carry out, including looking at its implications for our results. Here, we limit ourselves to commenting on this. Briefly, considering that the tonotopic (anterior-posterior) extent in A1 of rat was found to be 2 mm and that this represented approximately 6 octaves (1 to 60 kHz) (Kilgard & Merzenich 1999), we can deduce that the probability of connection is 0.1 up to an interneuronal distance of 0.6 octaves in A1. If neurons further apart are not connected at all, that would have little impact on our results beyond very slightly lowering lateral inhibition (excitation too narrow to be affected). Taking the above data as a starting point, what would need to be investigated in more detail is what happens *within* the 0.6 octaves closest to a stimulated neuron in either direction, i.e., the 1.2 octaves surrounding that neuron, something very important for our robustness study of chapter 3. Interestingly, very recently Levy & Reyes (2012) measured connection profiles in A1 of mice and gave values of widths in μm. Combining their data with the result of Stiebler, Neulist, Fichtel & Ehret (1997) that the mouse A1 covers about six octaves per mm, one also arrives at values of about 0.6 octaves for the intracortical connectivity widths of all types measured. However, their values for probability of release in said connections were generally higher than 0.1 (Levy & Reyes 2012).

Similarly, the fact that we used coupled intracortical connection profiles ($\lambda_I = \lambda_{\mathrm{II}} = \lambda_{\mathrm{EI}}$ and $\lambda_E = \lambda_{\mathrm{EE}} = \lambda_{\mathrm{IE}}$) in this work, while aiding us in limiting the complexity of the network, probably warrants a future scrutiny. Particularly, by relaxing this constraint, the elaborate network fit found in chapter 3 would probably have multiple companion solutions.

As for the heavy dependence of our work on the existence of thalamocortical depression, as we already argued, there are several studies supporting this (Abbott et al. 1997, Carandini et al. 2002, Chung et al. 2002, Elhilali et al. 2004). However, e.g., Richardson, Blundon, Bayazitov & Zakharenko (2009) point out that while thalamocortical synaptic transmission takes place through stubby dendritic spines in primarily layer IV (Nahmani & Erisir 2005), synaptic plasticity in pyramidal neurons has been shown to take place on mushroom spines (found in hippocampus; see Matsuo, Reijmers & Mayford 2008). In view of experiments demonstrating auditory

cortex plasticity during learning (e.g., Kilgard & Merzenich 1998*a*, Kilgard & Merzenich 1998*b*, Fritz, Shamma, Elhilali & Klein 2003, Froemke, Merzenich & Schreiner 2007), they speculate that this be a guarantor that "plastic changes in the auditory cortex during learning [...] not compromise the effectiveness of thalamocortical projections and consequently the reliability of the delivery of acoustic information to the auditory cortex". This is a valid point (although the studies cited were not all in auditory cortex or even cortex), but as we have seen and others (de la Rocha & Parga 2005) have shown, with enough variability the transmission of information is not disrupted by the presence of thalamocortical depression, even in the steady state reached after much input has been processed, provided the network is in the so-called fluctuation driven regime.

Finally, throughout this study we used an equal number of excitatory and inhibitory neurons, in order to limit the already considerable complexity involved in our task. Seeing as inhibitory neurons are typically considered to constitute only 15-30% of all cortical neurons, the repercussions of this simplification in our results would need to be elucidated in future work. In principle, however, by changing the relative input strength to each kind of neuron (nothing seems to be set in stone in neurophysiology), one could find equally satisfactory alternative solutions to the problems we have addressed.

5.2. Conclusions

5.2.1. Closing words

In summary, we have discovered a novel, biophysically plausible and robust mechanism for explaining the neural correlates of auditory streaming as measured in the primary auditory cortex in neurophysiological studies (notably Fishman et al. 2001, Fishman et al. 2004, Micheyl et al. 2005). The phenomenon we explain is called differential suppression, and we refer to our mechanism as dynamic frequency selectivity. Crucially, while this work has concentrated on the auditory modality, we see no reason why the same mechanism could not be used to explain selectivity improvement over time in any sensory modality.

Furthermore, we have proposed an organization of the intracortical connectivity of A1 that is consistent with several recent results in the literature (see, e.g., Wehr & Zador 2003, Wehr & Zador 2005, Liu et al. 2007, Wu et al. 2008, Tan & Wehr 2009, Wang et al. 2010). This same circuit can be used to sharpen responses of our feed-forward network with thalamocortical fan-out, thus showing a crucial role for the recurrent excitation and, to a

lesser extent, the (not exceedingly) lateral inhibition of our connectivity in maintaining our dynamic frequency selectivity mechanism robust when including this neuroanatomical constraint (shown by many authors, e.g., Miller et al. 2001)

Finally, we have proposed an attractor network model (model 2 of chapter 4) which reproduces data from a lexical decision task (Sebastián-Gallés et al. 2005), which predicted that Catalan-dominant bilinguals do not alter their phoneme categories, although showing signs of having stored a new word variation in the lexicon. This was then corroborated by our collaborators in a phoneme discrimination study, where the same subjects which displayed asymmetry in the LDT showed near-perfect discrimination of those very phonemes which were contrasted in the non-words of the LDT. These results, along with further follow-up experimental work employing ERPs, has been published in two different papers (Larsson et al. 2008, Sebastián-Gallés et al. 2009).

5.2.2. Future work

There are several interesting topics which have been touched upon during the work on this thesis, but for spatiotemporal reasons were not included. Here we mention a few.

First of all, while we found a very plausible and well-founded explanation for the neurophysiological data from auditory streaming correlate experiments using short tones (notably that by Fishman et al. 2004), we never attempted to reproduce results with longer tones (TD of 50 or 100 ms), which show an abrupt decline in RR for low Δf values when SOA becomes zero (i.e., for $PR = 20$ and 10 Hz, respectively). With the feed-forward solution of chapter 2 we are positive that this data can *not* be reproduced, for various reasons. First of all, with our input function ($\mu =$ 0.4 oct.) the difference in input at $\Delta f = 5$ % is at most around 10%, which means that for the response to B tones to be almost zero that close to φ_A, d_2 for the PR values in question would have to be unrealistically low in order to bring input below threshold (it would affect the rest of the fit negatively). Therefore, we are convinced that the full tonotopic network, possibly with some now left-out ingredient (see above) is needed in order to account for that data. This is something we definitely would like to look into in the immediate future.

Secondly, and possibly related to the previous point, it would be interesting to further investigate the relationship between neurophysiological results on forward masking and auditory streaming, as measured in A1.

We have been able to explain a great deal, but using slightly different parameter values for different phenomena. While we showed analytically that there is a certain redundancy in depression parameter values (two different sets of U and τ_D values can yield the same steady-state value of depression), it would still be desirable to find *one* network setup to explain 'all' relevant results. At the very least, we would like to employ the rigourous fit of the tonotopic network chapter 3 to simulating also the results by Wehr & Zador (2005), now only reproduced with the basic fit.

Thirdly, using the software we have developed, with slight changes, one could easily investigate the response of the tonotopic axis of A1 to frequency-modulated sweeps. During such a project, one could further elucidate the relative contributions of candidate mechanisms for direction selectivity, such as differential delays of excitatory inputs and spectral offset between excitatory and inhibitory inputs (Ye, Poo, Dan & Zhang 2010, Gittelman & Pollak 2011). As the authors of the cited studies point out, the spike generation mechanism, whose role in the present work was prominent, would also here play a role.

As the fourth point, we would like to emphasize the strong interest we have in extending our model to also study and attempt to explain mechanistically the full auditory streaming phenomenon, as manifested in the brain. While possibly a daunting challenge using such detailed models, a starting point could be to use our model as an input layer to a hypothetical higher-level layer where 'percepts' are formed based on the activity of the A1 layer.

Finally, we have some ideas for improvements for our model of a lexical decision task and phoneme discrimination. For one thing, the approach taken by Drew & Abbott (2003) in modelling birdsong syllables could work in our case too (LDT task). This would entail having up to eight pools for 'perceiving' three-syllable words, as those in our example stimuli. Another idea would be to have all pools represent syllables, which in our case would mean having a total of eight pools, representing /ga/, /lle/, /llɛ/, /da/, /fin/, /es/, /ɛs/, /tra/. However, seeing as there is still debate about whether words are represented in the brain by syllables or phonemes, or both (e.g., Siok, Jin, Fletcher & Tan 2003), one would have to design such a study with care.

Appendices

A Rate model and analytics

A.1. Rate model of local A1 population

The equation used to describe the temporal evolution of the mean firing rate $m(t)$ of a population of neurons in A1 is the rate model (Wilson & Cowan 1972)

$$\tau \frac{dm(t)}{dt} = -m(t) + [I_{Th}(t) - \theta]_+, \tag{A.1}$$

where the time constant $\tau = 20$ ms. The parameter θ represents the threshold of the population's transfer function

$$[x]_+ = \begin{cases} 0 & x \leq 0, \\ x & x \in (0, \theta), \\ 100\text{Hz} & \text{otherwise.} \end{cases} \tag{A.2}$$

The periodic function $I_{Th}(t)$ represents feed-forward sensory inputs (MGBv efferents) that innervate the A1 population. The input function to the population representing neurons at $\varphi_A = f_A$ in A1 is defined as

$$I_{Th}(t) = \int_{F_{low}}^{F_{high}} W(\varphi_A, F)\ R(F,t)\ d(F,t)\ dF, \tag{A.3}$$

where the distribution of thalamocortical synaptic weights is taken as the symmetric function

$$W(\varphi, F) = \frac{W_0}{2\nu} e^{-\frac{|\varphi - F|}{\nu}}. \tag{A.4}$$

Equation (A.3) is a convolution across thalamic tonotopic space, from F_{low} to F_{high}, which represent said space's extremes. According to the experimental results of Miller et al. (2001), thalamocortical projections from MBGv to A1 are focal and have a spread of at most one third of an octave. Hence, in our model we assume that the function W is narrow (small ν). The spatial range μ of the tuning curve in MGBv has been approximately adjusted using the data of Liu et al. (2007), where

thalamocortical excitatory input is broad, spanning up to 4-5 octaves at 60 dB (the sound intensity used by Fishman et al. 2004). The resulting tuning curve in the cortex is similar to the 'on' tuning curves recently measured in A1 (Fishman & Steinschneider 2009).

A.1.1. Thalamocortical input

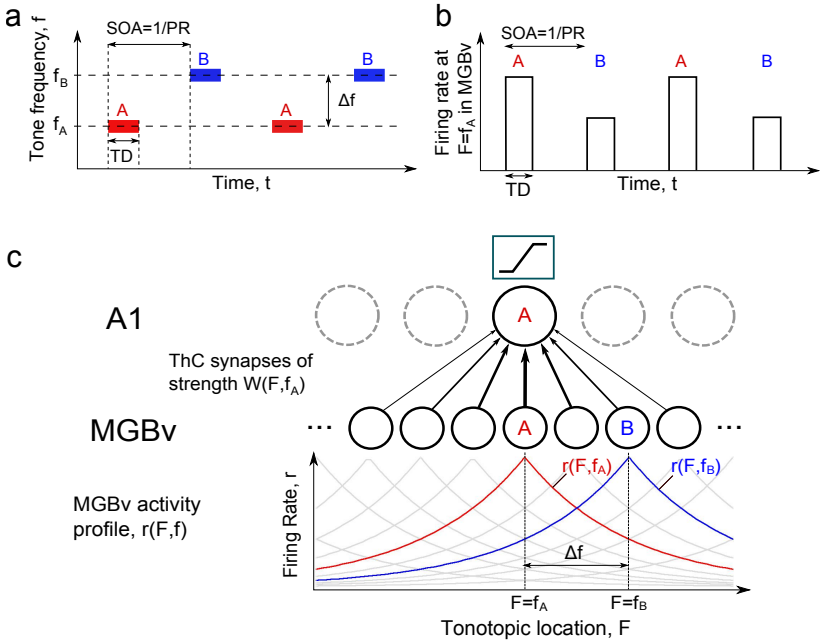


Figure A.1: Description of the model. **a**: Schematic representation of the stimuli used to study auditory streaming, consisting of sequences of alternating A and B pure tones of frequency $f = f_{A,B}$ and duration TD, presented at rate PR . **b**: The temporal dynamics of the firing rate at any particular location F of the thalamus (in this case $F = f_A$) is modelled as a series of short (20 ms) excitatory pulses of magnitude $r(F, f)$. **c**: Thalamic (MGBv) neurons send thalamocortical (TC) axons to make depressing synaptic connections with A1 neurons.

We consider that a pure tone of frequency f elicits an activity pattern in the MGBv described by (see figure A.1 **c**)

$$r(F, f) = \frac{r_0}{2\mu} e^{-\frac{|F-f|}{\mu}}, \tag{A.5}$$

where variable F denotes the tonotopic location in the MGBv. Additionally, μ represents the spatial extent of the activity (in octaves), and $r_0 = 100$Hz.

In the auditory streaming experimental paradigm, pure tones of two distinct frequencies are presented alternately in time, with presentation rate PR (see figure A.1 **a**).

Accordingly, the input function in the model is taken as

$$R(F,t) = r(F,f_A)\psi(t) + r(F,f_B)\psi(t-\frac{1}{PR}), \tag{A.6}$$

where $\psi(x)$ is the periodic function (with period $2/PR$) described as

$$\psi(x) = \begin{cases} 1 & x \in [0,\, TD], \\ 0 & x \in (TD, 2/PR), \end{cases} \tag{A.7}$$

with $x \bmod(2/PR)$. Note that the function ψ is simply a square function, so that the temporal aspects of the thalamic response are not taken into account. This approximation is justified by the fact that we deal with neural responses to short pure tones (25 ms, see Fishman et al. 2001, Fishman et al. 2004). Wallace, Anderson & Palmer (2007) showed that responses in the ventral part of the medial geniculate body (MGBv), are mainly represented by two types of cells: Onset cells, that respond transiently at the tone's onset, and on-sustained cells that show a sustained response with a firing rate above the background firing rate. Thus, we consider that the MGBv neural response simply consists of the transient part of the response, which we model as the square function (A.6) (we will later use time-varying stimuli when using a more detailed Hodgkin and Huxley model; see appendix B).

A.1.2. Synaptic depression model

This model is based on a model of synaptic depression in Tsodyks et al. (1998), with equation (A.8) governing the evolution of the fraction, $d(t)$, of transmitter available for release into the synaptic cleft.

$$\frac{dd(t)}{dt} = \frac{d_{start} - d(t)}{\tau_D} - \frac{d(t)\ U\ \delta(t - t_{pre})}{dt} \tag{A.8}$$

t_{pre} designates the time of arrival of a presynaptic action potential, at which time $d(t)$ depresses by a fraction U, which is a parameter of the model. The time constant in the model, τ_D, governs reactivation of used synaptic vesicles. d_{start} is the value to which the release probability converges exponentially in the absence of presynaptic input. Throughout this study (except when explicitly using steady-state values calculated in section A.2 at the outset of simulations), $d_{start} = 1$ (valid for very low spontaneous

rates in MGBv), whereas U and τ_D vary (in accordance with literature, as explained in main text). The fraction $d(t)$ is independent for each thalamocortical synapse connecting onto a cortical neuron.

A.2. Analysis of synaptic depression dynamics

A.2.1. Synaptic depression variable: Transient behavior and two-tone interactions (forward masking)

The differential equation for the synaptic depression variable (A.8) can be written as

$$\frac{D}{Dt}[d(t)\ e^{\beta t}] = \frac{1}{\tau_D} e^{\beta t} \tag{A.9}$$

where $\beta \equiv 1/\tau_D + UR(F,t)$. Considering the best frequency site φ_A, the function $R(F,t)$ (equation (A.6)) takes on the three constant values $r_A \equiv r(F_A, f_A)$, $r_B \equiv r(F_A, f_B)$ and 0 in a repetitive succession, $r_A, 0, r_B, 0, ...$ Therefore it is simple to integrate the equation (A.9) in these intervals, starting with a certain initial value of the synaptic depression variable $d(t=0) = d_{start}$. We have enumerated the intervals using the parameter n, starting with tone A. The equation (A.9) then has the solutions

$$d_{A,B;n} = \left(d_{A,B;n-1} - \frac{1}{\tau_D\, \beta_{A,B}} \right) e^{-\beta_{A,B} TD} + \frac{1}{\tau_D\, \beta_{A,B}}, \quad n \text{ odd} \tag{A.10}$$

$$d_{A,B;n} = (d_{A,B;n-1} - 1)\, e^{\frac{TD - \frac{1}{PR}}{\tau_D}} + 1, \quad n \text{ even}, \tag{A.11}$$

where $\beta_{A,B} \equiv 1/\tau_D + U\ r_{A,B}$. For large enough values of n, the steady state is reached and the solutions (A.10) and (A.11) converge to the four values $d_0 - d_3$ that are represented in figure A.2.

The lower panel of figure A.2 shows an enlarged view of the steady state of the variable d. The quantities d_{0-3} are defined as the extremes of the function $d(F_A, t)$ in the steady state (i.e., d_{0-3} are time-independent quantities). Using these, it is useful to define the two quantities

$$d_{A,B} \equiv d_{1,3} + (d_{0,2} - d_{1,3})/2, \tag{A.12}$$

that represent an *approximate* measure of the mean synaptic depression values during the presentation of an A or a B tone, respectively. Note that these measures are only defined in the steady state regime, as are the quantities d_{0-3}.

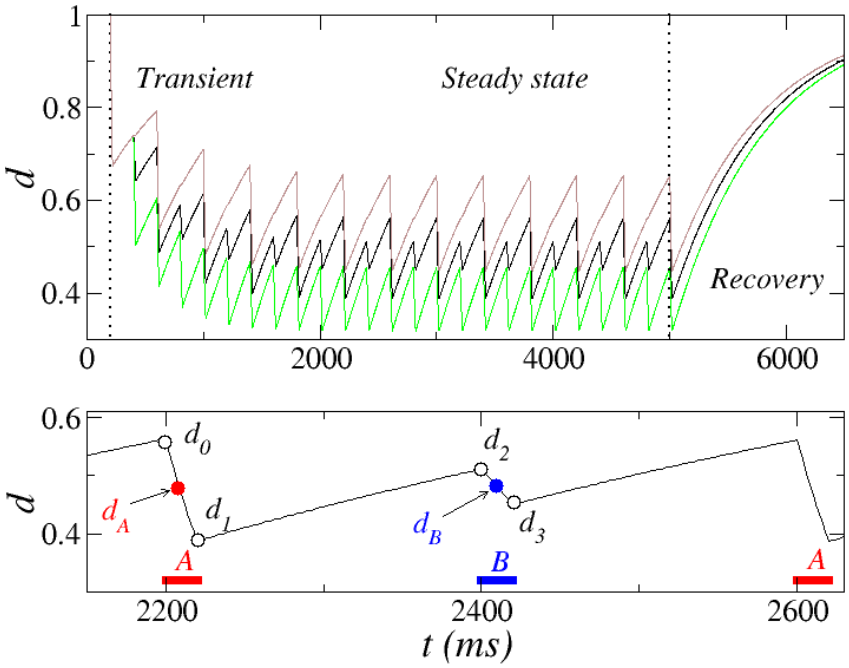


Figure A.2: Dynamics of the synaptic depression variable at the tonotopic location $F = F_A$. Upper panel: Transient behaviour before the steady state regime of the synaptic depression variable is reached, for sequences of inputs (applied in the time interval between the dotted lines) of the same frequency ($\Delta f = 0$; green line), and for different A and B tones, $\Delta f = 0.5$ (black) and $\Delta f = 2$ (grey) octaves, respectively. Lower panel: Zoomed view of the time course of the d variable (corresponding to $\Delta f = 0.5$) once the steady-state regime has been reached. Note that the different inputs A and B (represented in different colors) produce different degrees of depression, but the mean value of the depression variable in the steady state regime satisfies $d_A \approx d_B$ (parameters: $d_{start} = 1$, $U = 0.2$, $R_{\rm Th} = 100$ Hz, $\tau_D = 0.8$ s, $PR = 5$ Hz, $TD = 20$ ms, $\mu = 0.5$ oct.).

In figure (A.2) we have considered the synaptic depression variable to be fully recovered before the presentation of the sequence of alternating inputs (i.e. $d(t=0) \equiv d_{start} = 1$). Three different time series are shown. One with $\Delta f = 0$ (i.e. the sequence simply consists of a succession of A tones repeated with $PR = 5$ Hz) another one with $\Delta f = 0.5$ oct. and a final one with $\Delta f = 2$ oct. In this latter case the sequence consists of a sequence of A tones with half of the presentation rate, i.e. $PR = 2.5$ Hz.

Analysing further the equations (A.10,A.11), it is possible to see that the decay of the $d(t)$ function follows a decreasing exponential function with time constant

$$\tau_t = \frac{1}{TD\ PR\ \ U(r_0 + r_B) + 2/\tau_D} \tag{A.13}$$

Thus, the decay rate is slower when TD is smaller than $1/PR$, and increases with increasing PR.

The left panel of figure A.3 shows the dependence of d_A (see eq. (A.12)) on the PR for two values of Δf. The function d_A is a decreasing function of PR, simply because increasing the presentation rate increases the amount of synaptic transmission and thus, of synaptic depression. It is important to note that $d_A \approx d_B$, especially at high presentation rates and small values of Δf. This is shown in the inset of figure A.3 for the two corresponding values of Δf. In the main text, we make use of this property and employ the notation $\bar{d} \equiv \frac{d_A + d_B}{2}$ to indicate that then we are considering a unique value of the mean synaptic depression, independent of the presented tone.

A.2.1.1. Forward masking analysis

In the two tone paradigm used in forward masking experiments, a probe tone of a certain frequency $f = f_P$ (at the BF of a neuron from which is being measured) and intensity (here, rate) $r = r_P$ is presented some time S after the presentation of a masker tone of frequency $f = f_M$ and intensity (rate) $r = r_M$ (f_M and r_M may vary; Calford & Semple 1995, Brosch & Schreiner 1997, Wehr & Zador 2005). In this case the solutions (A.10,A.11) become

$$d_M = \frac{1}{\tau_D\, \beta_M} \left[1 + (\tau_D\ \ \beta_M\ d_{start} - 1) e^{-TD\, \beta_M}\right] \tag{A.14}$$

$$d_P = 1 + e^{\frac{TD - \frac{1}{PR}}{\tau_D}} \left[\frac{1}{\tau_D\, \beta_M} (\tau_D\, \beta_M\ d_{start} - 1) e^{-TD\, \beta_M} - 1\right] \tag{A.15}$$

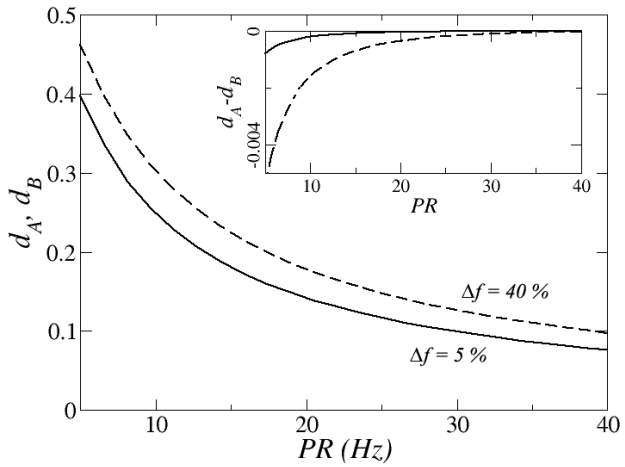


Figure A.3: Mean values of the synaptic depression quantity d_A defined in equation (A.12), as a function of PR for $\mu = 0.5$ octaves, $U = 0.2$ and $\Delta f = 40$ % (dashed curve), and $\Delta f = 5$ % (solid curve). Inset: Difference $d_A - d_B$, showing that $d_A \approx d_B$ for $\Delta f = 40$ % (dashed curve) and $\Delta f = 5$ % (solid curve), particularly at fast PR. All the curves have been analytically obtained solving the equations (A.10) and (A.11).

where $\beta_M \equiv 1/\tau_d + U r_M$. The time needed for the synaptic depression variable to recover to a certain value d_c can then be calculated using equation (A.15) and gives

$$t_{rec}(d_c) = \tau_D \log\left(\frac{e^{-TD\left(-\beta_M+\frac{1}{\tau_D}\right)}\left(e^{TD\,\beta_M}(1-\beta_M\tau_D)+d_{start}\;\beta_M\;\tau_D\;-1\right)}{\beta_M\;\tau_D\,(d_c-1)}\right). \tag{A.16}$$

With this equation (and taking into account the thresholding effect of the neural population) it is possible to calculate the recovery curves shown in figure 2.13.

A.2.2. Synaptic depression variable: Steady state regime

Integration of equation (A.9) between two arbitrary intervals is straightforward, and therefore for simple enough $R(F,t)$ functions it is possible to obtain analytic expressions for $d(t)$ and for the extremes d_{0-3} (see figure A.2). For such simple inputs, the equation (A.9) can be integrated in

the steady state regime, and we obtain the four equations:

$$d\ e^{t(1/\tau_D + U r_A)}\Big|_{d_0,t_0}^{d_1,t_1} = \frac{e^{t(1/\tau_D + U r_A)}}{1 + \tau_D\, U r_A}\Big|_{t_0}^{t_1} \tag{A.17}$$

$$d\ e^{t/\tau_D}\Big|_{d_1,t_1}^{d_2,t_2} = e^{t/\tau_D}\Big|_{t_1}^{t_2} \tag{A.18}$$

$$d\ e^{t(1/\tau_D + U r_B)}\Big|_{d_2,t_2}^{d_3,t_3} = \frac{e^{t(1/\tau_D + U r_B)}}{1 + \tau_D\, U r_B}\Big|_{t_2}^{t_3} \tag{A.19}$$

$$d\ e^{t/\tau_D}\Big|_{d_3,t_3}^{d_0,t_4} = \frac{e^{t/\tau_D}}{1 + \tau_D\, U r_B}\Big|_{t_3}^{t_4}, \tag{A.20}$$

where $t_0 = 0$, $t_1 = TD$, $t_2 = \frac{1}{PR}$, $t_3 = TD + \frac{1}{PR}$ and $t_4 = \frac{2}{PR}$ (relative to any A tone's onset, at steady state). This system can be solved for the unknowns d_{0-3}. In particular this permits us to obtain an analytical expression for the mean quantities $d_{A,B} = d_{A,B}(PR\,, \Delta f)$, using the definition (A.12).

An alternative way of calculating d_{0-3}, by way of their dynamic trajectories, $d_{0-3}[m]$, $m \in 1, 2, \ldots$, is to first formulate a system of difference equations, parting from the expressions (A.10) and (A.11). We get a system of four equations by writing down said expressions separately for r_A and r_B, and setting $d_0[m] \equiv d[i]$, $i = 4n$, $d_1[m] \equiv d[i]$, $i = 2n - 3$, $d_2[m] \equiv d[i]$, $i = 4n-2$, and $d_1[m] \equiv d[i]$, $i = 4n-1$, respectively ($n \in \mathbb{N}$). The resulting system is written

$$d_0[m] = (d_3[m-1] - 1)\, e^{\frac{TD\, -\, \frac{1}{PR}}{\tau_D}} + 1 \tag{A.21}$$

$$d_1[m] = \left(d_0[m] - \frac{1}{1+\xi_A}\right) e^{-\frac{TD\,(1+\xi_A)}{\tau_D}} + \frac{1}{1+\xi_A} \tag{A.22}$$

$$d_2[m] = (d_1[m] - 1)\, e^{\frac{TD\, -\, \frac{1}{PR}}{\tau_D}} + 1 \tag{A.23}$$

$$d_3[m] = \left(d_2[m] - \frac{1}{1+\xi_B}\right) e^{-\frac{TD\,(1+\xi_B)}{\tau_D}} + \frac{1}{1+\xi_B}, \tag{A.24}$$

where $\xi_{A,B} \equiv r_{A,B}\, U\, \tau_D$, and $d_3[0] = d_{start}$. Using the value $d_{start} = 1.0$, we obtain solutions for the dynamic trajectories $d_{0-3}[m]$ (extensive equations,

not shown due to spatial constraints). Then, by letting $m \to \infty$ in those expressions, we obtain the steady state, or stationary, solutions:

$$\begin{aligned} d_0 = & -\frac{1}{(1+\xi_A)(1+\xi_B)\left(e^{\frac{2+PR\ TD\ (\xi_A+\xi_B)}{PR\,\tau_D}}-1\right)} \\ & \times\left[\xi_B\left(1+(1+\xi_A)\left(e^{\frac{1+PR\ TD\ (1+\xi_A+\xi_B)}{PR\,\tau_D}}-e^{\frac{1+PR\ TD\,\xi_A}{PR\,\tau_D}}\right)\right)\right. \\ & \left.+(1+\xi_B)\left(\xi_A e^{\frac{TD\,(1+\xi_A)}{\tau_D}}-(1+\xi_A)e^{\frac{2+PR\ TD\ (\xi_A+\xi_B)}{PR\,\tau_D}}\right)+1\right], \end{aligned} \tag{A.25}$$

$$\begin{aligned} d_1 = & \frac{e^{-\frac{TD}{\tau_D}}}{(1+\xi_A)(1+\xi_B)\left(e^{\frac{2+PR\ TD\ (\xi_A+\xi_B)}{PR\,\tau_D}}-1\right)} \\ & \times\left[(1+\xi_B)\left(e^{\frac{2+PR\ TD\ (1+\xi_A+\xi_B)}{PR\,\tau_D}}+\xi_A e^{\frac{2+PR\ TD\,\xi_B}{PR\,\tau_D}}\right)\right. \\ & \left.+(1+\xi_A)\left(\xi_B e^{\frac{1}{PR\,\tau_D}}\left(1-e^{\frac{TD\,(1+\xi_B)}{\tau_D}}\right)-(1+\xi_B)e^{\frac{TD}{\tau_D}}\right)\right], \end{aligned} \tag{A.26}$$

$$\begin{aligned} d_2 = & \frac{1}{(1+\xi_A)(1+\xi_B)\left(e^{\frac{2+PR\ TD\ (\xi_A+\xi_B)}{PR\,\tau_D}}-1\right)} \\ & \times\left[(1+\xi_B)\left((1+\xi_A)e^{\frac{2+PR\ TD\ (\xi_A+\xi_B)}{PR\,\tau_D}}\right.\right. \\ & \left.-2\xi_A\sinh\left(\frac{TD\,(1+\xi_A)}{2\tau_D}\right)e^{\frac{2+PR\ TD\,(1+\xi_A+2\xi_B)}{2PR\,\tau_D}}\right) \\ & \left.-(1+\xi_A)(\xi_B e^{\frac{TD(1+\xi_B)}{\tau_D}}+1)\right], \end{aligned} \tag{A.27}$$

and

$$
\begin{aligned}
d_3 = & \frac{e^{-\frac{TD(7+2\xi_A+3\xi_B)}{\tau_D}}}{(1+\xi_A)(1+\xi_B)\left(e^{\frac{2+PR\ TD\ (\xi_A+\xi_B)}{PR\,\tau_D}}-1\right)} \\
& \times \Bigg[(1+\xi_A)e^{\frac{2+3PR\ TD(2+\xi_A+\xi_B)}{PR\,\tau_D}}\left(e^{\frac{TD(1+\xi_B)}{\tau_D}}+\xi_B\right) \\
& -(1+\xi_B)\Bigg((1+\xi_A)e^{\frac{TD(7+2\xi_A+3\xi_B)}{\tau_D}} \\
& +\xi_A e^{\frac{1+PR\ TD\ (6+2\xi_A+3\xi_B)}{PR\,\tau_D}}\left(e^{\frac{TD(1+\xi_A)}{\tau_D}}-1\right)\Bigg)\Bigg].
\end{aligned}
\tag{A.28}
$$

These solutions are identical to the ones that were obtained by integration.

B Hodgkin and Huxley Model

We developed a model of A1 using Hodgkin and Huxley type conductance-based models of regular-spiking (RS), pyramidal, excitatory neurons (Soto et al. 2006) and fast-spiking (FS) inhibitory interneurons (Pospischil et al. 2008). The membrane potential dynamics of one RS or FS neuron at point φ_X on the tonotopic axis of A1 follows

$$c_m \frac{dV_m}{dt} = -I_L - I_K - I_{Na} - I_{Th}(\varphi_X, t) - I_{C,*}(\varphi_X, t), \tag{B.1}$$

where $I_{C,*}(\varphi_X, t)$ stands for $I_{C,E}(\varphi_X, t)$ in exc. neurons and $I_{C,I}(\varphi_X, t)$ in inh. neurons, which represent the sum of cortical input (both exc. and inh.) to the respective neuron types (completely defined in equations (B.33) - (B.38)); $c_m = 1\mu\text{F/cm}^2$ is the membrane capacitance; $I_{Th}(\varphi, t)$ is the thalamocortical input current (see equation (B.25)); I_L is the leak current, I_K is the potassium current and I_{Na} is the sodium current. The latter three are given by

$$I_L = g_L(V_m - V_L), \tag{B.2}$$

$$I_{Na} = g_{Na} m^3 h (V_m - V_{Na}), \tag{B.3}$$

$$I_K = g_K n^4 (V_m - V_K). \tag{B.4}$$

Please note that m of equation (B.3), governing the activation of sodium ion channels, is entirely unrelated to the m of equation (A.1), describing the population firing rate in our rate model. There is no risk of confusion if this is kept in mind, as the latter m is the only one that appears in the discourse of the main text.
The values of conductances and reversal potentials in equations (B.2)-(B.4) are given in table B.1, and the gating variables h, m and n all satisfy the generic differential equation

$$\frac{dw}{dt} = \alpha_w(V)(1 - w) - \beta_w(V)w, \tag{B.5}$$

or, alternatively and equivalently,

$$\frac{dw}{dt} = \frac{w_\infty(V) - w}{\tau_w(V)}, \tag{B.6}$$

where

$$w_\infty(V) = (\alpha_w(V))/(\alpha_w(V) + \beta_w(V)), \tag{B.7}$$

$$\tau_w(V) = [\alpha_w(V) + \beta_w(V)]^{-1}. \tag{B.8}$$

For the regular-spiking pyramidal cell model, the specific $\alpha_w(V)$ and $\beta_w(V)$ functions are defined as

$$\alpha_h(V) = 0.128 \exp(-(50 + V)/18), \tag{B.9}$$

$$\beta_h(V) = \frac{4}{1 + \exp(-(V + 27)/5)}, \tag{B.10}$$

$$\alpha_n(V) = \frac{0.032(V + 52)}{1 - \exp(-(V + 52)/5)}, \tag{B.11}$$

$$\beta_n(V) = 0.5 \exp(-(57 + V)/40), \tag{B.12}$$

$$\alpha_m(V) = \frac{0.32(V + 50)}{1 - \exp(-(V + 50)/4)}, \tag{B.13}$$

$$\beta_m(V) = \frac{0.28(V + 27)}{\exp((V + 27)/5) - 1}. \tag{B.14}$$

In some simulations we also consider that the pyramidal cells have a non-inactivating K^+ current (Yamada, Koch & Adams 1989, Pospischil et al. 2008), which implements spike frequency adaptation (SFA). It is governed by

$$I_M = g_M n_2 (V_m - E_K) \tag{B.15}$$

$$\frac{dn_2}{dt} = \frac{n_{2,\infty}(V) - n_2}{\tau_{n_2}(V)}, \tag{B.16}$$

where

$$n_{2,\infty}(V) = \frac{1}{1 + \exp(-(V + 35)/10)} \tag{B.17}$$

$$\tau_{n_2}(V) = \frac{\tau_{\max}}{3.3 \exp((V + 35)/20) + \exp(-(V + 35)/20)}, \tag{B.18}$$

with $g_M = 0.35$ mS/cm^2 and $\tau_{\max} = 4$ s. The value of g_M was set so that, upon tonic stimulation, there is approximately a 35% reduction in firing rate (figure B.1; cf. figure 8*A* in Destexhe & Paré 1999). When this SFA current is in use, the term $-I_M$ is added to the r.h.s. of equation (B.1).

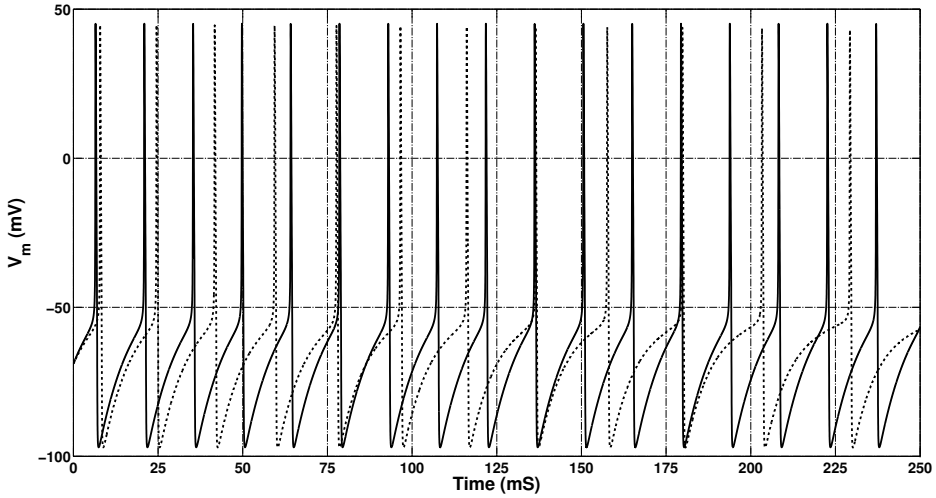


Figure B.1: Spike Frequency Adaptation (SFA) fit, based on Pospischil et al. (2008) equations and Destexhe & Paré (1999) data. $c_{TC} = 3\mu$A/cm^2, $g_M = 0.35$ mS/cm^2. Solid line: response without SFA. Dotted line: response with SFA.

As for the inhibitory cells, they are defined as follows. Pospischil et al. (2008) fit four types of cells to experimental data, among them RS excitatory and FS inhibitory cells. We make use of one of the FS cell fits

(Pospischil et al. 2008), defined by

$$\alpha_h(V) = 0.128 \exp(-(V - V_T - 17)/18) \tag{B.19}$$

$$\beta_h(V) = \frac{4}{1 + \exp(-(V - V_T - 40)/5)} \tag{B.20}$$

$$\alpha_n(V) = \frac{0.032(V - V_T - 15)}{1 - \exp(-(V - V_T - 15)/5)} \tag{B.21}$$

$$\beta_n(V) = 0.5 \exp(-(V - V_T - 10)/40) \tag{B.22}$$

$$\alpha_m(V) = \frac{0.32(V - V_T - 13)}{1 - \exp(-(V - V_T - 13)/4)} \tag{B.23}$$

$$\beta_m(V) = \frac{0.28(V - V_T - 40)}{\exp((V - V_T - 40)/5) - 1}, \tag{B.24}$$

where $V_T = -57.9$mV. This value is given in figure 4 of the paper (fitting data from somatosensory cortex), along with the other parameters that differ from the ones of the RS cell model we use. They are $V_{Na} = 58$ mV, $V_K = -90$ mV, $g_K = 3.9$ mS/cm^2 , $g_L = 0.038$ mS/cm^2 . Table B.1 gives an overview of the cellular biophysics parameter values that we fix throughout the work (unless stated otherwise).

Cell model	g_L	g_{Na}	g_K	V_L	V_{Na}	V_K	c_m
RS	0.1	100	80	-70	50	-100	1
FS	0.038	58	3.9	-70	50	-90	1

Table B.1: Values of biophysical parameters, fixed throughout this work (unless otherwise stated). Conductance values in mS/cm^2, potentials in mV and capacitance in μF/cm^2.

Returning to equation (B.1), the thalamocortical input current therein (at point φ_X of the cortical tonotopic axis and at instant t) is described by

$$I_{\mathrm{CTh}}(\varphi_X, t) = \int_{F_{low}}^{F_{high}} \sum_{i=1}^{K_n} I_{Th,i}(\varphi_X, F, t)\, dF, \tag{B.25}$$

where the limits of the integral are defined as $F_{low} = -L/2$ and $F_{high} = L/2$, L being the extension of the thalamic tonotopic axis, in octaves (relative to center of the axis; see figure B.2), and the sum runs over all

synapses originating at F on the aforementioned axis. $I_{Th,i}(\varphi, F, t)$ has two definitions, one for deterministic (det.) synapses and one for stochastic (stoch.) synapses (see section B.1):

$$I_{\mathrm{CTh,i}}(\varphi, F, t) = \begin{cases} g_{\mathrm{CTh}}(\varphi, F)\ d_i(F, t)\ s_i(F, t)\ (V_m - V_E), & \text{(det.)} \\ q_{\mathrm{CTh}}(\varphi, F)\ k, & \text{(stoch.)}, \end{cases} \tag{B.26}$$

where $d_i(F, t)$ denotes the synaptic depression variable at a specific thalamocortical (deterministic) synapse, whereas k is the total number of vesicles released from the functional contacts of (stochastic) synapse i (concepts defined in section B.1.2). The distributions of thalamocortical synaptic strengths (conductances and charges, respectively) are defined as

$$g_{\mathrm{CTh}}(\varphi, F) = g_{\mathrm{CTh}}^{\max}\ H(\varphi, F) \tag{B.27}$$

$$q_{\mathrm{CTh}}(\varphi, F) = q_{\mathrm{CTh}}^{\max}\ H(\varphi, F), \tag{B.28}$$

with $H(\varphi, F)$ modulating the strength of thalamocortical axonal projections across the cortical tonotopic axis. We consider three cases for this function, namely

$$H(\varphi, F) = \begin{cases} 1 & \forall \varphi, F \\ e^{-\frac{|\varphi - F|}{\nu}} & \\ e^{-\frac{(\varphi - F)^2}{2\nu^2}} & , \end{cases} \tag{B.29}$$

where ν is the fan-out or spread of axonal projections from thalamic 'neurons'. In the majority of simulations, we use $H(\varphi, F) \equiv 1$, since the fan-out or spread of thalamocortical axonal projections is already accounted for in the following fashion in our model.

We implement the spread in axonal projections (its value $\nu = 0.33$ octaves, based on data in Miller et al. 2001) in a quite concrete fashion, by having each thalamic (virtual) 'column' make K_n synapses onto each cortical 'column' within its reach, as per the value of ν. Thus, if we call our density of 'columns' per octave ρ (see values below), we get a spread in 'columns' of $\nu_C = \lfloor \rho\ \nu \rfloor$ (integer value due to discreteness of implementation). Consequently, each cortical site has a total of $K = K_n \times (2\ \nu_C + 1)$ afferent thalamocortical synapses, with the exception of those situated at a distance $\leq \nu_C$ from the extremes of the tonotopic axis, whose K value is less due to border effects. Naturally, each group of K_n synapses receives presynaptic spikes at the rate of firing of the thalamic 'neuron' on whose 'axon' the synaptic terminal is situated. We consider that each of the K_n synapses receives spikes at the same rate but at different exact times, since two different 'axons' may well belong to two distinct

thalamic 'neurons'. To this end, we maintain an independent Poisson process for each thalamocortical synapse.

About equations (B.25) and (B.26), furthermore, V_m is the membrane potential of the receiving cell, and the excitatory reversal potential is $V_E = 0$ mV, as we consider all these synapses to have α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA) type receptors. We approximate AMPA's fast rise time by an instantaneous rise of the synaptic gating variable, while the decay time constant has a value of $\tau_{AMPA} = 3$ ms, yielding the gating variable equation (used in equation (B.26))

$$\frac{ds_i(F,t)}{dt} = -\frac{s_i(F,t)}{\tau_{AMPA}} + \delta(t - t_{pre,F,i}), \tag{B.30}$$

in which $t_{pre,F,i}$ designates the time of arrival at synapse i of a presynaptic action potential via an efferent axon of site F of the MGBv.

We set the maximum thalamic input rate for pure tones to $R_{\mathrm{Th}}^{\max} = 100$ spikes per second, in accordance with literature (Creutzfeldt et al. 1980, Wallace et al. 2007). The synaptic conductance is set to $g_{\mathrm{CTh}}^{\max} = 0.02 \pm 0.004$ mS/cm^2, by fitting mean and s.d. of evoked mEPSPs (with depression) to values in Rose & Metherate (2005) (see details in section 3.2.1). We set $\sigma_{VL,E} = 1.6$ % and $\sigma_{VL,I} = 2.7$ %, also fit to that paper. The population, representing $L = 0/2/4/6$ octaves, consists of $N = 1/33/65/97$ 'columns' (for $N > 1$, density $\rho = 16$/octave) containing $M_E = 100/40/20/15$ exc. and $M_I = 100/40/20/15$ inh. neurons each. While M_I should be only about 15 - 25% (Meinecke & Peters 1987, Lampl & Okun 2009) of M_E, for simplicity we chose to give them the same value (we discuss the implications of this in chapter 5). For all values of N, if $\nu = 0$, each neuron receives input via $K = 100$ thalamocortical synapses, whereas for $\nu > 0$ this number varies with ρ, as discussed earlier. Each synapse receives input as an independent Poisson process of rate $R_{\mathrm{Th}}(F, f, t)$, as defined in equation B.41. The Poisson input and the non-zero $\sigma_{g_{\mathrm{CTh}}}$ and σ_{V_L} values all provide noise in our system, which proves important for some of our modelling solutions.

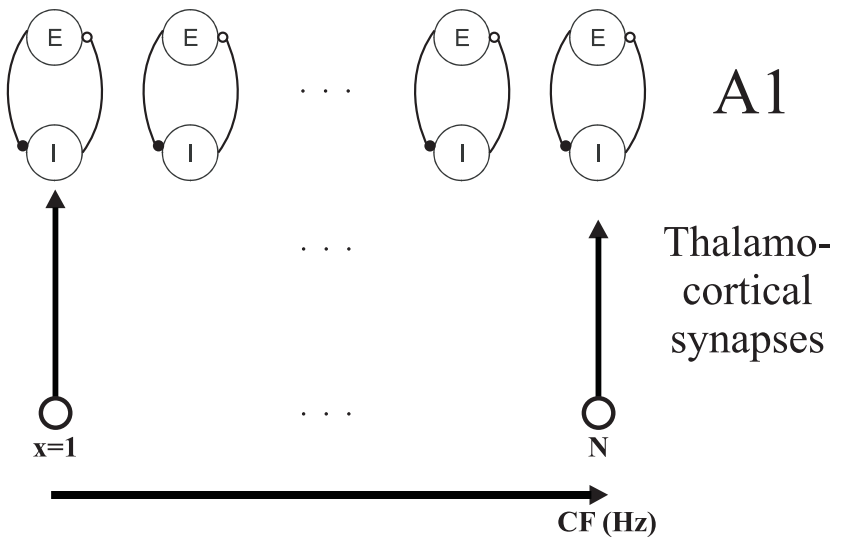


Figure B.2: Our primary auditory cortex (A1) model network consists of $N \times M$ excitatory and an equal number of inhibitory Hodgkin-Huxley type neurons, which form an array representing the tonotopic organization of A1 (N cortical 'columns' with distinct Characteristic Frequencies (CFs), each having M exc. and M inh. neurons). The array, of length L octaves, represents $\pm L/2$ octaves around a certain central frequency, f_0, at x_0. Each column is located at a certain position x in the tonotopic array and each of its constituent neurons has K thalamocortical synapses as well as intracortical synapses (see main text). The thalamic (MGBv) inputs are provided to each neuron by one independent Poisson spike train per thalamocortical synapse.

B.1. Thalamocortical synaptic depression model

We have employed two different models, one deterministic and one stochastic.

B.1.1. Deterministic depression model

We describe the mean thalamocortical synaptic depression value (fraction of vesicles available) at synapse i on an efferent axon from the MGBv

site F as

$$\frac{\partial d_i(F,t)}{\partial t} = \frac{d_{start} - d_i(F,t)}{\tau_D} - U\ R_{\mathrm{Th}}(F,f,t)\ d_i(F,t), \tag{B.31}$$

where U represents the utilization of synaptic vesicles, τ_D is the time constant of synapse recovery and d_{start} is the asymptotic value in the absence of presynaptic input. In order to set a baseline value for these parameters (as noted elsewhere, they vary greatly in the literature) we fit results in Rose & Metherate (2005) (described in section 3.2.1), which yielded $U = 0.4$ and $\tau_D = 250$ ms. These values are used in all single pure tone simulations, but are adjusted for simulating both forward masking with noise clicks and auditory streaming correlates with tones.

We emphasize that this model is mechanistically the same as that described in section A.1.2. Here however, our more detailed neuronal model and extended network require that we write $d_i(F,t)$ in equation (B.31), thus highlighting that depression varies independently across the tonotopic axis as well as across synapses at each tonotopic site.

B.1.2. Stochastic depression model

This model is implemented as in de la Rocha & Parga (2005), using vesicle turnover. Specifically, this model considers branched axons forming M_s so-called functional contacts (areas where exocytosis of synaptic vesicles containing neurotransmitter occurs) with a postsynaptic cell. Each of these contacts contains a releasable pool of N_s vesicles. Whenever a presynaptic spike is sent down an axon, each of the its functional contacts releases a synaptic vesicle with probability

$$p(n) = 1 - (1-U)^n, \tag{B.32}$$

where $n = 0,\ldots,N_s$. As is readily seen, when a pool contains a single vesicle ($n = 1$), this probability is U, a parameter of the model (on average, equal to the utilization parameter of the deterministic model – for that reason, and since the two different types of depression are never used simultaneously, we use the same symbol). Unless otherwise indicated, we use $N_s = 1$ and set $M_s = 7$, based on Gil et al. (1999) who find these numbers for the synapses of excitatory thalamic neurons projecting to the cortex. Again following de la Rocha & Parga (2005), we consider that, once a vesicle is released, the time it takes to replenish the pool with a vesicle be the first event from a Poisson process with homogeneous mean recovery time τ_D. We will refer to the total number of vesicles released across functional contacts, in response to a presynaptic spike, as n_{rel} (at instant t).

B.2. Intracortical connections

The model includes intracortical connectivity, which enables the study of the role of recurrent excitation and inhibition in shaping the cortical responses recorded in A1. We consider that inhibitory cells inhibit postsynaptic cells via $GABA_A$ synaptic currents and that excitatory cells excite postsynaptic cells via AMPA synaptic currents, only. The synaptic currents are modelled by

$$I_{C,E}(\varphi_X,t) = I_{C,EE}(\varphi_X,t) + I_{C,EI}(\varphi_X,t) \tag{B.33}$$

$$I_{C,I}(\varphi_X,t) = I_{C,II}(\varphi_X,t) + I_{C,IE}(\varphi_X,t) \tag{B.34}$$

$$I_{C,kj}(\varphi_X,t) = \int_{\varphi_{low}}^{\varphi_{high}} \sum_{i=1}^{M_j} g_{kj}(\varphi_X,\varphi)\ s_{kj,i}(\varphi,t)\ (V_{m,k}-V_j)\ d\varphi \tag{B.35}$$

$$g_{kj}(\varphi_X,\varphi) = g_{kj}\exp\left(\frac{-|\varphi_X-\varphi|}{\lambda_{kj}}\right) \tag{B.36}$$

$$\frac{ds_{*I,i}(\varphi,t)}{dt} = \alpha_I\left(1+\tanh\left(\frac{V_{m,I}}{4}\right)\right)(1-s_{*I,i}(\varphi,t)) - \beta_I\ s_{*I,i}(\varphi,t) \tag{B.37}$$

$$\frac{ds_{*E,i}(\varphi,t)}{dt} = \alpha_E\left(1+\tanh\left(\frac{V_{m,E}}{4}\right)\right)(1-s_{*E,i}(\varphi,t)) - \beta_E\ s_{*E,i}(\varphi,t), \tag{B.38}$$

where $k,j,* \in \{E,I\}$, with an index convention where EI means from inhibitory neurons to excitatory neurons; $\Delta\varphi \equiv |\varphi_X-\varphi|$ is the cortical tonotopic distance from the 'sending' to the 'receiving' neuron; $V_E = 0$ mV and $V_I = -80$ mV. For simplicity, $\lambda_E \equiv \lambda_{EE} = \lambda_{IE}$ and $\lambda_I \equiv \lambda_{II} = \lambda_{EI}$ (unless otherwise stated). Excitation was taken to be very narrow ($\lambda_E = 0.03$), as found by Liu et al. (2007) (but, for another possibility, see Kaur et al. 2004), whereas inhibition is quite broad ($\lambda_I = 0.4$), as can be deduced from data in Liu et al. (2007) and Wu et al. (2008) (for our reasoning on this, see section 3.2.1). α_E (α_I) is the rise factor and β_E (β_I) is the decay factor of excitatory (inhibitory) currents. We set $\alpha_E = 2.0$, $\beta_E = 0.5$, $\alpha_I = 0.5$ and $\beta_I = 0.3$, to give inhibitory currents a slower rise time and slower decay than excitatory currents. The conductance values g_{kj} are parameters of the model, whose values are given in the corresponding results sections.

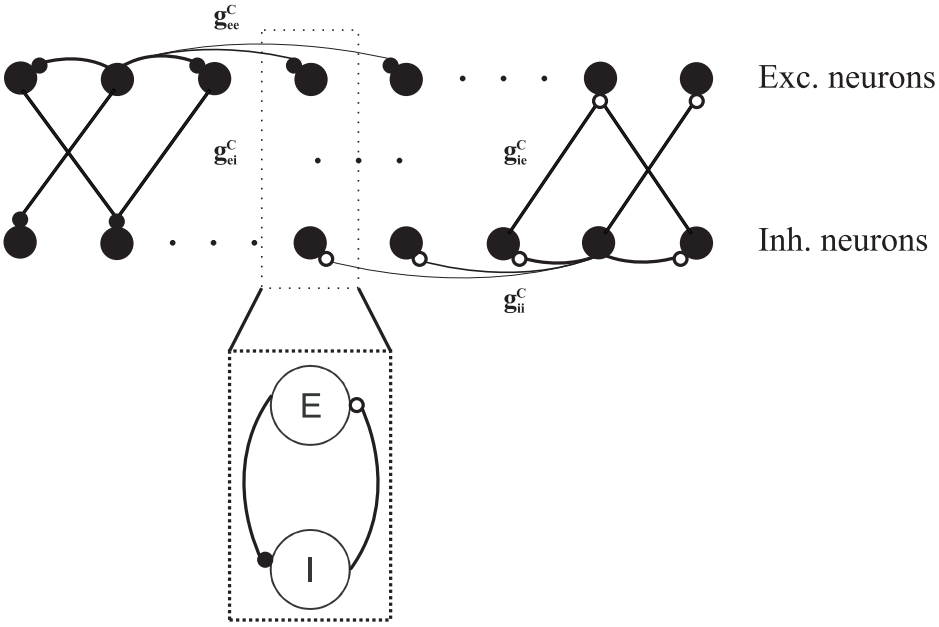


Figure B.3: Intracortical connectivity sketch. The neurons are coupled with synaptic strengths determined by equation (B.36), by setting the parameters g_{kj} and λ_{kj}. Note that, for clarity of presentation, on the left side of the figure, only excitatory connections are illustrated, while on the right only inhibitory connections are displayed. Omitted connections are possible.

B.3. Input function

We consider that the presentation of a pure tone of frequency f evokes an activity pattern in the ventral Medial Geniculate Body (MGBv) of the thalamus. The activity is centered at position $F = f$ of the thalamic tonotopic axis and its spatial extent is described by

$$R_{\mathrm{Th}}(F, f) = R_{\mathrm{Th}}^{\max} e^{-\frac{|F-f|}{\mu}}, \tag{B.39}$$

where $R_{\mathrm{Th}}^{\max}$ is a constant rate in units [1/time] and $\mu = 0.4$ octaves sets the (canonical) width of the pattern. We chose this width based on our fit of sub- and suprathreshold response data found by Liu et al. (2007) and Wu et al. (2008), primarily (see chapter 3). A recent study lends support to our use of an exponential shape, even though in a majority of MGBv neurons, it finds narrower widths than that we employ (see Bartlett et al. 2010).

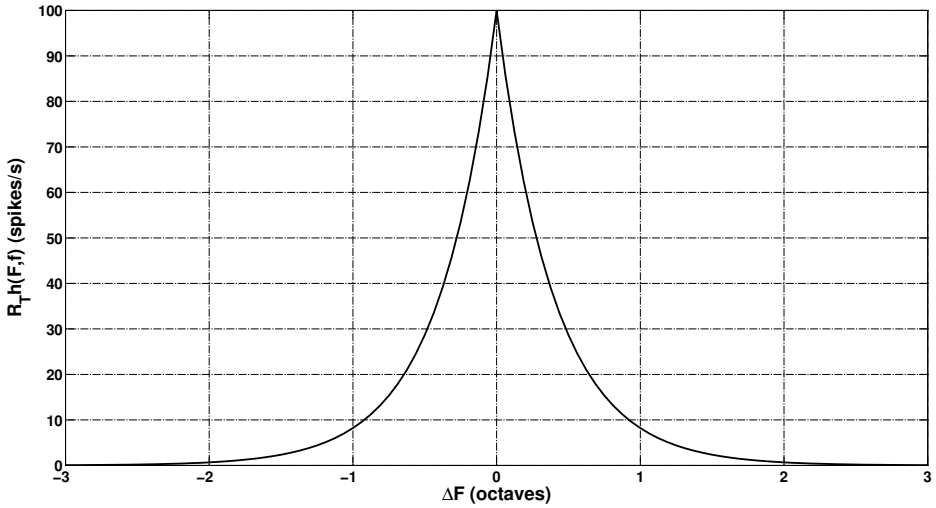


Figure B.4: Thalamic input rate $R_{\mathrm{Th}}(\Delta F)$, as a function of spectral distance from stimulus frequency f, $\Delta F \equiv F - f$, in octaves (equation (B.39)). Here, $\mu = 0.4$ oct.

We model both short tones (duration 25ms) and noise clicks (duration 5ms). Based on durations of responses to tones and clicks found in thalamus (Wallace et al. 2007, Wehr & Zador 2005), we consider the time course of both the noise click and the tone response to be governed by the function

$$\zeta(t) = \frac{\gamma}{\Sigma}\ t\ e^{-t/\Sigma\ +\ 1}, \tag{B.40}$$

where $\Sigma = 10$ ms for both the 5 ms noise clicks and the 25 ms tones. For noise, we use a modulation factor $\gamma = 2.0 \times 1.2 = 2.4$ (factor 2.0 reflects the higher amplitude of responses to a click as compared to a tone, found in Wallace et al. (2007), whereas the factor 1.2 reflects the difference in intensity of stimuli in Wehr & Zador (2005) (102 dB) and e.g. Wehr & Zador (2003) (60 dB)). For pure tones, $\gamma = 1.0$. The resulting product,

$$R_{\mathrm{Th}}(F, f, t) = R_{\mathrm{Th}}(F, f)\ \zeta(t), \tag{B.41}$$

is used as the rate in a non-homogenous Poisson process to generate spikes, subsequently fed to the model neurons through thalamocortical synapses (see equation (B.30)).

All differential equations were integrated numerically using the Heun method with step size 0.01 ms. The SIMD-oriented Fast Mersenne Twister

(SFMT) algorithm was used as random number generator for the external Poisson spikes as well as for generating spread around a mean parameter value. All (off-line) data analysis was carried out using MATLAB (MathWorks Inc.). Figures were made using MATLAB and Inkscape.

C Integrate and Fire model: Neural and synaptic dynamics

In this work, we use the mathematical formulation of the non-linear leaky integrate-and-fire neuron and its modelled synaptic input currents, as described in Brunel & Wang (2001). What follows is a presentation of the essentials in that formulation.

The evolution in time of the sub-threshold membrane potential $V(t)$ of a neuron is given by the equation:

$$C_m \frac{dV(t)}{dt} = -g_m(V(t) - V_L) - I_{syn}(t), \tag{C.1}$$

where C_m is the membrane capacitance, set to 0.5 nF for excitatory neurons and 0.2 nF for inhibitory neurons; g_m is the membrane leak conductance, with the values 25 nS for excitatory neurons and 20 nS for inhibitory neurons; V_L is the resting potential of -70 mV and I_{syn} the synaptic input current. The firing threshold is $V_{thr} = -50$ mV and the reset potential $V_{reset} = -55$ mV.

The synaptic current is the sum of recurrent excitatory currents ($I_{AMPA,rec}$ and $I_{NMDA,rec}$), an external excitatory current($I_{AMPA,ext}$) & an inhibitory current (I_{GABA}):

$$I_{syn}(t) = I_{AMPA,ext}(t) + I_{AMPA,rec}(t) + I_{NMDA,rec}(t) + I_{GABA}(t). \tag{C.2}$$

The currents are defined by:

$$I_{AMPA,ext}(t) = g_{AMPA,ext}(V(t) - V_E) \sum_{j=1}^{N_{ext}} s_j^{AMPA,ext}(t) \tag{C.3}$$

$$I_{AMPA,rec}(t) = g_{AMPA,rec}(V(t) - V_E) \sum_{j=1}^{N_E} w_{ji}^{AMPA} s_j^{AMPA,rec}(t) \tag{C.4}$$

$$I_{NMDA,rec}(t) = \frac{g_{NMDA}(V(t) - V_E)}{1 + [Mg^{++}]exp(-0.062V(t))/3.57} \times \sum_{j=1}^{N_E} w_{ji}^{NMDA} s_j^{NMDA}(t) \tag{C.5}$$

$$I_{GABA}(t) = g_{GABA}(V(t) - V_I) \sum_{j=1}^{N_I} w_{ji}^{GABA} s_j^{GABA}(t) \tag{C.6}$$

where $V_E = 0$ mV, $V_I = -70$ mV, w_j are the synaptic weights, s_j are the fractions of open channels for the different receptors and g_X designates the synaptic conductance for channel X. The NMDA synaptic current is dependent on the potential and controlled by the extracellular concentration of magnesium ($[\mathrm{M\,g}^{2+}] = 1$ mM). The values for the synaptic conductances for excitatory neurons are $g_{AMPA,ext} = 2.08$ nS, $g_{AMPA,rec} = 0.052$ nS, $g_{NMDA} = 0.1635$ nS and $g_{GABA} = 0.625$ nS and for inhibitory neurons $g_{AMPA,ext} = 1.62$ nS, $g_{AMPA,rec} = 0.0405$ nS, $g_{NMDA} = 0.129$ nS and $g_{GABA} = 0.4865$ nS. These values are obtained from the ones used in Brunel & Wang (2001) by correcting for the different number of neurons in our model. In their work the conductances were calculated so that in an unstructured network the excitatory neurons have a spontaneous spiking rate of 3 Hz and the inhibitory neurons a spontaneous rate of 9 Hz. The fractions of open channels are governed by:

$$\frac{ds_j^{AMPA,ext}(t)}{dt} = -\frac{s_j^{AMPA,ext}(t)}{\tau_{AMPA}} + \sum_k \delta(t - t_j^k) \tag{C.7}$$

$$\frac{ds_j^{AMPA,rec}(t)}{dt} = -\frac{s_j^{AMPA,rec}(t)}{\tau_{AMPA}} + \sum_k \delta(t - t_j^k) \tag{C.8}$$

$$\frac{ds_j^{NMDA}(t)}{dt} = -\frac{s_j^{NMDA}(t)}{\tau_{NMDA,decay}} + \alpha x_j(t)\left(1 - s_j^{NMDA}(t)\right) \tag{C.9}$$

$$\frac{dx_j(t)}{dt} = -\frac{x_j(t)}{\tau_{NMDA,rise}} + \sum_k \delta(t - t_j^k) \tag{C.10}$$

$$\frac{ds_j^{GABA}(t)}{dt} = -\frac{s_j^{GABA}(t)}{\tau_{GABA}} + \sum_k \delta(t - t_j^k), \tag{C.11}$$

where decay time constants are $\tau_{NMDA,decay} = 100$ ms for NMDA synapses, $\tau_{AMPA} = 2$ ms for AMPA synapses and $\tau_{GABA} = 10$ ms for GABA synapses. $\tau_{NMDA,rise} = 2$ ms is the rise time for NMDA synapses (the rise times for AMPA and GABA are in reality very short and are therefore neglected in the model) and $\alpha = 0.5$ ms^{-1}. Each sum over k represents a sum over spikes in the form of δ-peaks $\delta(t)$, emitted by presynaptic neuron j at time $t = t_j^k$.

The equations were in our implementation integrated numerically using a second order Runge-Kutta method with step size 0.02 ms. The Mersenne Twister algorithm was used as random number generator for the external Poisson spike trains and each of the typically 100 different trials for one specific parameter configuration differed in its random seed from the other trials.

D The mean-field formulation

The mean-field approximation used in the present work was derived in Brunel & Wang (2001). Its basic assumption is that the network of integrate-and-fire neurons has reached a stationary state. In this formulation the potential of a neuron is calculated as:

$$\tau_x \frac{dV(t)}{dt} = -V(t) + \mu_x + \sigma_x \sqrt{\tau_x}\eta(t) \tag{D.1}$$

where $V(t)$ is the membrane potential, and x labels the different populations. τ_x is the effective membrane time constant, μ_x is the mean value which the membrane potential would have in the absence of spiking and fluctuations, σ_x measures the magnitude of the fluctuations and η is a Gaussian process with an exponentially decaying correlation function and the time constant τ_{AMPA}. The quantities μ_x and σ_x^2 are given by:

$$\mu_x = \frac{(T_{ext}\nu_{ext} + T_{AMPA}n_x^{AMPA} + \rho_1 n_x^{NMDA})V_E + \rho_2 n_x^{NMDA}\langle V_x\rangle + T_I n_x^{GABA}V_I + V_L}{S_x} \tag{D.2}$$

$$\sigma_x^2 = \frac{g_{AMPA,ext}^2(\langle V_x\rangle - V_E)^2 N_{ext}\nu_{ext}\tau_{AMPA}^2\tau_x}{g_m^2\tau_m^2}. \tag{D.3}$$

where $\nu_{ext} = 3$ Hz $(+\lambda_{stim})$, ν_I is the spiking rate of the inhibitory pool, $\tau_m = C_m/g_m$ with the values for the excitatory or inhibitory neurons depending on the pool considered. The other quantities are given by:

$$S_x = 1 + T_{ext}\nu_{ext} + T_{AMPA}n_x^{AMPA} + (\rho_1 + \rho_2)n_x^{NMDA} + T_I n_x^{GABA} \tag{D.4}$$

$$\tau_x = \frac{C_m}{g_m S_x} \tag{D.5}$$

$$n_x^{AMPA} = \sum_{j=1}^{p} f_j w_{jx}^{AMPA} \nu_j \tag{D.6}$$

$$n_x^{NMDA} = \sum_{j=1}^{p} f_j w_{jx}^{NMDA} \psi(\nu_j) \tag{D.7}$$

$$n_x^{GABA} = \sum_{j=1}^{p} f_j w_{jx}^{GABA} \nu_j \tag{D.8}$$

$$\psi(\nu) = \frac{\nu\tau_{NMDA}}{1 + \nu\tau_{NMDA}} \left(1 + \frac{1}{1 + \nu\tau_{NMDA}} \sum_{n=1}^{\infty} \frac{(-\alpha\tau_{NMDA,rise})^n T_n(\nu)}{(n+1)!} \right) \tag{D.9}$$

$$T_n(\nu) = \sum_{k=0}^{n} (-1)^k \binom{n}{k} \frac{\tau_{NMDA,rise}(1 + \nu\tau_{NMDA})}{\tau_{NMDA,rise}(1 + \nu\tau_{NMDA}) + k\tau_{NMDA,decay}} \tag{D.10}$$

$$\tau_{NMDA} = \alpha\tau_{NMDA,rise}\tau_{NMDA,decay} \tag{D.11}$$

$$T_{ext} = \frac{g_{AMPA,ext}\tau_{AMPA}}{g_m} \tag{D.12}$$

$$T_{AMPA} = \frac{g_{AMPA,rec}N_E\tau_{AMPA}}{g_m} \tag{D.13}$$

$$\rho_1 = \frac{g_{NMDA}N_E}{g_m J} \tag{D.14}$$

$$\rho_2 = \beta \frac{g_{NMDA}N_E(\langle V_x \rangle - V_E)(J-1)}{g_m J^2} \tag{D.15}$$

$$J = 1 + \gamma \exp(-\beta\langle V_x \rangle) \tag{D.16}$$

$$T_I = \frac{g_{GABA}N_I\tau_{GABA}}{g_m} \tag{D.17}$$

$$\langle V_x \rangle = \mu_x - (V_{thr} - V_{reset})\nu_x \tau_x, \tag{D.18}$$

where p is the number of excitatory pools, f_x the fraction of neurons in the excitatory pool x, $w_{j,x}$ the weight of the connections from pool x to pool j, ν_x the spiking rate of the excitatory pool x, $\gamma = [Mg^{2+}]/3.57$ and $\beta = 0.062$.

The spiking rate of a pool as a function of the defined quantities can then be described by:

$$\nu_x = \phi(\mu_x, \sigma_x), \tag{D.19}$$

where

$$\phi(\mu_x, \sigma_x) = \left(\tau_{rp} + \tau_x \int_{\beta(\mu_x,\sigma_x)}^{\alpha(\mu_x,\sigma_x)} du\sqrt{\pi}\exp(u^2)[1 + \mathrm{erf}(u)] \right)^{-1} \tag{D.20}$$

$$\alpha(\mu_x, \sigma_x) = \frac{(V_{thr} - \mu_x)}{\sigma_x}\left(1 + 0.5\frac{\tau_{AMPA}}{\tau_x}\right) + 1.03\sqrt{\frac{\tau_{AMPA}}{\tau_x}} - 0.5\frac{\tau_{AMPA}}{\tau_x} \tag{D.21}$$

$$\beta(\mu_x, \sigma_x) = \frac{(V_{reset} - \mu_x)}{\sigma_x} \tag{D.22}$$

$\mathrm{erf}(u)$ is the error function and τ_{rp} is the refractory period, which is considered to be 2 ms for excitatory neurons and 1 ms for inhibitory neurons. To solve the equations defined by (D.19) for all x, we numerically integrate equation (D.18) and the differential equation below (D.23), whose fixed point solutions correspond to solutions to the equations (D.19):

$$\tau_x \frac{d\nu_x}{dt} = -\nu_x + \phi(\mu_x, \sigma_x). \tag{D.23}$$

The equations were integrated using the Euler method with step size 0.2 and 4000 iterations, enough for sufficient convergence to be attained.

Bibliography

Abbott, L. F. (2008), 'Theoretical neuroscience rising', *Neuron* **60**, 489–495.

Abbott, L. F., Varela, A., Sen, K. & Nelson, S. B. (1997), 'Synaptic depression and cortical gain control', *Science* **275**, 220–224.

Abeles, M. (1991), *Corticonics*, Cambridge University Press, New York, NY, USA.

Aitkin, L. M. (1973), 'Medial geniculate body of the cat: Responses to tonal stimuli of neurons in medial division', *Journal of Neurophysiology* **36(2)**, 275–283.

Anstis, S. & Saida, S. (1985), 'Adaptation to auditory streaming of frequency-modulated tones', *Journal of Experimental Psychology* **11**, 257–271.

Armstrong-James, M., Welker, E. & Callahan, C. A. (1993), 'The contribution of NMDA and non-NMDA receptors to fast and slow transmission of sensory information in the rat SI barrel cortex', *The Journal of Neuroscience* **13**, 2149–2160.

Asari, H. & Zador, A. M. (2009), 'Long-lasting context dependence constrains neural encoding models in rodent auditory cortex', *Journal of Neurophysiology* **102**, 2638–2656.

Bao, S., Chang, E. F., Woods, J. & Merzenich, M. M. (2004), 'Temporal plasticity in the primary auditory cortex induced by operant perceptual learning', *Nature Neuroscience* **7**(9), 974–981.

Bartlett, E. L., Sadagopan, S. & Wang, X. (2010), 'Fine frequency tuning in monkey auditory cortex and thalamus', *Journal of Neurophysiology* **106**, 849–859.

Bee, M. A. & Klump, G. (2005), 'Auditory stream segregation in the songbird forebrain: effects of time intervals on responses to interleaved tone sequences', *Brain, Behavior and Evolution* **66**, 197–214.

Bee, M. A. & Klump, G. M. (2004), 'Primitive auditory stream segregation: a neurophysiological study in the songbird forebrain', *Journal of Neurophysiology* **92**, 1088–1104.

Belelli, D., Harrison, N. L., Maguire, J., Macdonald, R. L., Walker, M. C. & Cope, D. W. (2009), 'Extrasynaptic $GABA_A$ receptors: Form, pharmacology, and function', *Journal of Neuroscience* **29(41)**, 12757–12763.

Bidet-Caulet, A. & Bertrand, O. (2009), 'Neurophysiological mechanisms involved in auditory perceptual organization', *Frontiers in Neuroscience* **3**, 182–191.

Bosch, L., Costa, A. & Sebastián-Gallés, N. (2000), 'First and second language vowel perception in early bilinguals', *European Journal of Cognitive Psychology* **12**, 189–222.

Braitenberg, V. & Schütz, A. (1991), *Anatomy of the Cortex*, Springer Verlag, Berlin, Germany.

Bregman, A. & Campbell, J. (1971), 'Primary auditory stream segregation and perception of order in rapid sequences of tones', *Journal of Experimental Psychology* **89**, 244–249.

Bregman, A. S. (1978), 'Auditory streaming is cumulative', *J. Exp. Psychol.: Hum. Percept. Perf.* **4**, 380–387.

Bregman, A. S. (1990), *Auditory Scene Analysis*, MIT Press, Cambridge, MA, USA.

Bregman, A. S., Ahad, P. A., Crum, P. A. C. & O'Reilly, J. (2000), 'Effects of time intervals and tone durations on auditory stream segregation', *Perception & Psychophysics* **62**, 626–636.

Brosch, M. & Schreiner, C. (1997), 'Time course of forward masking tuning curves in cat primary auditory cortex", *Journal of Neurophysiology* **77**, 923–943.

Brown, D., Feng, J. & Feerick, S. (1999), 'Variability of firing of Hodgkin-Huxley and Fitzhugh-Nagumo neurons with stochastic synaptic input', *Physical Review Letters* **82**(23), 4731–4734.

Brunel, N. & Wang, X. (2001), 'Effects of neuromodulation in a cortical network model of object working memory dominated by recurrent inhibition', *Journal of Computational Neuroscience* **11**, 63–85.

Bruno, R. M. & Sakmann, B. (2006), 'Cortex is driven by weak but synchronously active thalamocortical synapses', *Science* **312**, 1622–1627.

Calford, M. B. (1983), 'The parcellation of the medial geniculate body of the cat defined by the response properties of single units', *Journal of Neuroscience* **3(11)**, 2350–2364.

Calford, M. B. & Semple, M. N. (1995), 'Monaural inhibition in cat auditory cortex', *Journal of Neurophysiology* **73**, 1876–1891.

Carandini, M., Heeger, D. J. & Senn, W. (2002), 'A synaptic explanation of suppression in visual cortex', *Journal of Neuroscience* **22**(22), 10053–10065.

Chung, S., Li, X. & Nelson, S. B. (2002), 'Short-term depression at thalamocortical synapses contributes to rapid adaptation of cortical sensory responses in vivo', *Neuron* **34**, 437–444.

Clopper, C. G. & Pisoni, D. B. (2005), Perception of dialect variation, *in* D. B. Pisoni & R. E. Remez, eds, 'The Handbook of Speech Perception', Blackwell Publishing, Oxford, UK, pp. 313–337.

Connine, C. M. (2004), 'It's not what you hear but how often you hear it: on the neglected role of phonological variant frequency in auditory word recognition', *Psychonomic Bulletin & Review* **11**, 1084–1089.

Crair, M. C. & Malenka, R. C. (1995), 'A critical period for long-term potentiation at thalamocortical synapses', *Nature* **375**, 325–328.

Creutzfeldt, O., Hellweg, F. C. & Schreiner, C. (1980), 'Thalamocortical transformation of responses to complex auditory stimuli', *Experimental Brain Research* **39**, 87–104.

Cruikshank, S. J., Lewis, T. J. & Connors, B. W. (2007), 'Synaptic basis for intense thalamocortical activation of feedforward inhibitory cells in neocortex', *Nature Neuroscience* **10**, 462–468.

Cruikshank, S. J., Rose, R. J. & Metherate, R. (2002), 'Auditory thalamocortical synaptic transmission in vitro', *Journal of Neurophysiology* **87**, 361–384.

Cusack, R., Deeks, J., Aikman, G. & Carlyon, R. P. (2004), 'Effects of location, frequency region, and time course of selective attention on auditory scene analysis', *Journal of Experimental Psychology* **30**, 643–656.

Cutler, A., Weber, A. & Otake, T. (2006), 'Asymmetric mapping from phonetic to lexical representations in second-language listening', *Journal of Phonetics* **34**, 269–24.

Daw, M. I., Ashby, M. C. & Isaac, J. T. R. (2007), 'Coordinated developmental recruitment of latent fast spiking interneurons in layer iv barrel cortex', *Nature Neuroscience* **10**, 453–461.

de la Rocha, J., Marchetti, C., Schiff, M. & Reyes, A. (2008), 'Linking the response properties of cells in auditory cortex with network architecture: Co-tuning vs. lateral inhibition', *Journal of Neuroscience* **28**, 9151–9163.

de la Rocha, J. & Parga, N. (2005), 'Short-term synaptic depression causes a non-monotonic response to correlated stimuli', *Journal of Neuroscience* **25**(37), 8416–8431.

de la Rocha, J. & Parga, N. (2008), 'Thalamocortical transformations of periodic stimuli: the effect of stimulus velocity and synaptic short-term depression in the vibrissa-barrel system', *Journal of Computational Neuroscience* **25**, 122–140.

Deco, G. & Rolls, E. T. (2004), 'Synaptic and spiking dynamics underlying reward reversal in the orbitofrontal cortex', *Cerebral Cortex* **15**, 15–30.

Deco, G. & Rolls, E. T. (2005*a*), 'Attention, short term memory, and action selection: a unifying theory', *Progress in Neurobiology* **76**, 236–256.

Deco, G. & Rolls, E. T. (2005*b*), 'Neurodynamics of biased competition and cooperation for attention: a model with spiking neurons', *Journal of Neurophysiology* **94**, 295–313.

Deco, G. & Rolls, E. T. (2006), 'Decision-making and Weber's law: a neurophysiological model', *European Journal of Neuroscience* **24**, 901–916.

Denham, S. (2001), Cortical synaptic depression and auditory perception, *in* S. Greenberg & M. Slaney, eds, 'Computational models of auditory function', NATO ASI Series, IOS Press, Amsterdam, pp. 281–296.

Denham, S. & Winkler, I. (2006), 'The role of predictive models in the formation of auditory streams', *Journal of Physiology Paris* **100**, 154–170.

Destexhe, A. & Paré, D. (1999), 'Impact of network activity on the integrative properties of neocortical pyramidal neurons in vivo', *Journal of Neurophysiology* **81**, 1531–1547.

Drew, P. J. & Abbott, L. F. (2003), 'Model of song selectivity and sequence generation in area HVc of the songbird', *Journal of Neurophysiology* **89**, 2697–2706.

Eggermont, J. J. (1999), 'The magnitude and phase of temporal modulation transfer functions in cat auditory cortex', *Journal of Neuroscience* **19**, 2780–2788.

Eggermont, J. J. (2002), 'Temporal modulation transfer functions in cat primary auditory cortex: Separating stimulus effects from neural mechanisms', *Journal of Neurophysiology* **87**, 305–321.

Eisner, F. & McQueen, J. M. (2006), 'Stability over time in lexically-guided auditory perceptual learning', *Conference paper, Spring meeting of the Experimental Psychology Society, Birmingham, UK, April 10-12* .

Elhilali, M., Fritz, J. B., Klein, D. J., Simon, J. Z. & Shamma, S. A. (2004), 'Dynamics of precise spike timing in primary auditory cortex', *Journal of Neuroscience* **24**, 1159–1172.

Elhilali, M., Ma, L., Micheyl, C., Oxenham, A. J. & Shamma, S. A. (2009), 'Temporal coherence in the perceptual organization and cortical representation of auditory scenes', *Neuron* **61**, 317–329.

Escabí, M. A. & Read, H. L. (2005), 'Neural mechanisms for spectral analysis in the auditory midbrain, thalamus, and cortex', *International Review of Neurobiology* **70**, 207–252.

Fechner, G. J. (1880), *Elemente der Psychophysik*, Breitkopf und Härtel, Leipzig, Germany.

Fishman, Y., Arezzo, J. & Steinschneider, M. (2004), 'Auditory stream segregation in monkey auditory cortex: Effects of frequency separation, presentation rate and tone duration', *Journal of the Acoustical Society of America* **116**, 1656–1670.

Fishman, Y. I. & Steinschneider, M. (2009), 'Temporally dynamic frequency tuning of population responses in monkey auditory cortex', *Hearing Research* **254**, 64–76.

Fishman, Y., Reser, D. H. & Steinschneider, M. (2001), 'Neural correlates of auditory stream segregation in primary auditory cortex of the awake monkey', *Hearing Research* **151**, 167–187.

Fox, K., Sato, H. & Daw, N. (1989), 'The location and function of NMDA receptors in cat and kitten visual cortex', *The Journal of Neuroscience* **9**, 2443–2454.

Fritz, J., Shamma, S., Elhilali, M. & Klein, D. (2003), 'Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex', *Nature Neuroscience* **6**(11), 1216–1223.

Froemke, R. C., Merzenich, M. M. & Schreiner, C. E. (2007), 'A synaptic memory trace for cortical receptive field plasticity', *Nature* **450**, 425–429.

Gabernet, L., Jaddhav, S. P., Feldman, D. E., Carandini, M. & Scanziani, M. (2005), 'Somatosensory integration controlled by dynamic thalamocortical feed-forard inhibition', *Neuron* **48**, 315–327.

Garabedian, C. E., Jones, S. R., Merzenich, M. M., Dale, A. & Moore, C. I. (2003), 'Band-pass response properties of rat SI neurons', *Journal of Neurophysiology* **90**, 1379–1391.

Gil, Z. & Amitai, Y. (1996), 'Adult thalamocortical transmission involves both NMDA and non-NMDA receptors', *Journal of Neurophysiology* **76(4)**, 2547–2554.

Gil, Z., Connors, B. W. & Amitai, Y. (1997), 'Differential regulation of neocortical synapses by neuromodulators and activity', *Neuron* **19(3)**, 679–686.

Gil, Z., Connors, B. W. & Amitai, Y. (1999), 'Efficacy of thalamocortical and intracortical synaptic connections: Quanta, innervation, and reliability', *Neuron* **23**, 385–397.

Gittelman, J. X. & Pollak, G. D. (2011), 'It's about time: How input timing is used and not used to create emergent properties in the auditory system', *Journal of Neuroscience* **31**(7), 2576–2583.

Happel, M. F. K., Jeschke, M. & Ohl, F. W. (2010), 'Spectral integration in primary auditory cortex attributable to temporally precise convergence of thalamocortical and intracortical input', *Journal of Neuroscience* **30(33)**, 11114–11127.

Hebb, D. (1949), *The organization of behavior - A neuropsychological theory*, John Wiley, New York, NY, USA.

Hestrin, S., Sah, P. & Nicoll, R. (1990), 'Mechanisms generating the time course of dual component excitatory synaptic currents recorded in hippocampal slices', *Neuron* **5**, 247–253.

Higley, M. J. & Contreras, D. (2006), 'Balanced excitation and inhibition determine spike timing during frequency adaptation', *Journal of Neuroscience* **26(2)**, 448–457.

Hodgkin, A. L. & Huxley, A. F. (1952), 'A quantitative description of membrane current and its application to conduction and excitation in nerve', *Journal of Neurophysiology* **117**, 500–544.

Holmgren, C., Harkany, T., Svennenfors, B. & Zilberter, Y. (2003), 'Pyramidal cell communication within local networks in layer 2/3 of rat neocortex', *Journal of Physiology* **551.1**, 139–153.

Houde, J. F. & Jordan, M. I. (1998), 'Sensorimotor adaptation in speech production', *Science* **279**, 1213–1216.

Huang, E. P. & Stevens, C. F. (1997), 'Estimating the distribution of synaptic reliabilities', *Journal of Neurophysiology* **78**, 2870–2880.

Hull, C., Isaacson, J. S. & Scanziani, M. (2009), 'Postsynaptic mechanisms govern the differential excitation of cortical neurons by thalamic inputs', *The Journal of Neuroscience* **29**(28), 9127–9136.

Isaacson, J. S. & Scanziani, M. (2011), 'How inhibition shapes cortical activity', *Neuron* **72**, 231–243.

Jahr, C. & Stevens, C. (1990), 'Voltage dependence of NMDA-activated macroscopic conductances predicted by single-channel kinetics', *Journal of Neuroscience* **10**, 3178–3182.

Kanwal, J. S., Medvedev, A. V. & Micheyl, C. (2003), 'Neurodynamics for auditory stream segregation: tracking sounds in the mustached bat's natural environment', *Network: Computation in Neural Systems* **14**, 413–435.

Kaur, S., Lazar, R. & Metherate, R. (2004), 'Intracortical pathways determine breadth of subthreshold frequency receptive fields in primary auditory cortex', *Journal of Neurophysiology* **91**, 2551–2567.

Kidd, G. & Feth, L. L. (1982), 'Effects of masker duration in pure-tone forward masking', *J. Acoust. Soc. Am.* **72**, 1384–1386.

Kilgard, M. P. & Merzenich, M. M. (1998*a*), 'Cortical map reorganization enabled by Nucleus Basalis activity', *Science* **1**, 1714–1718.

Kilgard, M. P. & Merzenich, M. M. (1998*b*), 'Plasticity of temporal information processing in the primary auditory cortex', *Nature Neuroscience* **1**, 727–732.

Kilgard, M. P. & Merzenich, M. M. (1999), 'Distributed representation of spectral and temporal information in rat primary auditory cortex', *Hearing Research* **134**, 16–28.

Kilgard, M. P., Pandya, P. K., Vázquez, J., Gehi, A., Schreiner, C. E. & Merzenich, M. M. (2001), 'Sensory input directs spatial and temporal plasticity in primary auditory cortex', *Journal of Neurophysiology* **86**(1), 326–338.

Koch, K. W. & Fuster, J. M. (1989), 'Unit activity in monkey parietal cortex related to haptic perception and temporary memory', *Exp. Brain Res.* **76**, 292–306.

Kowalski, N., Depireux, D. A. & Shamma, S. A. (1996), 'Analysis of dynamic spectra in ferret primary auditory cortex. I. Characteristics of single-unit responses to moving ripple spectra.', *Journal of Neurophysiology* **76**, 3503–3523.

Kraljic, T. & Samuel, A. G. (2005), 'Perceptual learning for speech: Is there a return to normal?', *Cognitive Psychology* **51**, 141–178.

Krukowski, A. E. & Miller, K. D. (2001), 'Thalamocortical NMDA conductances and intracortical inhibition can explain cortical temporal tuning', *Nature Neuroscience* **4**(4), 424–430.

Lampl, M. & Okun, I. (2009), 'Balance of excitation and inhibition', *Scholarpedia* **4**(8), 7467.

Larsson, J. P., Montbrió, E. & Deco, G. (2011), 'A biophysically detailed model of the primary auditory cortex explains physiological forward

masking, co-tuning of excitation and inhibition and cortical signal amplification', *BMC Neuroscience* **12**(Suppl 1), P66.

Larsson, J. P., Vera-Constán, F., Sebastián-Gallés, N. & Deco, G. (2008), 'Lexical plasticity in early bilinguals does not alter phoneme categories: I. neurodynamical modeling', *Journal of Cognitive Neuroscience* **20(1)**, 76–94.

Levy, R. B. & Reyes, A. D. (2011), 'Coexistence of lateral and co-tuned inhibitory configurations in cortical networks', *PLOS Computational Biology* **7**(10), 1–14.

Levy, R. B. & Reyes, A. D. (2012), 'Spatial profile of excitatory and inhibitory synaptic connectivity in mouse primary auditory cortex', *Journal of Neuroscience* **32**(16), 5609–5619.

Liu, B., Li, Y., Ma, W., Pan, C., Zhang, L. I. & Tao, H. W. (2011), 'Broad inhibition sharpens orientation selectivity by expanding input dynamic range in mouse simple cells', *Neuron* **71**, 542–554.

Liu, B., Wu, G. K., Arbuckle, R., Tao, H. W. & Zhang, L. I. (2007), 'Defining cortical frequency tuning with recurrent excitatory circuitry', *Nature Neuroscience* **10**, 1594–1600.

Loebel, A., Nelken, I. & Tsodyks, M. (2007), 'Processing of sounds by population spikes in a model of primary auditory cortex', *Frontiers in Neuroscience* **1**, 197–209.

Loebel, A. & Tsodyks, M. (2002), 'Computation by ensemble synchronization in recurrent networks with synaptic depression', *Journal of Computaional Neuroscience* **13**, 111–124.

Ma, X. & Suga, N. (2009), 'Specific and nonspecific plasticity of the primary auditory cortex elicited by thalamic auditory neurons', *Journal of Neuroscience* **29(15)**, 4888–4896.

Machens, C. K., Romo, R. & Brody, C. D. (2005), 'Flexible control of mutual inhibition: A neural model of two-interval discrimination', *Science* **307**, 1121–1124.

Macmillan, N. & Creelman, C. (1991), *Detection Theory: A user's guide*, Cambridge University Press, Cambridge.

Markram, H., Lübke, J., Frotscher, M. & Sakmann, B. (1997), 'Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs', *Science* **275**, 213–215.

Matsuo, N., Reijmers, L. & Mayford, M. (2008), 'Spine-type-specific recruitment of newly synthesized AMPA receptors with learning', *Science* **319**, 1104–1107.

McAdams, S. & Bregman, A. S. (1979), 'Hearing musical streams', *Computational Music Journal* **3**, 26–63.

McClelland, J. L. & Elman, J. L. (1986), 'The TRACE model of speech perception', *Cognitive Psychology* **18**, 1–86.

McCormick, D., Connors, B., Lighthall, J. & Prince, D. A. (1985), 'Comparative electrophysiology of pyramidal and sparsely spiny stellate neurons in the neocortex', *Journal of Neurophysiology* **54**, 782–806.

Meinecke, D. L. & Peters, A. (1987), 'GABA immunoreactive neurons in rat visual cortex', *Journal of Comparative Neurology* **261**, 388–404.

Micheyl, C., Carlyon, R. P., Gutschalk, A., Melcher, J. R., Oxenham, A., Rauschecker, J., Tian, B. & Wilson, E. C. (2007), 'The role of auditory cortex in the formation of auditory streams', *Hearing Research* **229**, 116–131.

Micheyl, C., Tian, B., Carlyon, R. P. & Rauschecker, J. P. (2005), 'Perceptual organization of tone sequences in the auditory cortex of awake macaques', *Neuron* **48**, 139–148.

Miller, G. A. (1947), 'The masking of speech', *Psychological Bulletin* **44**(2), 105–129.

Miller, G. A. & Heise, G. A. (1950), 'The trill threshold', *J. Acoust. Soc. Am.* **22**, 637–638.

Miller, K. D., Chapman, B. & Stryker, M. P. (1989), 'Visual responses in adult cat visual cortex depend on N-methyl-D-aspartate receptors', *Proc. Natl. Acad. Sci. USA* **86**, 5183–5187.

Miller, L. M., Escabí, M. A., Read, H. L. & Schreiner, C. E. (2001), 'Functional convergence of response properties in the auditory thalamocortical system', *Neuron* **32**, 151–160.

Miller, L. M., Escabí, M. A., Read, H. L. & Schreiner, C. E. (2002), 'Spectrotemporal receptive fields in the lemniscal auditory thalamus and cortex', *Journal of Neurophysiology* **87**, 516–527.

Miller, P. & Wang, X.-J. (2002), 'Inhibitory control by an integral feedback signal in prefrontal cortex: A model of discrimination between sequential stimuli', *Proc. Natl. Acad. Sci. USA* **36**, 955–968.

Montbrió, E., Almeida, R., Larsson, J. P. & Deco, G. (2012), 'Synaptic mechanism of dynamic frequency selectivity in the primary auditory cortex', *In preparation* .

Montbrió, E., Larsson, J. P., Almeida, R. & Deco, G. (2009), 'A model of the primary auditory cortex response to sequences of pure tones', *BMC Neuroscience* **10**(Suppl 1), P151.

Moore, B. C. J. & Gockel, H. (2002), 'Factors influencing sequential stream segregation', *Acta Acustica* **88**, 320–333.

Nahmani, M. & Erisir, A. (2005), 'VGluT2 immunochemistry identifies thalamocortical terminals in layer 4 of adult and developing visual cortex', *Journal of Comparative Neurology* **484**, 458–473.

Nicoll, R. A., Eccles, J. C., Oshima, T. & Rubia, F. (1975), 'Prolongation of hippocampal inhibitory postsynaptic potentials by barbiturates', *Nature* **258**, 625–627.

Noreña, A. J., Gourévitch, B., Aizawa, N. & Eggermont, J. J. (2006), 'Spectrally enhanced acoustic environment disrupts frequency representation in cat auditory cortex', *Nature Neuroscience* **9**(7), 932–939.

Norris, D. G., McQueen, J. M. & Cutler, A. (2000), 'Merging information in speech recognition: Feedback is never necessary', *Behavioural and brain sciences* **23**, 299–370.

Norris, D. G., McQueen, J. M. & Cutler, A. (2003), 'Perceptual learning in speech', *Cognitive Psychology* **47**, 204–238.

Obleser, J., Lahiri, A. & Eulitz, C. (2004), 'Magnetic brain response mirrors extraction of phonological features from spoken vowels', *Journal of Cognitive Neuroscience* **16**, 31–39.

O'Mard, L. P. & Meddis, R. (1997), 'Computer exploration of the auditory system with LUTEar', *Br. J. Audiol.* **31**, 125–126.

Oswald, A.-M. M., Schiff, M. L. & Reyes, A. D. (2006), 'Synaptic mechanisms underlying auditory processing', *Current Opinion in Neurobiology* **16**, 371–376.

Pallier, C., Bosch, L. & Sebastián-Gallés, N. (1997), 'A limit on behavioral plasticity in speech perception', *Cognition* **64**, B9–B17.

Pallier, C., Colomé, A. & Sebastián-Gallés, N. (2001), 'The influence of native-language phonology on lexical access: exemplar-based vs. abstract lexical entries', *Psychological Science* **12**, 445–449.

Pickles, J. O. (2008), *An introduction to the physiology of hearing*, Elsevier, Amsterdam.

Porter, J. T. & Nieves, D. (2004), 'Presynaptic $GABA_B$ receptors modulate thalamic excitation of inhibitory and excitatory neurons in the mouse barrel cortex', *Journal of Neurophysiology* **92**, 2762–2770.

Pospischil, M., Toledo-Rodriguez, M., Monier, C., Piwkowska, Z., Bal, T., Frégnac, Y., Markram, H. & Destexhe, A. (2008), 'Minimal hodgkin-huxley type models for different classes of cortical and thalamic neurons', *Biological Cybernetics* **99**, 427–441.

Pressnitzer, D., Sayles, M., Micheyl, C. & Winter, M. (2008), 'Perceptual organization of sound begins in the auditory periphery', *Current Biology* **18**, 1124–1128.

Richardson, R. J., Blundon, J. A., Bayazitov, I. T. & Zakharenko, S. S. (2009), 'Connectivity patterns revealed by mapping of active inputs on dendrites of thalamorecipient neurons in the auditory cortex', *Journal of Neuroscience* **29**(20), 6406–6417.

Rolls, E. T. & Deco, G. (2002), *Computational Neuroscience of Vision*, Oxford University Press, Oxford, UK.

Rolls, E. T. & Treves, A. (1998), *Neural networks and brain function*, Oxford University Press, Oxford, UK.

Rose, H. J. & Metherate, R. (2005), 'Auditory thalamocortical transmission is reliable and temporally precise', *Journal of Neurophysiology* **94**, 2019–2030.

Sadagopan, S. & Wang, X. (2010), 'Contribution of inhibition to stimulus selectivity in primary auditory cortex of awake primates', *Journal of Neuroscience* **30(21)**, 7314–7325.

Saeb, S., Gharibzadeh, S., Towhidkhah, F. & Farajidavar, A. (2007), 'Modeling the primary auditory cortex using dynamic synapses: Cand synaptic plasticity explain the temporal tuning?', *Journal of Theoretical Biology* **248**, 1–9.

Salin, P. & Prince, D. A. (1996), 'Spontaneous $GABA_A$ receptor-mediated inhibitory currents in adult rat somatosensory cortex', *Journal of Neurophysiology* **75**, 1573–1588.

Sally, S. L. & Kelly, J. B. (1988), 'Organization of auditory cortex in the albino rat: Sound frequency', *Journal of Neurophysiology* **59(5)**, 1627–1638.

Schiff, M. L. & Reyes, A. D. (2012), 'Characterization of thalamocortical responses of regular-spiking and fast-spiking neurons of the mouse auditory cortex in vitro and in silico', *The Journal of Neurophysiology* **107**, 1476–1488.

Sebastián-Gallés, N., Echeverría, S. & Bosch, L. (2005), 'The influence of initial exposure on lexical representation: Comparing early and simultaneous bilinguals', *Journal of Memory and Language* **52**, 240–255.

Sebastián-Gallés, N., Rodríguez-Fornells, A., Diego-Balaguer, R. & Diaz, B. (2006), 'First- and second-language phonological representations in the mental lexicon', *Journal of Cognitive Neuroscience* **18**, 1277–1291.

Sebastián-Gallés, N. & Soto-Faraco, S. (1999), 'Online processing of native and non-native phonemic contrasts in early bilinguals', *Cognition* **72**, 112–123.

Sebastián-Gallés, N., Vera-Constán, F., Larsson, J. P., Costa, A. & Deco, G. (2009), 'Lexical plasticity in early bilinguals does not alter phoneme categories: II. Experimental evidence', *Journal of Cognitive Neuroscience* **21**, 2343–2357.

Sejnowski, T. J., Koch, C. F. & Churchland, P. S. (1988), 'Computational neuroscience', *Science* **241**, 1299–1306.

Shadlen, M. N. & Newsome, W. T. (1998), 'The variable discharge of cortical neurons: Implications for connectivity, computation and information coding', *Journal of Neuroscience* **18**, 3870–3896.

Shamma, S. A., Elhilali, M. & Micheyl, C. (2011), 'Temporal coherence and attention in auditory scene analysis', *Trends in Neurosciences* **34**(3), 114–123.

Shamma, S. & Micheyl, C. (2010), 'Behind the scenes of auditory perception', *Current Opinion in Neurobiology* **20**, 1–6.

Siok, W. T., Jin, Z., Fletcher, P. & Tan, L. H. (2003), 'Distinct brain regions associated with syllable and phoneme', *Human Brain Mapping* **18**, 201–207.

Smith, P. H. & Populin, L. C. (2001), 'Fundamental differences between the thalamocortical recipient layers of the cat auditory and visual cortices', *The Journal of Comparative Neurology* **436**, 508–519.

Snyder, J. S. & Alain, C. (2007), 'Toward a neurophysiological theory of auditory stream segregation', *Psychological Bulletin* **133**, 780–799.

Soto, G., Kopell, N. & Sen, K. (2006), 'Network architecture, receptive fields and neuromodulation: Computational and functional implications of cholinergic modulation in primary auditory cortex', *Journal of Neurophysiology* **96**, 2972–2983.

Spruston, N., Jonas, P. & Sakmann, B. (1995), 'Dendritic glutamate receptor channel in rat hippocampal CA3 and CA1 pyramidal neurons', *Journal of Physiology* **482**, 325–352.

Stiebler, I., Neulist, R., Fichtel, I. & Ehret, G. (1997), 'The auditory cortex of the house mouse: Left-right differences, tonotopic organization and quantitative analysis of frequency representation', *J Comp Physiol A* **181**, 559–571.

Tan, A. Y. Y. & Wehr, M. (2009), 'Balanced tone-evoked synaptic excitation and inhibition in mouse auditory cortex', *Neuroscience* **163**, 1302–1315.

Tan, A. Y. Y., Zhang, L. I., Merzenich, M. M. & Schreiner, C. E. (2004), 'Tone-evoked excitatory and inhibitory synaptic conductances of primary auditory cortex neurons', *Journal of Neurophysiology* **92**, 630–643.

Thomson, A. M. & Deuchars, J. (1994), 'Temporal and spatial properties of local circuits in neocortex', *Trends in Neurosciences* **17**, 119–126.

Tsodyks, M., Uziel, A. & Markram, H. (2000), 'Synchrony generation in recurrent networks with frequency-dependent synapses', *Journal of Neuroscience* **20**, 1–5.

Tsodyks, M. V. & Markram, H. (1997), 'The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability', *Proc. Natl. Acad. Sci. USA* **94**, 719–723.

Tsodyks, M. V., Pawelzik, K. & Markram, H. (1998), 'Neural networks with dynamic synapses', *Neural Computation* **10**, 821–835.

Tuckwell, H. (1988), *Introduction to theoretical neurobiology*, Cambridge University Press, Cambridge, UK.

Ulanovsky, N., Las, L., & Nelken, I. (2003), 'Processing of low-probability sounds by cortical neurons', *Nature Neuroscience* **6**(4), 391–398.

Ulanovsky, N., Las, L., Farkas, D. & Nelken, I. (2004), 'Multiple time scales of adaptation in auditory cortex neurons', *Journal of Neuroscience* **24**(46), 10440–10453.

van Noorden, L. P. A. S. (1975), Temporal coherence in the perception of tone sequences, PhD thesis, Eindhoven Technische Hogeschool, Eindhoven, the Netherlands.

Viaene, A. N., Petrof, I. & Sherman, M. (2011), 'Synaptic properties of thalamic input to layers 2/3 and 4 of primary somatosensory and auditory cortices', *Journal of Neurophysiology* **105**, 279–292.

Vliegen, J. & Oxenham, A. J. (1999), 'Sequential stream segregation in the absence of spectral cues', *J. Acoust. Soc. Am.* **105**(1), 339–346.

Volkov, I. & Galazjuk, A. (1991), 'Formation of spike response to sound tones in cat auditory cortex neurons: Interaction of excitatory and inhibitory effects', *Neuroscience* **43**(2-3), 307–321.

Wallace, M. N., Anderson, L. A. & Palmer, A. R. (2007), 'Phase-locked responses to pure tones in the auditory thalamus', *Journal of Neurophysiology* **98**, 1941–1952.

Wang, H.-P., Spencer, D., Fellous, J.-M. & Sejnowski, T. J. (2010), 'Synchrony of thalamocortical inputs maximizes cortical reliability', *Science* **328**, 106–109.

Wang, X. (2007), 'A sharper view from the top', *Nature Neuroscience* **10**, 1509–1511.

Wang, X.-J. (2002), 'Probabilistic decision making by slow reverberation in cortical circuits', *Neuron* **36**, 955–968.

Wehr, M. & Zador, A. M. (2003), 'Balanced inhibition underlies tuning and sharpens spike timing in auditory cortex', *Nature* **426**, 442–446.

Wehr, M. & Zador, A. M. (2005), 'Synaptic mechanisms of forward suppression in rat auditory cortex', *Neuron* **47**, 437–445.

Wilson, F. A., Scalaidhe, S. P. & Goldman-Rakic, P. S. (1994), 'Functional synergism between putative gamma-aminobutyrate-containing neurons and pyramidal neurons in prefrontal cortex', *Proc. Natl. Acad. Sci. USA* **91**, 4009–4013.

Wilson, H. R. & Cowan, J. D. (1972), 'Excitatory and inhibitory interactions in localized populations of model neurons', *Biophysical Journal* **12**(1), 1–24.

Winer, J. A. & Lee, C. C. (2007), 'The distributed auditory cortex', *Hearing Research* **229**, 3–13.

Wong, K.-F. & Wang, X.-J. (2006), 'A recurrent network mechanism of time integration in perceptual decisions', *Journal of Neuroscience* **26**, 1314–1328.

Wu, G. K., Arbuckle, R., Liu, B., Tao, H. W. & Zhang, L. I. (2008), 'Lateral sharpening of cortical frequency tuning by approximately balanced inhibition', *Neuron* **58**, 132–143.

Xiang, Z., Huguenard, J. R. & Prince, D. A. (1998), '$GABA_A$ receptor mediated currents in interneurons and pyramidal cells of rat visual cortex', *Journal of Physiology* **506**, 715–730.

Yamada, W. M., Koch, C. & Adams, P. R. (1989), Multiple channels and calcium dynamics, *in* C. Koch & I. Segev, eds, 'Methods in neuronal modeling', MIT Press, pp. 97–134.

Yamauchi, T., Hori, T. & Takahashi, T. (2000), 'Presynaptic inhibition by muscimol through $GABA_B$ receptors', *European Journal of Neuroscience* **12**, 3433–3436.

Ye, C., Poo, M., Dan, Y. & Zhang, X. (2010), 'Synaptic mechanisms of direction selectivity in primary auditory cortex', *Journal of Neuroscience* **30**(5), 1861–1868.

Zhang, L. I., Tan, A. Y., Schreiner, C. E. & Merzenich, M. M. (2003), 'Topography and synaptic shaping of direction selectivity in primary auditory cortex', *Nature* **424**, 201–205.

Zhou, X. & Merzenich, M. M. (2008), 'Enduring effects of early structured noise exposure on temporal modulation in the primary auditory cortex', *Proc. Natl. Acad. Sci. USA* **105**, 4423–4428.